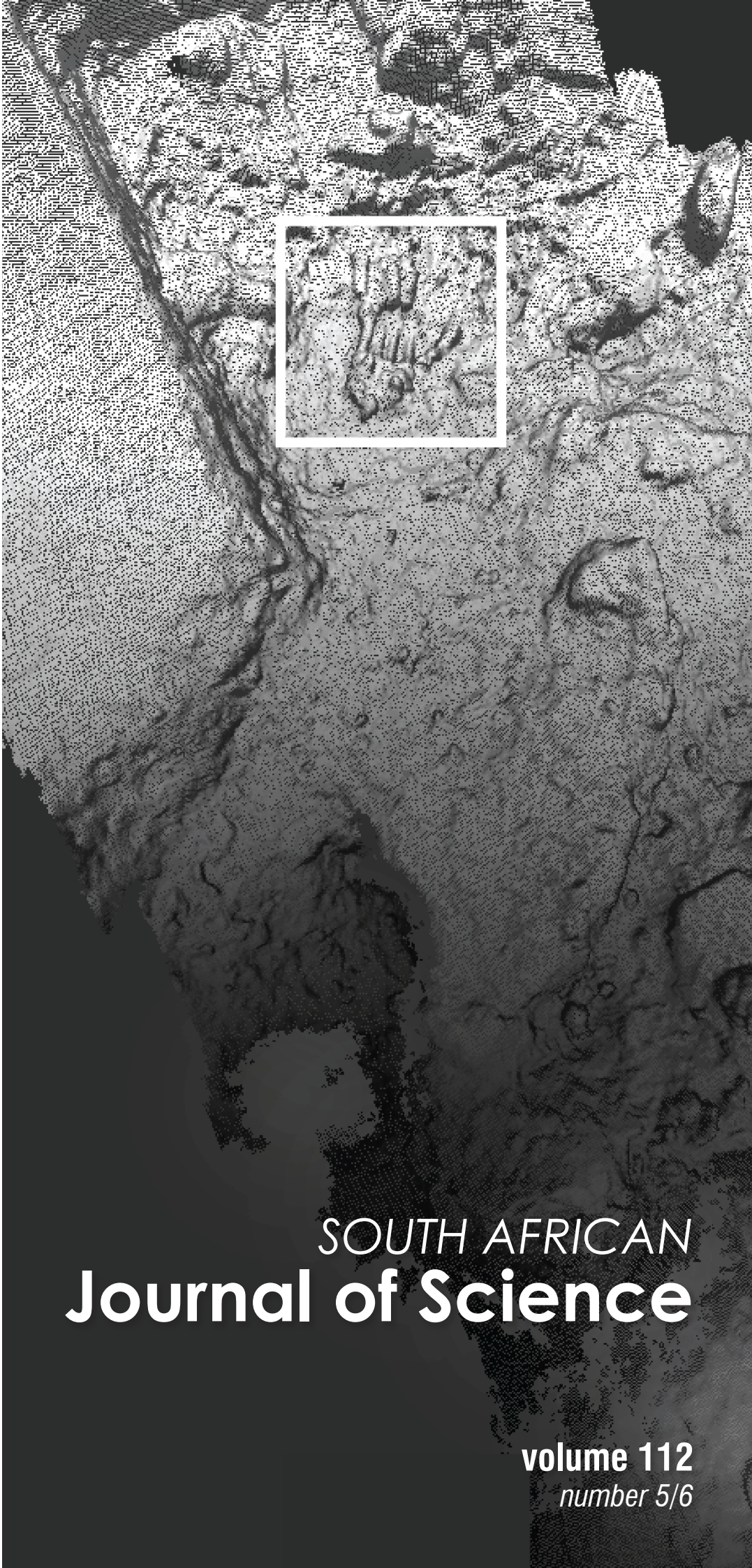
Multimodal spatial mapping and visualisation of Rising Star Cave

Climate change threatens coastal tourism in South Africa

Does the DHET research output subsidy model penalise high-citation publication?

Citizen science tools for ecological research in South Africa



MAY/JUNE 2016

eISSN: 1996-7489

Applying scientific thinking in the service of society

20 YEARS

ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

SOUTH AFRICAN Journal of Science

volume 112
number 5/6

MAY/JUNE

eISSN: 1996-7489

SOUTH AFRICAN
Journal of Science

volume 112

number 5/6

Published by
the Academy of Science of South Africa (www.assaf.org.za) with financial assistance from the Department of Science & Technology.

Design and layout
SUN MeDIA Bloemfontein
T: 051 444 2552
E: admin@sunbloem.co.za

Correspondence and enquiries
sajs@assaf.org.za

Submissions
Submissions should be made at http://mc.manuscriptcentral.com/sajs

Cover caption
A photogrammetry scan of the excavation area – in the Dinaledi Chamber of the Rising Star Cave (Cradle of Humankind, South Africa) – in which an adult hand of *Homo naledi* was found, visualised as point cloud data. Kruger and colleagues (p. 103) developed and used a multimodal set of recording and survey methods to record the fossils and map the Dinaledi Chamber.

Research Article

Research Letter

Corrigendum

Errors appear in the Research Article on page 46 of Volume 111 (11/12).

The second author's name and affiliation were reported incorrectly. The correct spelling of her name is Kathrine (full name: Kathrine E. Theron) and her affiliation is [2]Department of Medical Sciences, Public Health and Health Promotion, University of Limpopo, Polokwane, South Africa.

Reference

Adefolaju GA, Theron KE, Hosie MJ. In-vitro effects of protease inhibitors on BAX, BCL-2 *and apoptosis in two human breast cell lines. S Afr J Sci. 2015;111(11/12), Art. #2014-0417, 9 pages. http://dx.doi.org/10.17159/sajs.2015/20140417*

Flowers born to blush unseen?

Ten years ago, I was a member of a team that went to visit a school in a far outlying location of the Cabo del Dado Province in northern Mozambique, the home of more elephants than had (at the time) been counted. The team included one of my then colleagues, and three associates from the Aga Khan Foundation who were based in the town of Pemba.

To reach the school, my colleague and I flew for 2 h from Johannesburg to Pemba, spent a night there, and were then driven for 3 h in a four-wheel 'monster' to reach the school, located in the middle of the Quirimbas National Park, which stretches from inland areas to the Mozambican coastline.

This Leader is not, however, about the school, but about the lunch we were served – a salad, made of leaves that were entirely foreign to me, and served with a peanut sauce. I asked a teacher from the school what it was we were eating and he said, with great surprise at my ignorance, 'Moringa, of course'. I ate on, without disappointment – the salad was delicious.

On returning to South Africa, I looked for details about *Moringa oleifera* and discovered that its leaves are comparable to Andean quinoa grain: a remarkably nutritious food source that is an antioxidant, is rich in amino acids, and which contains, amongst other nutrients, calcium, magnesium, iron, iodine, lutein, zinc and vitamins A and B3.

It is not surprising, then, that all of 10 years later, the *South African Journal of Science* has received an unexpected number of papers that provide research findings on the nutritional and/or health values of a wide variety of indigenous, or indigenised, African plants. Many of the papers were derived from the theses and dissertations of postgraduate students who studied at universities in Africa. Not all of the manuscripts have been suitable for publication, but many that have been, have dealt with various uses of *Moringa oleifera* leaves, bark and roots, with emphasis not only on nutrition but also on the production and uses of oils and creams from various parts of the tree.

The papers most recently published speak to a range of issues, although the three areas most frequently addressed are nutritional value, food security and medicinal properties. As the papers are all about plants, these areas of focus are not surprising, although the range of research approaches and findings are remarkably diverse.

While they do not intentionally set out to do so, two papers reveal the opposite sides of the debates around 'indigenous knowledge'. In a paper dealing with sweet potatoes[1] ('indigenised' since the 1500s), the authors' focus is on nutrition and drought resistance – both critical in contributing to food security in times of drought, such as much of the country faces at present. On the other hand, the authors of a paper on wild vegetables[2] (*Corchorus olitorius* and *Amaranthus cruentus*) conclude that although claims have been made that these particular wild vegetables have superior micronutrient content to that of 'exotic' vegetables, this is demonstrably not the case. In the first case, indigenous knowledge is confirmed – sweet potatoes, a staple source of food, are suited to our climate, and to climatic changes. Conventional wisdom prevails. In the second case, however, wild vegetables are shown not to have properties superior to 'exotics'.

As might be expected, a number of papers focus on the medicinal or health value of plants. The most unusual of these, 'Plants, people and health: Three disciplines at work in Namaqualand'[3] is a paper that investigates not just plants but also multidisciplinary approaches to research. The theoretical focus of the paper is multidisciplinary while the subject focus is medicinal plants found in Namaqualand.

The main conclusion of the paper is that

> *...the primary orienting, although often unstated, question in chemical studies of plant medicine concerns pharmacologically active ingredients for antibacterial, antifungal, antiviral and/or anti-inflammatory properties. Yet this orientation depends on equating health and illness with the eradication of a particular taxonomy of pathogens. If the orientation to health includes a wider array of toxins and taxonomies that contribute to the experience of having energy or vitality, ...then biochemical research need not necessarily begin with the particular pathway of seeking compounds related to pathogen elimination.*[3]

Sadly, the paper does not go on to suggest what the other 'particular pathways' might possibly be. While the earlier papers suggested contested notions of indigenous knowledge, this paper leaves the understanding of plants in a rather ambiguous position.

Two other plant-medicinal papers proceed and conclude from what some scholars might consider to be a more limited (or, in the views of some, simply wrong), empiricist position – but they have the advantage of offering results that can be tested and readily contested.

The researchers[4] who studied the mutagenic and antioxidant effects of uMakhonya®, advocated by traditional healers for many years, concluded that it could well have value in treating the symptoms of HIV/AIDS. Similarly, a study[5] of 48 South African medicinal plants – whose antimutagenic and antioxidant properties might be effective in treating acne – concluded that, used alone or in combination with one another, the extracts have considerable potential in the successful treatment of skin ailments.

Their intrinsic values apart, these and related papers have two other contributions to make. Firstly, they draw attention to the debate around indigenous knowledge, some of which is of enormous importance and value; and some of which might well be matters of folklore. Secondly, they are powerful reminders of Einstein's often forgotten, but challenging, comment that we cannot solve our problems with the same thinking we used when we created them.

References

1. Motsa NM, Modi AT, Mabhaudhi T. Sweet potato (*Ipomoea batatas* L.) as a drought tolerant and food security crop. S Afr J Sci. 2015;111(11/12), Art. #2014-0252, 8 pages. http://dx.doi.org/10.17159/sajs.2015/20140252
2. Mavengahama S, De Clercq WP, McLachlan M. Trace element composition of two wild vegetables in response to soil-applied micronutrients. S Afr J Sci. 2014;110(9/10), Art. #2013-0339, 5 pages. http://dx.doi.org/10.1590/sajs.2014/20130339
3. Green L, Gammon DW, Hoffman MT, Cohen J, Hilgart A, Morrell RG, et al. Plants, people and health: Three disciplines at work in Namaqualand. S Afr J Sci. 2015;111(9/10), Art. #2014-0276, 12 pages. http://dx.doi.org/10.17159/sajs.2015/20140276
4. Ngcobo M, Gqaleni N, Ndlovu V. Antimutagenic and antioxidant effects of a South African traditional formulation used as an immune booster. S Afr J Sci. 2016;112(3/4), Art. #2015-0152, 7 pages. http://dx.doi.org/10.17159/sajs.2016/20150152
5. Sharma R, Lall N. Antibacterial, antioxidant activities and cytotoxicity of plants against *Propionibacterium acnes*. S Afr J Sci. 2014;110(11/12), Art. #2013-0293, 8 pages. http://dx.doi.org/10.1590/sajs.2014/20130293

HOW TO CITE: Butler-Adam J. Flowers born to blush unseen? S Afr J Sci. 2016;112(5/6), Art. #a0159, 1 page. http://dx.doi.org/10.17159/sajs.2016/a0159

Citizen science tools available for ecological research in South Africa

AUTHOR:
Joseph M. Hulbert[1]

AFFILIATION:
[1]Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
Joseph Hulbert

EMAIL:
joey.hulbert@fabi.up.ac.za

POSTAL ADDRESS:
Forestry and Agricultural Biotechnology Institute, University of Pretoria, Private Bag X20, Pretoria 0002, South Africa

KEYWORDS:
science education; public engagement; ecology; web-based platforms; social media

HOW TO CITE:
Hulbert JM. Citizen science tools available for ecological research in South Africa. S Afr J Sci. 2016;112(5/6), Art. #a0152, 2 pages. http://dx.doi.org/10.17159/sajs.2016/a0152

Citizen science is a term for scientific research conducted by non-scientists. Ordinary citizens can participate in research from their home computer, in their own gardens, or in the great outdoors – without any expertise in the field. Many citizen science projects and opportunities exist in South Africa – ranging from monitoring bird migrations to identifying and mapping distributions of fungi. In this article, three citizen science tools available to researchers in South Africa are discussed and three interesting plant-focused research projects which currently utilise these tools are introduced. Incorporating citizen science tools into ecological research provides many benefits. The use of citizen science tools increases sampling distributions, both temporally and spatially[1], and engagement of local citizens in research provides 'many scientific eyes'[2], which is especially useful for geographical studies, such as monitoring invasive species or mapping populations of rare species over large spatial scales. Not only does the research benefit from multiple volunteer observers, but also from greater access to private lands.[3] Thus, incorporating citizen science tools ultimately increases the feasibility of many research projects through the culmination of these benefits. Each of the tools discussed below may already be in use by large numbers of citizens who are actively contributing to projects which will be available for future research endeavours. However, citizen science is recognised as an inherent avenue for outreach,[3] and should only be pursued through projects for which there is a commitment and interest in improving scientific literacy by engaging a greater population. Citizen science is not only recognised as a tool for research, but also as a tool for scientific education. Citizen science projects provide ample outside-of-the-classroom learning opportunities for participants. They also provide educational benefits such as skills for accurate data collection, critical thinking and scientifically informed decision-making[2] – which ultimately increases scientific capacity, better informs decisions and improves social capital in South Africa.

Citizen science tools

Three not-for-profit tools available for ecological research projects in South Africa are discussed below. Each tool is briefly summarised and the method of application in research is introduced. Reviewing the qualities and limitations of each tool is not in the scope of this article. All three tools are web-based platforms on which citizens can register, upload images and locations, and share observations.

Ispot

www.ispotnature.org

Ispot is an international platform with a South African based initiative sponsored by the South African National Biodiversity Institute (SANBI). The project launched in South Africa in June 2012, and has contributed to the nearly 400 000 international observations of 30 000 different species reported by mid-2014.[4] The project strives to connect citizens to experts in the field through a social network, who can work together to identify organisms and learn about ecology. Researchers can create 'projects' on Ispot for which they can collect observations using a tag system. For example, the tag 'dyingfynbos' is used to filter and organise observations for a research project about plant disease in fynbos vegetation. Complexity can also be added to projects on Ispot by adding species interactions to observations (i.e. experts and citizens can work together to identify, from a single image, both the bee species and the flower species being pollinated).[5]

WhatSpecies

www.whatspecies.com

WhatSpecies is a citizen science tool that was developed in South Africa. A parent who wanted to help her children identify insects and plants launched the platform to aid people to 'live out their passion and learn about identifying nature' (Cruywagen M 2015, personal communication, November 6). The data on the platform are openly accessible and the layout of the website caters well to a young audience. A second version of WhatSpecies is expected to launch in late March 2016. The platform strives to engage youth through other forms of social media such as Facebook and blogging and is committed to protecting the ownership of uploaded images. In contrast to Ispot, researchers organise projects by creating groups that citizens can join to upload observations.

Virtual Museum

http://vmus.adu.org.za

Virtual Museum is a South African based platform hosted and organised by the Animal Demography Unit at the University of Cape Town. The platform is project oriented and is currently hosting 17 different geographical projects for mapping distributions of organisms that range from fungi and orchids to dung beetles and birds, with each comprising many genera. Using this tool, researchers create a project and citizens share observations by uploading images and selecting locations specifically for that project. Growth in the platform has been exceptional, with more than a million records of bird distributions added since the beginning of the Southern African Bird Atlas Project (SABAP2) in 2007.[6] The data are stored under a Creative Commons licence and registered users can request static maps and species lists.

Selecting a citizen science tool

Each tool discussed above has strengths and weaknesses, which are highly dependent on the scope of the research project itself. Therefore, it is important to try and consider all three when initiating a citizen science component of a research project in South Africa. Alternatively, researchers could choose to create projects in all three tools discussed, merging observations for their own analysis. All three tools are limited to geographical studies, but each could be used to recruit participants for more interactive projects that involve physical sampling. Data quality, target audience, social engagement and user-abundance (e.g. number of active citizen scientists) are important components to consider when selecting a tool. Both Ispot and WhatSpecies use crowd-sourcing to identify observations, but observation identifications on Virtual Museum are at the discretion of the project leaders and a selected 'expert panel' through checking of photographs. All three tools are suitable for mapping distributions supported by images, but each is limited to projects in which organisms can be identified from photographs (e.g. plants, fungi and birds). However, each tool can be used to direct sampling efforts to confirm observations or conduct a more thorough investigation. For example, Cape Citizen Science (discussed below) is a citizen science initiative for a research project to study plant disease throughout the Fynbos Biome. Because microscopic organisms cause plant disease, images showcasing symptoms of disease cannot be used to conclusively identify the microorganism. However, the project is currently using two of the tools discussed above to find sampling locations and potentially identify new hotspots of plant disease emergence.

South African citizen science projects

There are many citizen science projects ongoing within South Africa. As mentioned previously, there are 17 different projects that citizens can participate in hosted on Virtual Museum alone. One project on Virtual Museum and two projects on Ispot that focus on plant communities are discussed below.

OrchidMap

http://orchidmap.adu.org.za

OrchidMap is a project hosted by Virtual Museum in which citizens can upload images and locations of orchid observations. The purpose of the project is to improve the understanding of the distributions of African orchids. Nearly 3000 geo-referenced records for orchids have been uploaded to Virtual Museum for this project since it was initiated in September 2014.[6] In contrast to Ispot and WhatSpecies, Virtual Museum uses a grid system to share location data for individual observations and does not provide explicit coordinates to avoid potential abuse of the platform. This approach is important in the case of access to distributions of rare and endangered species with economical incentives, such as many orchid species. Ispot and WhatSpecies do allow users to hide location data when sharing observations, but users may be unaware of the risk.

Cape Citizen Science

http://citsci.co.za

The Forestry and Agricultural Biotechnology Institute of the University of Pretoria initiated Cape Citizen Science for a project to survey plant disease in the Fynbos Biome. The research is designed to focus on a group of microorganisms called *Phytophthora*, which translates from Greek as 'plant destroyers'. As part of the initiative, the researchers have created a project on Ispot in which citizens can contribute observations of dying plants using the tag 'dyingfynbos'. A group also exists to organise observations for this project on WhatSpecies. The reported locations will be used to choose sampling locations and are expected to help researchers and land managers respond more quickly to new invasions and diseases. Since the project was initiated, the researchers have also started collaborating with a group from the Cape Peninsula University of Technology that is studying invasive aboveground fungi in the same system. These research projects will directly benefit from shared observations from citizen scientists and will ultimately contribute positively to the conservation of the biodiversity in the Fynbos Biome.

Aliens of the Cape Peninsula

http://www.ispotnature.org/projects/aliens-of-the-cape-peninsula

Aliens of the Cape Peninsula is a project exclusively on Ispot. The aim of the project is to catch new introductions of alien plants and map the current distributions of known alien species. The project page on Ispot indicates that there are currently over 800 species of alien plants – a third of the total species present – on the Cape Peninsula and provides a long list of species for which to look. The project was initiated in early January 2016 and has already received more than 1000 observations.

Conclusion

The tools discussed above are freely and readily available for implementation in southern Africa research projects. They are incredible resources that are useful for improving public well-being through educational benefits whilst also providing future societies with basic research. These tools also enable the coupling of science education and hypothesis-driven research, thereby benefitting society and the planet by engaging the public in the scientific process.

References

1. Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg KV, et al. Citizen science: A developing tool for expanding science knowledge and scientific literacy. BioScience. 2009;59(11):977–984. http://dx.doi.org/10.1525/bio.2009.59.11.9
2. Dickinson JL, Shirk J, Bonter D, Bonney R, Crain RL, Martin J, et al. The current state of citizen science as a tool for ecological research and public engagement. Front Ecol Environ. 2012;10(6):291–297. http://dx.doi.org/10.1890/110236
3. Meentemeyer RK, Dorning MA, Vogler JB, Schmidt D, Garbelotto M. Citizen science helps predict risk of emerging infectious disease. Front Ecol Environ. 2015;13(4):189–194. http://dx.doi.org/10.1890/140299
4. Silvertown J, Harvey M, Greenwood R, Dodd M, Rosewell J, Rebelo T, et al. Crowdsourcing the identification of organisms: A case-study of iSpot. ZooKeys. 2015;480:125–146. http://dx.doi.org/10.3897/zookeys.480.8803
5. Pocock MJO, Evans DM, Fontaine C, Harvey M, Julliard R, McLaughlin Ó, et al. Chapter two – The visualisation of ecological networks, and their use as a tool for engagement, advocacy and management. In: Woodward G, Bohan DA, editors. Advances in ecological research. Cambridge, MA: Academic Press; 2016. p. 41–85. http://dx.doi.org/10.1016/bs.aecr.2015.10.006
6. Animal Demography Unit. Annual report (January to November). Cape Town: Department of Biological Sciences, University of Cape Town; 2015.

Comparison of Holocene temperature data (Boomplaas Cave) and oxygen isotope data (Cango Caves)

AUTHOR:
J. Francis Thackeray[1]

AFFILIATION:
[1]Evolutionary Studies Institute, School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Francis Thackeray

EMAIL:
Francis.thackeray@wits.ac.za

POSTAL ADDRESS:
Evolutionary Studies Institute, School of Geosciences, University of the Witwatersrand, Private Bag 3, Wits 2050, South Africa

KEYWORDS:
Quaternary; palaeoclimate; South Africa

HOW TO CITE:
Thackeray JF. Comparison of Holocene temperature data (Boomplaas Cave) and oxygen isotope data (Cango Caves). S Afr J Sci. 2016;112(5/6), Art. #a0156, 2 pages. http://dx.doi.org/10.17159/sajs.2016/a0156

Boomplaas Cave is situated near the town of Oudtshoorn in the Western Cape Province, South Africa. It was excavated by Hilary Deacon[1] in the 1970s, yielding not only Late Quaternary artefacts that were of archaeological importance but also fossils of rodents and insectivores (analysed by Margaret Avery[2]) as well as bovids and equids (analysed by Richard Klein[3]) that were important for the reconstruction of palaeoenvironments in the context of changes in climate within part of the Late Pleistocene (12 000–80 000 years BP) and Holocene (younger than 12 000 years BP). Also of palaeoclimatic importance is a stalagmite from the adjacent Cango Caves, analysed by John Vogel and Siep Talma. In 1992, Talma and Vogel[4] used a transfer function to estimate Holocene temperatures from oxygen isotope data from the Cango speleothem. Their basic oxygen isotope data (Figure 1) can be compared to temperature estimates obtained from Thackeray's[5] multivariate analysis of fossil rodents and insectivores represented in the late Quaternary sequence from Boomplaas Cave (Table 1). Here, for the first time, a comparison is made between the results obtained by Talma and Vogel[4] with those obtained by Thackeray[5], both of which relate to temperature but use independent sources of data.

The oxygen isotope values for the Cango stalagmite are shown in Figure 1a. These values can be compared to the relative changes in Thackeray's temperature estimates for the Boomplaas sequence (Figure 1b). The implication is that there is remarkably good agreement between the *raw* oxygen isotope data from the speleothem[4], and the calibrated temperature curve for the Boomplaas Holocene sequence, based on Thackeray's[5] multivariate analysis of fossil rodents and insectivores.

It would appear that the *raw* oxygen isotope data reflect variation in mean annual temperature, without using a complicated transfer function of the kind employed by Talma and Vogel[4].

Using data from the correlations indicated in Figure 1, the following regression equation can be obtained for a relationship between mean annual temperature (MAT) and oxygen isotope ratio (OIR):

$$MAT = -0.557\ OIR + 13.55$$

($r = 0.93$)

These results are potentially relevant to interpretations of oxygen isotope ratios from other speleothems, at least for part of the Quaternary in regions adjacent to Boomplaas Cave and the Cango Caves.

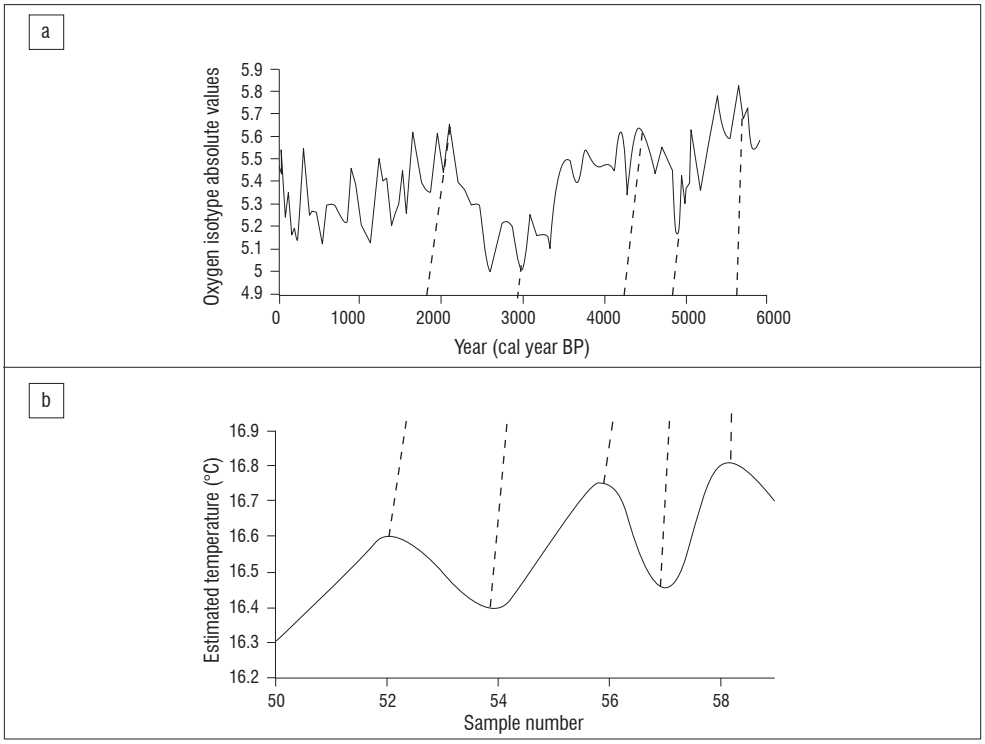


Figure 1: (a) Raw oxygen isotope data from a speleothem in the Cango Caves, obtained by Talma and Vogel[4] with calibrated radiocarbon dates within the past 6000 years. Data were provided for this study by Talma. (b) Holocene mean annual temperature curve for samples 50–59 from Boomplaas Cave, as calculated by Thackeray[5]. These Holocene samples are dated between circa 1000 and 6000 BP (see radiocarbon dates in Table 1). The two curves appear to be correlated (dashed lines).

Table 1: Calibrated temperatures for a set of Boomplaas Holocene samples (numbered from 50 to 59), with radiocarbon dates[2]

Boomplaas sample	Calibrated temperature (°C)	Date (BP)
50	16.3	1630
51	16.45	
52	16.6	
53	16.5	
54	16.4	
55	16.6	
56	16.75	
57	16.45	6400
58	16.8	
59	16.7	

Acknowledgements

I am grateful to Siep Talma for access to the tabulated oxygen isotope data for the Cango stalagmite. I thank Jennifer Fitchett for assistance with Figure 1. Margaret Avery identified the rodents and insectivores from excavations at Boomplaas, directed by Hilary Deacon. The correlations presented are the result of ongoing research on palaeoclimates and faunal analyses, supported by grants awarded to J.F. Thackeray from the National Research Foundation (South Africa) and the A.W. Mellon Foundation.

References

1. Deacon HJ. Two Late Pleistocene-Holocene archaeological depositories from the southern Cape, South Africa. S Afr Archaeol Bull. 1995;50:121–131. http://dx.doi.org/10.2307/3889061
2. Avery DM. Micromammals as palaeoenvironmental indicators and an interpretation of the Late Quaternary in the Southern Cape Province, South Africa. Ann S Afr Mus. 1982;85:183–374.
3. Klein RG. Environmental and ecological implications of large mammals from Upper Pleistocene and Holocene sites in southern Africa. Ann S Afr Mus. 1980;81:223–283.
4. Talma AS, Vogel JC. Late Quaternary paleotemperatures derived from a speleothem from Cango Caves, Cape Province, South Africa. Quaternary Res. 1992;37:203–213. http://dx.doi.org/10.1016/0033-5894(92)90082-T
5. Thackeray JF. Late Quaternary environmental changes inferred from small mammalian fauna, southern Africa. Climatic Change. 1997;10:285–305. http://dx.doi.org/10.1007/BF00143907

An actuarial artificial intelligence for the game rock-paper-scissors

AUTHOR:
Michael B. Jordan[1]

AFFILIATION:
[1]Student Member of the Actuarial Society of South Africa, Cape Town, South Africa

CORRESPONDENCE TO:
Michael Jordan

EMAIL:
bayman1991@gmail.com

POSTAL ADDRESS:
PO Box 50362, Waterfront 8001, South Africa

KEYWORDS:
game theory; AI; machine learning; actuarial science

HOW TO CITE:
Jordan MB. An actuarial artificial intelligence for the game rock-paper-scissors. S Afr J Sci. 2016;112(5/6), Art. #a0154, 3 pages. http://dx.doi.org/10.17159/sajs.2016/a0154

Rock-paper-scissors is a simple game played by two players who use hand gestures that resemble a 'rock' (a fist), a piece of 'paper' (a flat hand) or a pair of 'scissors' (index and middle fingers in the shape of a V). Each player plays a gesture at the same time, with the scoring as follows: 'rock' beats 'scissors'; 'paper' beats 'rock'; and 'scissors' beats 'paper'.

If the game is played only once, no dominating strategy exists and the result is down to chance. If the game is played multiple times the chance element reduces as a player's next gesture is influenced by two factors: the previous result (win, loss or draw) and the previous gesture ('rock', 'paper' or 'scissors').

Previous attempts by RoShamBot to create an artificial intelligence (or AI) for rock-paper-scissors have relied on simple frequency analysis (i.e. the user has played 'rock' the most, therefore the AI must play 'paper' next) or history matching (i.e. the AI matching the last four rounds to a large database of previous games and determining what gesture the user will play next and then countering that gesture).

In an actuarial spirit, I take a stochastic approach and instead of determining the next gesture with certainty, I flex the probabilities for 'rock', 'paper' and 'scissors' and then let my AI randomly choose one. The probabilities are flexed based on previous results and previous gestures.

An outline of various strategies

The question arises as to how the previous results and previous gestures influence the players' next gestures. I propose that there are 10 potential strategies that a player can adopt in a multi-round game of rock-paper-scissors. An example of one strategy is: a player repeats a gesture when they win, alternates gestures when they lose and chooses randomly when they draw. Thus for each of the three results, there are three potential responses (repeat, alternate, random). There are therefore nine (three x three) potential strategies. Another example of this kind is when a player alternates their gestures when they win, chooses randomly when they lose and alternates their gestures when they draw. The tenth strategy is to identify which of the nine strategies the other player has adopted and then to counter that strategy. For example, if an opponent tends to repeat their gestures when they win and if they had won the previous round with the gesture of 'rock', then a player would counter their strategy by using 'paper' in the next round.

It is difficult for a human to accurately determine which strategy their opponent is playing. In order for them to do so they would need to keep track of eight variables:

- The previous gesture
- The previous result
- A count of the number of times their opponent alternates after winning
- A count of the number of times their opponent repeats after winning
- A count of the number of times their opponent alternates after losing
- A count of the number of times their opponent repeats after losing
- A count of the number of times their opponent alternates after drawing
- A count of the number of times their opponent repeats after drawing

The player would also need to do a few calculations. For example, is the count for alternating after winning larger than the count for repeating after winning? If the answer is yes, then what does that say about their next gesture and how do you counter it? The time between game rounds might not be sufficient for an untrained human to adopt the tenth strategy, but it would be very simple for a computer.

Defining the variables

The first thing my AI needs is starting probabilities for each gesture. I assign 33.3% each for 'rock', 'paper' and 'scissors' because, for the first round, it is a game of chance. I create six variables that will track how my opponent plays. I create two variables to store the previous result and previous gesture. I also create an AI Confidence and an AI ForgetLimit (these are discussed later). Finally, I create a random variable that can take any value between 0 and 100.

```
//Thinking
private double repeatStyleWin = 0;
private double repeatStyleLose = 0;
private double repeatStyleDraw = 0;
private double alterStyleWin = 0;
private double alterStyleLose = 0;
```

```
private double alterStyleDraw = 0;
//Counters
private String historyChoice = "none";
private String historyResult = "none";
//Memory
private double confidence;
private double forgetLimit;
//Personality
```

The basics

The human opponents to my AI are allowed to input 'rock', 'paper' or 'scissors' as gestures. Inputting a gesture calls the play method, which works as follows: A random number is generated. If it is less than the probability for 'rock', then my AI chooses 'rock'. If it is less than the probability for 'rock' + 'paper', then my AI chooses 'paper'. If it is larger than the probabilities for 'rock' + 'paper', then my AI chooses 'scissors'. The gesture randomly chosen by the AI is then compared to the opponents input gesture, a result is determined and the scores are appropriately updated. Next the AI reviews the previous result and previous gesture and makes the necessary count that will later determine the opponent's strategy.

```
public void rock(View v){
  play("rock");
}
public void paper(View v){
  play("paper");
}
public void scissors(View v){
  play("scissors");
}
```

Making the AI forget

I then call a method known as Forget&Adjust, which is not very important for the first few rounds but is developed so that my AI can continue to counter its opponent if they change their strategy later on. Opponents are likely to change their strategy if they are losing to my AI so it is important for the AI to detect and adapt if they do so. Forget&Adjust is achieved by simply capping the difference between the alternate counts and the repeat counts. For example, if the repeat count for win is 5 and the alternate count for win is 1, the difference between them is 4. If the cap limit is set as 3, the repeat count will be reduced by 1. In this way, the AI still acknowledges that the opponent is repeating for wins but the repeat count does not increase so much that my AI becomes predictable for an extended period of rounds. Capping the counts allows my AI to react sooner to an opponent who changes their strategy.

The cap limit – that is, the AI ForgetLimit – does not have a fixed value. In the example above I set it to 3, but it could be 1 or 5 or 10. The optimal cap limit is unknown, so I let my AI decide for itself. Before each game with a new opponent, the AI ForgetLimit is determined randomly. At the end of the game the AI notes if that cap limit was successful or not. When a new opponent starts the game, a new ForgetLimit is randomly chosen and the success of that game is noted. After this training period, my AI determines which ForgetLimit was empirically the most successful and then adopts that value. This value is semi-permanent. A series of overall game losses would result in the AI wanting to enter training again. The AI will communicate this eventuality to me and I will have the ability to override the decision, as it would not be desirable to enter the AI into a competition when its cap limit is randomly high, because it would be potentially vulnerable to an opponent who changes their strategy. The more games the AI completes in training mode, the more likely an optimal value for the ForgetLimit will be found.

```
public void forgetAndAdjust(){
  if (alterStyleWin > repeatStyleWin + forgetLimit){
    alterStyleWin = alterStyleWin - 1;
  }
  if (repeatStyleWin > alterStyleWin + forgetLimit){
    repeatStyleWin = repeatStyleWin - 1;
  }
  if (alterStyleDraw > repeatStyleDraw + forgetLimit){
    alterStyleDraw = alterStyleDraw - 1;
  }
  if (repeatStyleDraw > alterStyleDraw + forgetLimit){
    repeatStyleDraw = repeatStyleDraw - 1;
  }
  if (alterStyleLose > repeatStyleLose + forgetLimit){
    alterStyleLose = alterStyleLose - 1;
  }
  if (repeatStyleLose > alterStyleLose + forgetLimit){
    repeatStyleLose = repeatStyleLose - 1;
  }
}
```

Resetting the probabilities

Before I adjust the probabilities to counter my opponent's strategy, I reset the probabilities for 'rock', 'paper' and 'scissors' to 33.3%. I want my AI to make its next decision based only on my opponent's strategy, the last result and the last gesture they played. Their strategy already incorporates all historical results and gestures and if not reset, the probabilities for 'rock', 'paper' and 'scissors' would be distorted. Therefore I flex the probabilities, that is, I adjust them from 33.3% to a new value, and then reset them to 33.3% before I adjust them again.

```
paperChance = 33.3;
rockChance = 33.3;
```

Flexing the probabilities

Once the probabilities have been reset, I flex them to counter my opponent's strategy. (The logic is not straightforward so I will provide an example at the end.) The method starts by looking at the opponent's previous gesture, and comparing the repeating count with the alternating count for the previous result. There are three possible outcomes. Outcome 1: if the repeating count is greater than the alternating count, then it follows that the opponent's result (win/loss/draw) with a particular gesture ('rock', 'paper' or 'scissors') is likely to prompt that same gesture again. Therefore I increase the probability for the gesture that beats that one and slightly increase the probability for the same gesture. Outcome 2: if the alternating count is greater than the repeating count, then it follows that the opponent's result with that gesture is likely to prompt a different gesture in the next round. Therefore I increase the probability for the gesture that loses to that one and slightly increase the probability for the winning gesture. Outcome 3: if the repeating count is the same as the alternating count, I change nothing.

For example: if the opponent previously won with 'rock', and they tend to repeat winning gestures, then I assume that they are more likely to play 'rock' again. To counter this play, the probabilities are flexed to favour 'paper', which would result in a win for the AI. The probabilities are also flexed to slightly increase the chance for 'rock' as the second best result is a draw; the probability for 'scissors' would be decreased.

```
public void adjustStrategy(String choice, double alter, double repeat){
  if (choice.equalsIgnoreCase("rock")){
    if (repeat > alter) {
      rockChance = rockChance + (confidence/2);
```

```
    paperChance = paperChance + confidence;
    //She results with rock, likely to do rock again, paper is best
    strategy next
    //Second best is to play rock and get draw
  }
  else
  if(alter > repeat){
    rockChance = rockChance - (confidence / 2);
    paperChance = paperChance - confidence;
    //She results with rock, likely not to do rock again, scissors is
    best strategy next
    //Second best is to play rock and go for a 50/50 win/lose
  }
```

Starting confidence

I do not know by how much the probabilities should flex. Therefore I create a variable known as AI Confidence. As for AI ForgetLimit, the AI will learn the optimal value by going through a training session in which it assigns random values for AI Confidence and then selects the value that has the highest end game success rate. I call this type of learning Global Machine Learning, as the learning occurs across various games with various opponents. Unlike AI ForgetLimit, AI Confidence can change during the game. If the AI wins, then its confidence increases; but if the AI loses, then its confidence decreases – I call this Local Machine Learning, as the learning occurs during the game with one opponent.

I am careful when it comes to increasing or decreasing confidence. I only increase confidence when the AI wins with the gesture that has the highest probability, and I only decrease confidence when the AI loses with the gesture that has the highest probability. For example, with the probabilities of 'rock' at 50%, 'paper' at 30% and 'scissors' at 20%, if my opponent plays 'scissors' and my AI plays 'paper' it will lose, but it should not be punished because 'paper' was not the highest probability. If my opponent plays 'scissors' and my AI plays 'rock' it would win and should be rewarded, as 'rock' was the highest probability. The higher the AI Confidence, the more the probabilities will be increased or decreased. Another way of thinking about it is that the AI Confidence represents the clarity of the opponent's strategy. An opponent who always repeats after a win and who always alternates after a loss will cause the AI Confidence to increase. An opponent who is completely random with their gestures will cause the AI Confidence to decrease. Note that it is possible for the AI to lose overall and yet end with a higher confidence and it is possible for the AI to win overall and yet end with a lower confidence. Although this scenario is unlikely as result and confidence are correlated to some extent, it is not improbable because we are rewarding only some wins and punishing only some losses. In the example above, if the opponent plays 'paper' and by chance the AI plays 'scissors' it will win, but should not be rewarded as the opponent has not played what we anticipated them to play; in fact, one could argue that after this type of win the AI should be punished and its confidence reduced. This scenario opens up another possibility – one could base the confidence not on the result and the AI's choice but rather on whether the opponent plays the gesture that they were most anticipated to play. In this way, if the opponent is following a clear strategy, the probabilities will flex more to counter. The worst case scenario is if no clear strategy for the opponent can be identified and thus AI Confidence will reduce and will leave the probabilities at 33.3% for each gesture. In this scenario the final outcome will be based on luck.

My AI can be considered successful if it wins the majority of games, it is possible for it to lose and it is not designed to guarantee a win every time. One of the benefits of taking on a stochastic nature is that my AI cannot be outwitted by its opponent. If my AI was deterministic, then patterns would emerge which could be taken advantage of by its opponents. To further prevent being outwitted, the final values for AI Confidence and AI ForgetLimit should be kept secret from all parties, including myself as the creator.

A colleague pointed out the weakness of these check digit verification (or CDV) checks and recommended using a stochastic variable in the key to improve security. My AI could use this more secure method of encryption for protecting its AI Confidence and AI ForgetLimit.

Fluctuating confidence

One consideration is by how much the confidence should increase or decrease. Again, the optimal value of this variable is unknown. I could make it random and let the AI decide the optimal value through the same training process for AI Confidence and AI ForgetLimit. But that option is not ideal as the more variables that require optimisation, the more extended the AI's training period becomes. That being said, I arbitrarily adjust confidence by 1 point, but I acknowledge that this value may not be optimal. I could train the AI for AI Confidence and AI ForgetLimit and then create AI Adjust as a random variable and train again. But optimising two variables and then another is inferior to simultaneously optimising all three because of potential dependency.

Another consideration is whether there should be limits for AI Confidence – the minimum could be zero but what should the maximum be? Also, should the probabilities be limited? During my testing of earlier versions, I did occasionally get negative probabilities and probabilities exceeding 100%; these unrealistic values distort the AI's performance and can be avoided by either limiting the probabilities or limiting AI Confidence. I also question by what fraction of confidence the second gesture should be adjusted. The higher the fraction, the lower I must limit the AI Confidence if I want to avoid impossible probabilities. The limit for AI Confidence must be less than $(33.3(x/x+1))$, where $1/x$ represents the selected fraction. Therefore, if the second gesture changes by half the confidence then the limit is 22.2, and if the second gesture changes by a third of the confidence then the limit is 25.

Future applications

Humanity is on a journey to make a cognitive computer that can rival our own intellect. We are going to create machines that humans can interact with on a social and emotional level. Playing games together is one way we can socially connect with them.

When we think of an AI we think of it as an individual computer, but it could very well take on a hive model. This AI that plays rock-paper-scissors could be just one of millions of AIs in a hive. Each AI could be assigned certain tasks instead of having one system that does everything. It is my hope that the AI described above gets adopted into future hives and is the one called upon when a computer is challenged to a game of rock-paper-scissors.

This application is just the beginning. This AI can be extended to determine strategies for more complicated games like American football. AI could also provide solutions in education. AI is about using rules, semantics and logic to program a machine to learn – why not reverse engineer these instructions into a study method that helps students to understand information in the most efficient way possible?

I have already started building an AI for American football and have already started developing the most efficient way to study. Actuarial science has provided us with a skill set to approach this subject and I believe actuaries will play a major role in creating the cognitive computer.

Marc Andreesen predicted that there would be two kinds of people in tomorrow's world: those who tell computers what to do and those who are told by computers what to do. I think those in the actuarial profession will agree that we do not want to be the latter.

Download an app of this AI here: https://play.google.com/store/apps/details?id=mjprojects.rockpaperscissors&hl=en

Or watch the YouTube video here: https://youtu.be/K4uH3BqPe3c

A global roadmap for climate change action: From COP17 in Durban to COP21 in Paris

AUTHOR:
Debra Roberts[1,2]

AFFILIATIONS:
[1]Environmental Planning and Climate Protection Department, eThekwini Municipality, Durban, South Africa

[2]School of Life Sciences, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Debra Roberts

EMAIL:
debra.roberts@durban.gov.za

POSTAL ADDRESS:
Environmental Planning and Climate Protection Department, eThekwini Municipality, PO Box 680, Durban 4000, South Africa

KEYWORDS:
Paris Agreement; UNFCCC; carbon dioxide; global warming; super-wicked problem

HOW TO CITE:
Roberts D. A global roadmap for climate change action: From COP17 in Durban to COP21 in Paris. S Afr J Sci. 2016;112(5/6), Art. #a0158, 3 pages. http://dx.doi.org/10.17159/sajs.2016/a0158

Late one Saturday night – 12 December 2015 – on the outskirts of Paris while sitting in the chilly prefabricated buildings of a repurposed airport known as Le Bourget, the world's climate change negotiators watched as a new climate agreement was gavelled into existence. This new, legally binding agreement known as the Paris Agreement, effectively set the world – including South Africa – onto a new, climate-focused development path.

Human-induced climate change is acknowledged to be the greatest challenge facing our society and economy.[1] Our excessive and unsustainable use of fossil fuels and destruction of the earth's natural ecosystems have resulted in greenhouse gases (GHGs) – such as carbon dioxide (CO_2) and methane (CH_4) – which enter the atmosphere and trap heat, thereby driving an increase in global average temperatures. Climate change has been classified as a super-wicked problem[2] – in other words, a global environmental problem for which the time available to find a solution is rapidly decreasing and where our irrational discounting of the impacts (that is, spending only small amounts now to combat climate change and ramping up slowly over time) pushes the required responses and associated costs into the future. Ironically, super-wicked problems are caused by those seeking the solution, and finding solutions is further impeded by the fact that the central authority responsible for action is weak or non-existent. However, the Paris Agreement, together with the Sustainable Development Goals adopted in New York in September 2015, offer us the first real hope that we might be able to address climate change in a more timeous, responsible, united, equitable and sustainable way. This article provides a brief background to the United Nations Framework Convention on Climate Change's (UNFCCC) COP21 session in Paris and highlights some of the key elements and implications of the Paris Agreement going forward.

What is the UNFCCC?

The international political response to climate change began when the United Nations Framework Convention on Climate Change (UNFCCC) was adopted at the United Nations Headquarters in New York on 9 May 1992. The Convention opened for signature at the Rio Earth Summit in June 1992 and was entered into force on 21 March 1994. Currently, the UNFCCC has a near-universal membership of 197 parties (aka 'countries' in UNFCCC language). The aim of the UNFCCC is to stabilise atmospheric concentrations of GHGs to avoid 'dangerous anthropogenic interference with the climate system'[3]. At the end of every year, a conference of parties (COP) is held to review the state of the Convention's implementation. The first COP took place in Berlin in 1995, and while the last two decades have featured notable highs (e.g. COP3 during which the Kyoto Protocol was adopted) and disappointing lows (COP15 in Copenhagen where an agreement to succeed the Kyoto Protocol was not achieved), COP17 in Durban in 2011 marked a particularly important turning point in the negotiating process.

In Durban's International Convention Centre, the world's governments committed to a new universal legally binding climate change agreement by 2015, to be operationalised in 2020 at the end of the second commitment period of the Kyoto Protocol. The Kyoto Protocol is a critical part of the international climate change regime as it commits developed countries to internationally binding emission reduction targets (there are no equivalent commitments for developing countries). The first commitment period ran from 2008 to 2012, and the Doha Amendment established a second commitment period from 2013 to 2020. At COP17 there was uncertainty about the form that the new agreement (to replace the Kyoto Protocol) would take and as a result the final decision indicated that it could be 'a protocol, another legal instrument or an agreed outcome with legal force under the Convention applicable to all Parties'[4], leaving everyone wondering what might ultimately emerge. The Ad Hoc Working Group on the Durban Platform for Enhanced Action (or ADP) was established as the negotiating platform for this important work.

So as climate change negotiators from around the world arrived in Paris at the end of November 2015 for COP21 they were all aware that the goal was to finish the job started at COP17 in Durban. As noted by President Zuma in his address during the leaders' session on the opening day:

> *As leaders, we are here today because of the historic and bold decisions we took in Durban to enhance the implementation of the Convention...For South Africa, the Durban decision to enhance the implementation of the Convention was of paramount importance, because the impacts of climate change are harshest on the poor. Climate change is a major global challenge that requires an urgent global response.*[5]

Three important climate milestones reached just before COP21 highlighted what was at stake if the global community failed to fulfil the Durban commitments:

1. 2015 was set to be the hottest year on record.
2. The UK Meteorological Office announced that climate change was set to pass the milestone of 1 °C of warming since pre-industrial times (1750) by the end of 2015 – putting the world half way to the stated global policy goal of limiting global temperature rise to 2 °C above pre-industrial levels. This target was first proposed by the European Union (EU) in 1996, with support from some environmentalists and scientists, and subsequently included in the 2009 Copenhagen Accord.
3. The World Meteorological Organization announced that 2016 would be the first year in which the concentration of CO_2 in the atmosphere would exceed 400 ppm on average as a result of the continued burning of fossil fuels. This figure is up from about 278 ppm in pre-industrial times (1750). A figure of 400 ppm is regarded as a 'symbolic milestone', but science suggests that the 'safe' level of CO_2 is much closer to 350 ppm.[6]

What was the outcome of COP21?

When the COP21 President, Laurent Fabius, brought down the gavel at Le Bourget after two long weeks and several sleepless nights, the Paris Agreement was unanimously adopted (even though it seemed that Nicaragua wanted to take the floor to object). Despite its name, the Paris Agreement is two different documents. Firstly, the Agreement itself, which is a legally binding treaty on climate action containing emission reduction commitments from 187 countries (this number will increase as more countries submit their commitments prior to 2020). Secondly, the Paris Decision which passes the Agreement, prepares for its implementation once it enters into force and sets out a number of less legally binding ways to accelerate climate action with immediate effect. What is significant about the Paris Agreement is that it charts a new political course for the world. It ends the strict differentiation between developed and developing countries enshrined in the Kyoto Protocol, and marks the beginning of a new era of collective will. It is an acknowledgement that climate change cannot be addressed effectively unless everyone acts according to their respective capabilities and resources.

Some of the key Articles contained within the Agreement are outlined below:

Article 2 – The temperature goal

This Article reaffirms the goal of limiting global temperature increase to well below 2 °C, while pursuing efforts to limit the increase to 1.5 °C.[7] This is a much stronger outcome than many countries had thought possible, but still falls well short of the aspirations of many small island states and least developed countries who had wanted 1.5 °C established as an absolute limit. The Paris Decision also invites the Intergovernmental Panel on Climate Change to produce a Special Report on the impacts of 1.5 °C and related global GHG emission pathways by 2018.

Article 4 – Mitigation

In terms of the long-term emissions goal it was decided to peak global GHG emissions as soon as possible (recognising that this will take longer for developing countries) and then undertake rapid reductions so that all anthropogenic emissions are balanced with 'removal by sinks'[7] in the second half of the century. Effectively, this means reaching net-zero emissions after 2050. In order to achieve this goal, Article 4 establishes binding commitments for all countries to make 'nationally determined contributions'[7] (NDCs) and to pursue domestic measures aimed at achieving them. All countries have to submit new NDCs every 5 years, with the expectation that they will 'represent a progression' beyond previous ones (as specified in Article 3) and reflect the 'highest possible ambition'.[7] It is important to note, however, that implementation of the NDCs is not a binding obligation.

Article 7 – Adaptation

In his leaders' day address, President Zuma stressed that: 'A global goal for adaptation must…be part of the Paris Agreement. This goal must express adaptation as a global responsibility that requires a global response'[5]. This objective was fully realised through the 'global goal' on adaptation in Article 7 which focuses on 'enhancing adaptive capacity, strengthening resilience and reducing vulnerability to climate change'[7]. Article 7 also links the amount of adaptation and its cost to the level of mitigation action (i.e. the less mitigation, the more adaptation required and the greater the cost). The Paris Agreement is the first agreement of its kind to put the need for adaptation on an equal footing with the need for mitigation. This is a remarkable achievement given that an adaptation goal was unthinkable even 2 years before COP21.

Article 8 – Loss and damage

Loss and damage refers to the permanent loss and residual damage that remains after mitigation and adaptation efforts. In the Paris Agreement loss and damage has its own Article which makes permanent the Warsaw International Mechanism for Loss and Damage (established as an interim body at COP19 in Warsaw in 2013). This Article is regarded as an important political statement, putting loss and damage on a par with mitigation and adaptation, and is regarded as a victory for small island countries and other countries highly vulnerable to climate impacts. Nevertheless, at the insistence of developed countries (led by the USA), the accompanying Paris Decision specifies that the loss and damage provision 'does not involve or provide a basis for any liability or compensation'[7] as these countries wanted to remove the possibility of climate reparation claims arising from their responsibility for most of the GHGs currently in the atmosphere.

Article 9 – Finance

Finance is always a contentious issue and COP21 was no exception. In Paris, poorer developing countries worked to ensure that financial support for their actions would be increased, while developed countries lobbied for wealthier developing countries to contribute as well. Both succeeded to some extent, given that Article 9 places a legal obligation on developed countries to continue to provide climate finance for mitigation and adaptation to developing countries ('in continuation of their existing obligations under the Convention'[7]) and encourages other countries, for the first time, to provide support voluntarily. Many of the financial details were, however, moved to the decision text including the provision that, prior to 2025, countries should agree a 'new collective quantified goal'[7] from the floor of USD100 billion per year, which is the current aspiration.

Article 13 – Transparency

The Paris Agreement rests heavily on transparency as a means of holding countries accountable. Article 13 commits all countries to report regularly on their emissions and 'progress made in implementing and achieving'[7] their NDCs, and to undergo international technical expert review. This review process is intended to be 'facilitative, non-intrusive, non-punitive'[7]. The rules on transparency were a priority for the USA and EU, who wanted to ensure that China was equally scrutinised in terms of its emission reduction efforts.

Article 14 – Global stocktake

An assessment undertaken of the national climate plans of 146 countries as of 1 October 2015[8] showed that the intended nationally determined contributions (INDCs) submitted at that point could limit global temperature increase to approximately 2.7 °C by 2100. Because the current commitments by countries are not sufficient to reach the 2 °C temperature goal, a two-stage process to increase ambition over time was agreed to. The first stage will involve a facilitative dialogue held in 2018 to take stock of the collective efforts of countries which should help update and enhance individual country plans. This process will then be repeated every 5 years, with the first post-2020 stocktake occurring in 2023.[7]

The importance of non-party stakeholders

The goals outlined in the Paris Agreement are substantially more ambitious than many believed possible prior to COP21, but such ambitions cannot be realised by national governments alone. Recognising this, the Paris Decision welcomes the efforts of all non-party stakeholders in addressing and responding to climate change, including those of civil society, the private sector, financial institutions, cities and other sub-national authorities.[7] Non-party stakeholders are also invited to scale up their efforts and support actions to reduce emissions and/or to build resilience and decrease vulnerability to the adverse effects of climate change and to showcase these efforts via the Non-State Actor Zone for Climate Action (NAZCA). A large number of commitments was made to reduce emissions and increase adaptive capacity prior to and during COP21 by countries, regions, cities, investors and companies, often along with governments, under the umbrella of the Lima Paris Action Agenda (LPAA). The LPAA is an initiative led by France, Peru, the UN Secretary-General and secretariat of the UNFCCC, and its objective is to showcase such commitments and partnerships. A good example is 'Mission Innovation'. This is an initiative of 20 countries to accelerate global clean energy innovation, including doubling their current research and development investments in the sector and is coupled with a private sector effort called the 'Breakthrough Energy Coalition', in which 28 billionaire investors from 10 countries, led by Bill Gates, will invest private capital in clean energy.

The role of cities and local government

Acknowledging the important role of cities and local governments in tackling climate change, Anne Hidalgo, Mayor of the City of Paris, and Michael R. Bloomberg, the UN Secretary-General's Special Envoy for Cities and Climate Change – in partnership with the major global cities and local government networks – co-hosted the 'Climate Summit for Local Leaders' on 4 December during COP21. Attended by 1000 city and regional officials, the summit was the largest ever global convening of mayors, governors and local leaders on climate change and culminated in the *Paris City Hall Declaration*[9]. This declaration pledges local and regional leaders' support for 100% renewable energy and 80% reduction in emissions by 2050 and for the production and implementation of participatory resilience strategies and action plans to adapt to climate change hazards by 2020. It also commits these leaders to reduce urban GHG emissions by up to 3.7 Gt per annum by 2030 – the equivalent of up to 30% of the difference between current national commitments and the 2 °C emissions reduction pathway. This commitment is in line with recent research[10] which shows that urban policy decisions made by 2020 could determine up to a third of the remaining global carbon budget that is not already 'locked in' by past decisions.

So how good are the Paris outcomes?

As many negotiators pointed out during the COP21 closing plenary, the Paris Agreement is a compromise document and as such 'is good, but not perfect'. Despite its imperfections, it is clear that the Paris Agreement will shape climate action for decades to come. It has broken new ground and placed adaptation, resilience and response to climate impacts at the heart of the new regime. At the same time, the unexpectedly ambitious goals, universal nature and the near-universal coverage of NDCs, the 5-year review cycles and the transparency framework, provide much-needed signals to global markets to redirect investments to low-carbon and climate-resilient development. As noted by May Boeve, Executive Director 350.org: 'There is no way to meet the targets laid out in this agreement without keeping coal, oil and gas in the ground. The text should send a clear signal to fossil fuel investors: divest now.'[11] Future investment will therefore need to be compatible with a zero carbon world, marking the end of the era of fossil fuels. The need to achieve a fossil-fuel free future means that governments and investors will have to manage an orderly transition away from a fossil fuel dominated economy in a way that avoids stranded assets and negative impacts on workers. This transition will need to be a rapid one as new research suggests that if the 2 °C target is to be met, no new GHG emitting infrastructure can be built after 2017 unless other electricity infrastructure is retired early or retrofitted with carbon capture technologies.[12] The implications for South Africa are obvious and serious. It is clear that we need to start re-imagining our future in the post-Paris world – business-as-usual will no longer do.

Where to from here?

The Paris Agreement is a remarkable, but still fragile, agreement and its success will depend on the extent to which countries carry the spirit of Le Bourget forward. The agreement will enter into force once 55 countries accounting for at least 55% of global emissions have acceded to it. Countries will need to sign the Paris Agreement in New York between 22 April 2016 and 21 April 2017, and also adopt it within their own legal systems (through ratification, acceptance, approval or accession). On 22 April 2016, 175 parties signed the Paris Agreement at the UN Headquarters in New York – a record for first-day signatures to an international agreement. It is critical that all countries continue to act quickly as time is rapidly running out – on 9 April 2016, the Mauna Loa Observatory recorded the highest ever daily average measurement of CO_2: 409.44 ppm.[13]

References

1. World Economic Forum. The global risks report 2016. 11th ed. Geneva: World Economic Forum; 2016.
2. Levin K, Cashore B, Bernstein S, Auld G. Overcoming the tragedy of super wicked problems: Constraining our future selves to ameliorate global climate change. Policy Sci. 2012;45:123–152. http://dx.doi.org/10.1007/s11077-012-9151-0
3. United Nations. United Nations framework convention on climate change [document on the Internet]. c1992 [cited 2016 Apr 16. Available from: https://unfccc.int/files/essential_background/background_publications_htmlpdf/application/pdf/conveng.pdf
4. United Nations Framework Convention on Climate Change. Report of the Conference of the Parties on its seventeenth session, held in Durban from 28 November to 11 December 2011 [document on the Internet]. c2011 [cited 2016 Apr 16]. Available from: http://unfccc.int/resource/docs/2011/cop17/eng/09a01.pdf
5. United Nations Framework Convention on Climate Change. Statement by H.E. President Jacob Zuma to the opening session of the Paris Climate Change Conference, 30 November 2015, Paris [document on the Internet]. c2015 [cited 2016 Apr 16]. Available from: https://unfccc.int/files/meetings/paris_nov_2015/application/pdf/cop21cmp11_leaders_event_south_africa.pdf
6. Hansen J, Sato M, Kharecha P, Beerling D, Berner R, Masson-Delmotte V, et al. Target atmospheric CO_2: Where should humanity aim? Open Atmos Sci J. 2008;2:217–231. http://dx.doi.org/10.2174/1874282300802010217
7. United Nations Framework Convention on Climate Change. Paris Agreement [document on the Internet]. c2015 [cited 2016 Apr 16]. Available from: http://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf
8. United Nations Framework Convention on Climate Change. Synthesis report on the aggregate effect of the intended nationally determined contributions [document on the Internet]. c2015 [cited 2016 Apr 16]. Available from: http://unfccc.int/resource/docs/2015/cop21/eng/07.pdf
9. United Cities and Local Governments. Paris City Hall Declaration. A decisive contribution to COP21 [document on the Internet]. c2015 [cited 2016 Apr 16]. Available from: http://www.uclg.org/sites/default/files/climate_summit_final_declaration.pdf
10. Erickson P, Tempest K. Keeping cities green: Avoiding carbon lock-in due to urban development. SEI Working Paper No. 2015-11. Seattle, WA: Stockholm Environment Institute; 2015.
11. 350.org. 350.org & Bill McKibben react to COP21 climate text [homepage on the Internet]. c2015 [cited 2016 Apr 16]. Available from: https://350.org/press-release/cop21-reaction/
12. Pfeiffer A, Millar R, Hepburn C, Beinhocker E. The '2°C capital stock' for electricity generation: Committed cumulative carbon emissions from the electricity generation sector and the transition to a green economy. Appl Energy. In press 2016. http://dx.doi.org/10.1016/j.apenergy.2016.02.093
13. CO_2.Earth. Daily CO_2 [database on the Internet]. No date [cited 2016 Apr 16]. Available from: https://www.co2.earth/daily-co2

Genres of science news in the Nigerian press

AUTHOR:
Bankole Falade[1]

AFFILIATION:
[1]Faculty of Linguistics and Literary Studies, University of Bielefeld, Bielefeld, Germany

CORRESPONDENCE TO:
Bankole Falade

EMAIL:
bankolefalade@hotmail.com

POSTAL ADDRESS:
Universität Bielefeld, Fakultät für Linguistik und Literaturwissenschaft, Postfach 10 01 31, D-33501 Bielefeld, Germany

KEYWORDS:
newspapers; content analysis; correspondence analysis; public interest

HOW TO CITE:
Falade B. Genres of science news in the Nigerian press. S Afr J Sci. 2016;112(5/6), Art. #a0157, 3 pages. http://dx.doi.org/10.17159/sajs.2016/a0157

The decade 2001 to 2009 was one of contrasting scientific achievements in Nigeria: while the public responded positively to technology – with telephone penetration increasing from less than 900 000 telephone landlines and a teledensity of 0.73 in 2001 to 111 million Global System for Mobile Telephony (GSM) lines and a teledensity of 63.1 in 2010[1] – the use of the oral polio vaccine was banned in some parts of the country following a revolt premised on a conspiracy against Muslims. The revolt, which resulted in about 5000 new cases, spread the disease to hitherto free neighbouring countries and created an image of a country that is pro-religion and anti-science. This commentary is about the place of science in Nigerian newspapers as a proxy for public opinion.

Newspapers are a functional part of the public sphere: a realm of our social life in which something approaching public opinion can be formed.[2] The idea of a public sphere is that of a body of private persons assembled to discuss matters of public concern or common interest and newspapers, radio, television, the Internet, etc. are the media required for transmitting information and influencing those who receive it.[2,3] The media's role in transmitting information, argues McQuail, makes it an instrument of social power.[4] Being an instrument of social power does not, however, make the media our only source of social reality, but it plays a central role in the thinking society dynamic.[5] The definition of social reality, of facts, norms, values and expectations,[6] is, however, contextual; and as journalists are themselves part of society, how they practise their profession will be largely influenced by their culture, value systems and corporate practices. Livingstone[7], however, argues that audiences are heterogeneous in their interpretation and, at times, even resistant to the dominant meanings encoded in news.

Genres of science news

Newspapers are traditionally categorised by genre – which separates them into classes such as broadsheet, tabloid and business. There are also genres based on content such as news, features, editorial and sports, which cut across news media. Genres thus have similarities in strategies or forms and are recognisable by a set of shared features.[8,9] Genres also share the same communicative purpose for sense making and act as frames for social action or guideposts to explore the unfamiliar.[10,11] For Franzozi[12], the term denotes a class of texts with distinctive and invariant qualities, including an evaluation. There are also genres which describe articles that share implicit characteristics for sense making across sections and newspapers. Moscovici[13] evaluated the articles in his study of psychoanalysis in the French press based on implicit communicative purposes and categorised them into three groups: propagation, propaganda and diffusion. According to Castro and Gomes[5], Moscovici's proposals rest on a recognition of the role of contradiction and the existence of a number of ways of dealing with it. They argue that about a certain social issue, there is 'belief A' and 'belief non-A'. Belief A for this study categorises articles which are highly positive about a scientific issue; belief non-A comprises those articles that are highly critical, while those in between are mainly neutral and present largely balanced views. This format of evaluation is also called sentiment analysis.

The digital age has made the media (print, radio, television and social media) increasingly intertextual. Thus, it has become more important to categorise news across both traditional and new media. Society has also become highly polarised on some issues (nuclear power, climate change, stem cells, etc.) and finding out percentage coverage has become inadequate to categorise news. Multivariate statistical analysis (correspondence analysis, cluster analysis, etc.) has also made it possible to show how all the genres (newspapers, sources, sections, topics, evaluations, etc.) relate to each other, thereby showing the likelihood of a genre sharing similar communicative purpose for sense making with some members of one cluster more than with others. Genres of news content using correspondence analysis can thus categorise news for collective sense making derived from shared and dissimilar features. The concept is stable and statistically reproducible and the approach is often validated by the output.[14]

Using traditional content analysis, the objectives of this Commentary are to examine the following: (1) What was the intensity of coverage? (2) Which fields of science were covered? (3) How were the articles sourced? (4) How were the articles presented? (5) Visualisation of coverage as genres of science news.

Sampling strategy

Schafer's[15] meta-analysis of media coverage of science worldwide found that most authors analysed national elite newspapers or broadsheets owing to their relevance as opinion leaders and inter-media agenda setters. *The Guardian* newspaper, analysed here, is regarded as the flagship of the Nigerian press and an opinion leader in the industry. A relevant article will contain science in at least two paragraphs and categories reported were chosen to answer the research questions. The categories include source of news, story writer, academic field, strategic technology present, primary news value, discourse of benefit and risk, and positive and negative values. The valuation was based on a scale (from +7 to -7) of perceived benefits and projected risk or damage by the technology reported in the article.

The sample size is a cost–benefit issue[16] but Stempel[17] found that increasing the sample size beyond 12 in the sampling of a full year of daily newspapers did not produce marked differences in findings. Thus, for this study, every 25th edition starting from an arbitrarily chosen date in January (to allow for variations in days) was selected, leading to two constructed weeks. The years 2001, 2005 and 2009 were selected for cost–benefit reasons and the newspapers were library hard copies.

Establishing reliability requires some doubling of effort[18] by the same person after a time interval or two people at the same time. A sub-sample was re-coded 1 year after the first coding was done. Different levels of agreement were expected because of the varying number of sub-categories and often overlapping functions of the journalists' beats in the media. Distinguishing between primary and secondary actors for an article about a major oil spill in the

Niger Delta is a case in point. The oil company involved, the Niger Delta and the damage to the environment are strong news leaders presenting the journalist's dilemma of multiple leads. For most categorisations, the level of reliability is often increased when the sub-categories are reduced. Banerjee et al.[19] argue that from 0.40 and above is good agreement beyond chance when there are several categories, with fewer categories necessitating higher values of agreement. For this study, Scott's pi for source of news was 0.67; academic field, 0.92; story teller, 0.91 and strategic technology, 0.97. For the valuation, a scale variable, Spearman rho, was 0.997; exact agreement was 61% and plus or minus 2 points range was 98%.

Findings

What was the intensity of coverage?

This question examines coverage as an indicator of rising or falling interest in science. Is science fashionable in the Nigerian culture? The space allocated in terms of column inches was about 6.8% of the total in 2001, 7.7% in 2004 and 8.3% in 2009. This proportion compares well with the 5–6% devoted to science by the British press[20]; both data are, however, separated in time. The Pew Research Center's annual report[21] on US journalism showed the proportion of science stories was 9.2% (cumulated from science and technology at 1.7%; environment at 1.8% and health and medicine at 5.7%), but the focus was on cover page stories. The interest here is in the direction of the movement and the Nigerian data show increasing science news coverage which supports a public interested in reading about science.

Which fields of science were covered?

The analysis here maps the field of coverage using sub-categories delineated by academic fields such as physics, medicine, social science, technology and engineering. Also important for the analysis is the presence or absence of certain technologies strategic to modern societies such as nuclear power, genetic engineering, the Internet and the environment.

Almost half of the articles published (45.3%) were categorised as strategic (by the definition given for this research). Prominent among these articles were those on energy and telecoms, at 14% each. The telecommunication sector experienced a boom in the review period and perennial concerns over energy supplies also remained high on the public agenda. Thus the prominence of this sector in the news validates the level of attention it received in society. Articles on the environment comprised 6% compared with those on nuclear power which comprised 1.6%. Environmental concerns locally are about pollution in the oil rich Niger delta and the presence of nuclear energy related articles shows the local significance of this topic. About 5% of the articles were about HIV/AIDS, which reflects ongoing concerns about the prevalence of this disease in the country.

The academic field was dominated by technology at 49% and medical at 36%. The dominance of technology was sustained from 2001 to 2009. While the percentage of articles on medicine rose by 3% in 2005, it dropped again in 2009 to a level lower than that in 2001. Technology and engineering articles rose by 6% between 2005 and 2009.

How were the articles sourced?

Official sources – press releases, seminars and conferences – accounted for 45% of articles. Scientific journals and research reports also featured as sources of news. The wire services, including AFP and Reuters, played a substantial role (14.6%) as sources of science news from abroad. The source of news can be read alongside primary news value (novelty, elite or institutions, good or bad news, protest, others) which is an indication of why the article was published. Novelty stories were sourced primarily from scientific journals (30%) and press releases (21%). The elite's main entry point to the news was through press releases (56.8%). Not surprisingly, most good news (61.6%) emerged from press releases.

The Guardian parades a sizeable number of specialist writers, including science writers, in its various sections; but because of the multiple nature of news beats, some science stories originate from political (National Assembly) or business (stock market) sectors. Stories written by science reporters comprised 21%; 59% were written by columnists and non-specialist reporters, while the rest (20%) were written by foreign journalists. More stories were written by science writers in 2009 (28%) than in 2001 (18%). Most (58%) of the contributions of foreign journalists to science in the Nigerian press were in the biological and medical sciences category; physical and engineering accounted for 36%, while arts and social sciences accounted for 6%. The distribution for local journalists was 31% in medical news, 60% in engineering and 9% in arts.

How were the articles presented?

The categories here are more implicit and indicate a valuation of the article. The primary location of consequences includes socio-cultural and moral which accounted for about half (48.5%) of the articles. Science also had applications in economic and financial fields (34.4%) while about 6.9% were in science itself; 8.3% were in environment and ecological and 1.8% in politics and power. Also interesting was the finding that most of the articles (66.2%) promoted the benefits of science, 13% both benefits and risk while 10% focused on risk alone. Similarly, about 83% of the articles were free of any scientific controversy and 80% were about mature and recurring issues and awareness.

The valuation genre (+7 to -7) was based on positive or negative consequences of science using a rating scale. Constructing a stable and reproducible scale is difficult and this study adopted the reach or spread of perceived societal benefits and dangers as proxy. The valuations also represent the two opposing ends of public attitudes to science in society: optimism/progress and pessimism/fear. Only 3% of all articles were in the high negative category, 14% low negative, 75% low positive and 7% high positive.

Visualisation of coverage as genres of science news

Content analysis turns words into numbers[22] and correspondence analysis is an excellent method for analysing its results. The primary goal of correspondence analysis is to transform a table of numbers into a graphical display, thus facilitating its interpretation.[23] There was a significant association between primary news value and valuation ($X^2=367$, $p=0.01$, Cramer's V=0.4); discourse of benefits and risk and valuation ($X^2=459.7$, $p=0.01$, Cramer's V=0.4); and strategic technology and valuation ($X^2=172$, $p=0.01$, Cramer's V=0.3). Cramer's V shows strong enough variation in the data to support a geometric display for the three combinations. A joint relationship of discourse of benefits and risk, strategic technologies and primary news value with valuation can be visualised in a table composed of several two-way cross tabulations. The geometric representation retains as much as possible of the individual spatial arrangements in the simple correspondence analyses.

Dimension one in the output shown in Figure 1 (left to right) contrasts the positive and negative values and this is named the 'valuation dimension' of optimism versus pessimism/fears. The second dimension separates the technologies and issues we are familiar with from those in which the risks/dangers are unknown and this is named the 'familiarity' dimension. The two dimensions account for 93% of the inertia and therefore the error in a two-dimensional interpretation is less than 7%.

Nuclear, environment, cancer and HIV/AIDS overlap between optimism and pessimism on dimension one and familiar and unfamiliar on dimension two. The dual perceptions of these issues are validated by the geometric display. Nuclear power has positive uses in power generation but disasters constantly remind the public of the global nature of their accidents. There has been a lot of progress in preventing HIV transmission but it remains a serious problem locally and stories of cancer are both of survival and death. We can thus derive genres of science news on the basis of familiar and unfamiliar issues and progress/optimism versus fear/pessimism. These categories of genres complement existing ones based on newspaper orientation and news content categories.

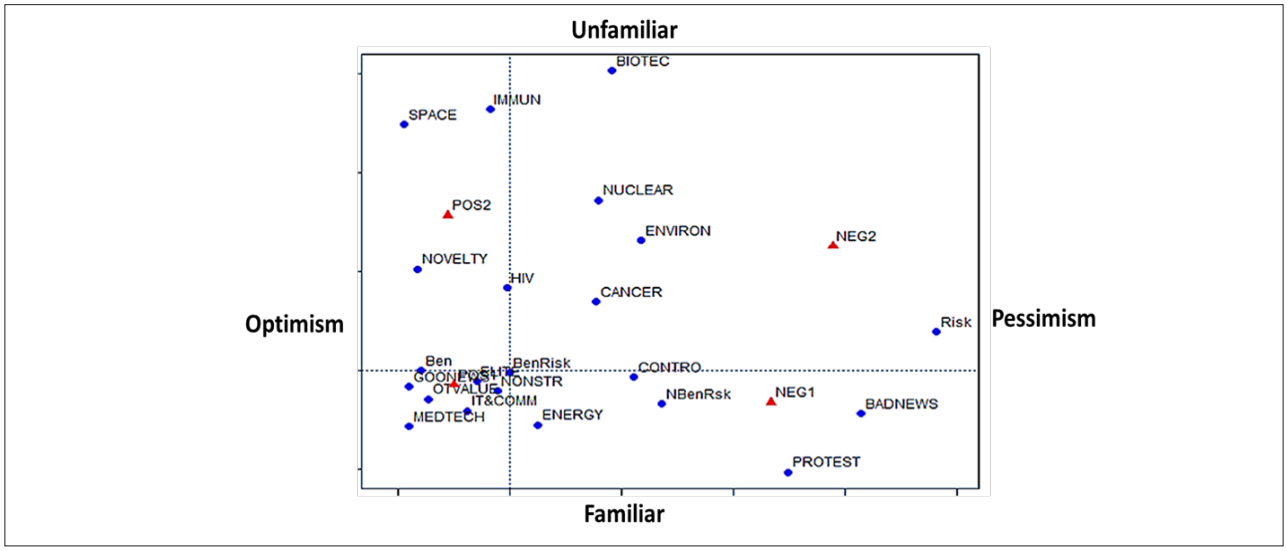


Figure 1: Correspondence analysis biplot of cross tabulations.

References

1. Nigerian Communications Commission. Subscriber statistics [homepage on the Internet]. c2013 [cited 2013 Feb 13]. Available from: http://www.ncc.gov.ng/index.php?option=com_content&view=article&id=125:art-statistics-subscriber-data&catid=65:cat-web-statistics&Itemid=73.
2. Habermas J, Lennox S, Lennox F. The public sphere: An encyclopedia article (1964). New Ger Crit. 1974;3:49–55. http://dx.doi.org/10.2307/487737
3. Fraser N. Rethinking the public sphere: A contribution to the critique of actually existing democracy. Soc Text. 1990;25/26:56–80. http://dx.doi.org/10.2307/466240
4. McQuail D. McQuail's mass communication theory. 6th ed. Los Angeles, CA / London: SAGE; 2010.
5. Castro P, Gomes I. Genetically modified organisms in the Portuguese press: Thematization and anchoring. J Theory Soc Behav. 2005;35(1):1–17. http://dx.doi.org/10.1111/j.0021-8308.2005.00261.x
6. McQuail D. The influence and effects of mass media. In: Curran J, Gurevitch M, Woolacott J, editors. Mass communication and society. Los Angeles, CA / London: SAGE; 1979. p. 70–93.
7. Livingstone S. The challenge of changing audiences or, what is the audience researcher to do in the age of the Internet? Eur J Commun. 2004;19(1):75–86. http://dx.doi.org/10.1177/0267323104040695
8. Miller CR. Genre as social action. Quart J Speech. 1984;70:151–167. http://dx.doi.org/10.1080/00335638409383686
9. Douglas D. The multi-dimensional approach to linguistic analyses of genre variation: An overview of methodology and findings. Comput Hum. 1992;25(5–6):331–345. http://dx.doi.org/10.1007/BF00136979
10. Luzon MJ. Genre analysis in technical communication. IEEE Trans Prof Commun. 2005;48(3):285–295. http://dx.doi.org/10.1109/TPC.2005.853937
11. Bazerman C. The life of genre, the life in the classroom. In: Bishop W, Ostrom H, editors. Genre and writing: Issues, arguments, alternatives. Portsmouth, NH: Boynton/Cook-Heinemann; 1997. p. 19–26.
12. Franzozi R. Quantitative narrative analysis. London: SAGE; 2010.
13. Moscovici S. Psychoanalysis: Its image and its public. Cambridge: Polity Press; 2008.
14. Greenacre MJ, Blasius J. Correspondence analysis in the social sciences: Recent developments and applications. London: Academic Press; 1994.
15. Schafer MS. Taking stock: A meta-analysis of studies in the media's coverage of science. Public Underst Sci. 2010;21(6):650–663. http://dx.doi.org/10.1177/0963662510387559
16. Krippendorf K. Content analysis: An introduction to its methodology. London: SAGE; 2004.
17. Stempel GH. Sample size for classifying subject matter in dailies: Research in brief. Journalism Quarterly. 1952;29:333–334.
18. Bauer M. Classical content analysis: A review. In: Bauer MW, Gaskell G, editors. Qualitative researching with text, image and sound: A practical handbook. London: SAGE; 2000.
19. Banerjee M, Capozzoli M, McSweeney L, Sinha D. A review of interrater agreement measures. Can J Statistics. 1999;27(1):3–23. http://dx.doi.org/10.2307/3315487
20. Bauer M, Ragnarsdottir A, Rudolfsdottir A, Durant J. Science and technology in the British press, 1946 to 1990: A systematic content analysis of the press. In: Schiele B, Amyot M, Benoit C, editors. Proceedings of When Science becomes Culture; 1994 April 11–13; Montreal, Canada. Available from: http://www.cirst.uqam.ca/pcst3/PDF/Communications/BAUER.PDF
21. Pew Research Center. 2011 data sets [data set on the Internet]. c2011 [cited 2011 Oct 09]. Available from: http://people-press.org/category/datasets/2011/.
22. Franzosi R. From words to numbers: Narrative, data, and social science. Cambridge: Cambridge University Press; 2004.
23. Bartholomew DJ, Steel F, Moustaki I, Galbraith JI. Analysis of multivariate social science data. New York: Taylor and Francis; 2008.

Scientific networks in the production of knowledge in South Africa

AUTHOR:
Radhamany Sooryamoorthy[1]

AFFILIATION:
[1]Sociology Programme, School of Social Sciences, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Radhamany Sooryamoorthy

EMAIL:
sooryamoorthyr@ukzn.ac.za

POSTAL ADDRESS:
School of Social Sciences, MTB, Howard College Campus, University of KwaZulu-Natal, Durban 4041, South Africa

KEYWORDS:
productivity; professional activities; collaboration; networks; publication

HOW TO CITE:
Sooryamoorthy R. Scientific networks in the production of knowledge in South Africa. S Afr J Sci. 2016;112(5/6), Art. #a0155, 3 pages. http://dx.doi.org/10.17159/sajs.2016/a0155

It is now recognised that the systems of knowledge production have become a hybrid network system.[1] This transition is an indication that professional contacts and networks are important in the production of scientific knowledge. Networking is indispensable in countries where scientific advancement is both a concern and a priority. Scientific networks are able to expand the horizons of knowledge sources and stimulate the exchange of ideas that lead to the creation of more new knowledge.[2] Applied as a strategy in research, networking builds crucial professional contacts to improve the potential to produce knowledge.[3] The scientific capital of a scientist tends to increase as the number of scientists with whom they connect increases.[4]

Professional networks turn out to be meaningful and beneficial to scientists in several ways. Firstly, professional contacts and networks assist scientists in forging new research alliances. These new research alliances eventually lead to the co-production of knowledge. Secondly, scientific networks are the building blocks in research which serve the function of disseminating information.[5] Thirdly, networks improve the possibilities of expanding the capacities of knowledge production.[6] Fourthly, networks can widen the circle of contacts, which leads to extensive participation in scientific matters, and to collaborative research enterprises.[6]

Literature on scientific networks enables us to conceptualise social structures as a set of ties and helps us to focus on the characteristics of the ties in the scientific system.[7] Professional contacts are intentionally established by scientists, with a view to conduct, produce and publish new knowledge. The network of ties that is involved in scientific research influences not only the quantity but also the quality of production. It has been found that the greater the number of international ties of an author, the greater the likelihood that their work will be widely cited.[8] Expected citation of a publication depends largely on the number of authors and the number of countries that participated in the production.[8]

Networks are often discussed in relation to collaboration. Collaboration consists of nodes (research units such as scientists, institutions and countries) and relational ties (collaborative relations).[9] Despite the importance of networks in scientific research, the impact of the networks on the publication productivity of scientists has not been adequately investigated,[10] more so in Africa.

This essay presents the findings of a series of empirical studies conducted in South Africa over the past 10 years.[3,11-14] It focuses on scientific networks from the perspectives of collaboration and research productivity.

Scientists create and maintain networks to advance their research outcomes. Institutions often are positioned in the central and peripheral positions in co-authorship networks.[9] Networks have been used as a proxy variable of collaboration in the study of South African scientists.[15] The importance of social and professional networks in co-authorship has also been illustrated.[16] In some cases, productivity, networks and research activities occur prior to the start of collaboration.[17]

Although there is a large reserve of empirical studies on scientific networks, theoretical models that explain them are rather inadequate.[18] The network formation game is an exception. This model presents how network formation materialises as individuals make decisions in their scientific relationships through a link formation game.[18] The decision of researchers to form a collaboration link with talented researchers is in line with a trade-off between rewards and costs involved in collaboration. The model is particularly useful in grasping the conditions under which researchers organise themselves into unequal and hierarchical scientific networks.

The formation of scientific networks is considered to be a strategy of academics[19] that involves co-authoring and support networks. The number of strong professional ties in the support network is related to the quality of publications as well.[19]

Relationship between scientific networks and scientific productivity

There are several assumptions regarding the relationship between scientific productivity and networks. One is that scientific productivity is correlated with the existing networks that scholars maintain in their career. The results of the South African studies used in this essay support the hypothesis that the productivity of academics and scientists (i.e. the production of edited books, co-production of papers in national journals and co-authored books in particular) is positively associated with their scientific networks (total number of networks, total network locations, countrywide network, local networks, domestic networks and international networks) of all the respondents in the sample.

The relationship between scientific productivity and scientific networks varies between academics (in universities) and scientists (in research institutes). The results from the studies showed that the production of edited books was significantly related to the total network size and total network locations. A number of network variables (total networks, total network location, countrywide networks and domestic networks) were significant in their association with the co-production of papers in national journals. For scientists in research institutes, no association between any productivity or network variables was evident. The relationship between scientific productivity and scientific networks varied across the sample, i.e. academics and scientists. The connection between productivity and networks was not as prominent for scientists as for the academics in South Africa.

There is a view that professional activities of scholars are positively associated with their existing scientific networks; the data support this positive association with some key network variables. This relationship was not

manifest in the case of scientists in research institutes, as much as it was significant for all scholars and academics in universities.

The study of South African academics and scientists has also uncovered the inter-relationships among scientific research, networks, professional activities and productivity. The South African scientific system holds characteristics that are revealed in the dimensions of professional activities, scientific networks and the production of knowledge. Scientific networks are essential for the professional activities and productivity of those who are engaged in scientific research. The scientific productivity and networks of the scholars, in many instances, were positively related. Measured under several indicators, the scientific productivity of scholars is seemingly relevant in influencing network variables. An increase in the size of the networks, particularly that of international links, is instrumental in the increase in this specific publication productivity. The relationship between the size of networks based in countries outside South Africa emphasises the international links the respondents maintain in their career. These links will also have positive effects on the professional activities of scholars. One additional dimension of this correlation between productivity and networks is co-publication.

Some of the co-publication productivity measures have connection with network measures. Evident also is the prominence of local and domestic networks in the co-publication outputs of the respondents. Scholars were found to maintain their networks with colleagues and peers, both in the region and within the country. In a way, this is the effect of the domestic collaboration that the respondents currently pursue in their professional life. The increased extent of contacts and networks, both within and outside the country, led to increased levels of productivity. Co-publications and networks are stronger than sole publications and networks.

Scholars are found to be different in their productivity and in the nature and size of scientific networks. For academics, the total network size and total network locations were crucial in the production of edited books. In the co-production of papers in national journals, the size of total networks, network locations, countrywide networks and domestic networks were applicable to academics, but not to scientists in research institutes. Co-production of papers in national journals brings in local collaboration. The increased size of the networks that were locally based (local, countrywide and domestic) had an influence on the co-production of publications in local journals. The contacts and networks academics built within their domestic surroundings have resulted in production that is local, but not international.

The above pattern in the production and co-production of research publications by academics is different from that of scientists in research institutes. None of the scientific network variables has shown any effect (or cause) on the productivity of scientists in research institutes. Networks seem to have no influence at all on the scientific productivity, individual or co-production, among scientists. This finding does not mean that scientists in research institutes do not produce at all. The contrast between academics and scientists in both the total productivity and co-productivity was very prominent. One explanation for this contrast is that scientists in research institutes are not expected to publish but only to conduct research. Publication is not a primary concern but research is. For academics, publications are crucial for their career advancement. In the current context in which publications are supported through government funding, universities encourage academics to be actively productive. There is an expectation at the national level that the academics, depending on their rank, produce a particular number of papers every year. The same expectation does not exist for scientists in research institutes in the country, whose core activity is limited to conducting research. This being the case, scientists are not keen to maintain their contacts for the purposes of producing publications in future. When there is a possibility or potential to conduct research together and to publish jointly then the contacts and networks become crucial and are therefore maintained. Normally, for scientists, their research itself is the end product. Other than personal motivation, there is no incentive for scientists to publish their research findings in a form other than reports. The research they conduct is either institution-based or inter-institutional but is ring-fenced within the country. Networks are therefore irrelevant for them in advancing their productivity or co-productivity.

Professional activities can lead to the creation (and maintenance) of networks. Attendance of seminars and conferences and the review of papers and books present opportunities to connect with peers. Most of the network variables in the studies reported here were positively associated with such professional activities. Network size, network location size, the size of the networks within the country, and international networks were connected to their professional activities. Professional activities are crucial to one's expanse of scientific networks, as scientific networks are also important in productivity. An explicit and straightforward relationship between many of the professional activities and networks among the respondents in general and academics in particular was discernible. In contrast, the same type of relationship did not exist for all the network variables among scientists in research institutes. Except in continental and international networks, professional activities are irrelevant for scientists. One reason for this trend among scientists versus academics is that the academics are more professionally active than scientists. They are more likely to avail themselves of professional opportunities as these are indispensable for their professional and career advancement. Academics participate in professional associations; organise and attend seminars, conferences and workshops; peer review for journals and publishers; and serve on committees. These activities put academics in an advantageous position to expand their connections with peers in their field, both nationally and internationally. Eventually these activities get translated into networks, contributing to their network contacts and productivity. Because of the nature of the work and the pressure of time-bound research projects, scientists are more focused on their research than on professional activities. Spending time on work such as peer reviewing or other professional activities is therefore not in their career interests. They do, however, attend seminars and conferences, although not as frequently as their academic colleagues.

The results of the studies[3,11-14] confirm the significance of variables that influence the scientific networks of academics and scientists. The size of the networks could be predicted on the basis of the professional activities. The size of the total network locations was also dependent on professional activities, namely, meetings. Also significant in the size of the networks was the amount of time they spent on research pursuits and total productivity. These findings contribute to the understanding of the productivity, professional activities and scientific networks of academics and scientists in South Africa.

There is a close association between professional activities, scientific networks and productivity. The amount of time scholars usually spend on research-related activities is reflected generally in the diversity of networks they have. The contrast in the networks between the two sets of respondents, academics and scientists, is obvious.

It is appropriate to compare the situation in South Africa to that reported by other studies elsewhere. In a study of US academics, it was observed that the frequency of the use of networks was obvious in the benefit of scholarly productivity.[20] The greater the number of collaborators in a network, the greater the opportunity to access knowledge and skills.[21] Correlation between production and collaboration – a proxy for networks – in a specific discipline has also been reported.[6] Highly productive authors have higher than average rates of collaboration.

Studies that concur with the South African case showed that there is a positive and significant correlation between production and networks.[22] A positive relationship between the output of authors and the centrality measures of authors is supported.[22] Considering the network as a form of social capital, scientists employ a variety of networks configured in a way to enable their work.[23] This has positive consequences for their productivity. In a university centre in the USA, professional linkages and network ties are instrumental in harnessing resources for productivity.[24]

The strong correlation between international research contacts and the publishing activity of scientists has also been established.[25] It has also been demonstrated that networks through electronic means of communication increase productivity, and collaborations with strong ties are more productive.[26]

The above analysis on the scientific system of a developing country like South Africa, with its scientific edge over other African countries, offers some insight on the growth of scientific production. South Africa is a leading country in the production of scientific knowledge in sub-Saharan Africa.[27] For South Africa to expand and consolidate its position in the production of scientific knowledge, due consideration of the factors identified is important. Measures to encourage professional activities, networking and collaboration can lead to a further increase in productivity of academics and scientists in the country.

The findings have significance beyond the borders of South Africa, particularly for other African countries. Building networks of contacts that eventually lead to collaborative research enterprises of scholars is to be viewed as a strategy for the production of knowledge. The linkage between collaboration and productivity further strengthens the need for professional contacts and networks in scientific research. This study of the South African research system once again reinforces the connection among networks, collaboration and production of knowledge.

Acknowledgement

Funding received from the National Research Foundation (South Africa) to conduct the studies reported in this essay is gratefully acknowledged.

References

1. Heimeriks G, Horlesberger M, Besselaar PVD. Mapping communication and collaboration in heterogeneous research networks. Scientometrics. 2003;58(2):391–413. http://dx.doi.org/10.1023/A:1026296812830
2. Lee DH, Seo IW, Choe HC, Kim HD. Collaboration, network patterns and research performance: The case of Korean public research institutions. Scientometrics. 2012;91:925–942. http://dx.doi.org/10.1007/s11192-011-0602-8
3. Sooryamoorthy R. Internet technology and the epistemic strategies of scientists in post-apartheid South Africa: Race as a decisive factor. Tec Soc. 2015;40:73–81. http://dx.doi.org/10.1016/j.techsoc.2014.08.003
4. Jonkers K, Tijssen R. Chinese researchers returning home: Impacts of international mobility on research collaboration and scientific productivity. Scientometrics. 2008;77(2):309–333. http://dx.doi.org/10.1007/s11192-007-1971-x
5. Corley EA, Boardman PC, Bozeman B. Design and the management of multi-institutional research collaborations: Theoretical implications from two case studies. Res Pol. 2006;35(7):975–993. http://dx.doi.org/10.1016/j.respol.2006.05.003
6. De Granda-Orive JI, Villanueva-Serrano S, Aleixandre-Benavent R, Valderrama-Zurián JC, Alonso-Arroyo A, García Rio F, et al. World-wide collaboration among medical specialties in smoking research: Production, collaboration, visibility and influence. Res Eval. 2009;18(1):3–12. http://dx.doi.org/10.3152/095820209X393163
7. Otte E, Rousseau R. Social network analysis: A powerful strategy, also for the information sciences. J Inf Sci. 2002;28(6):441–453. http://dx.doi.org/10.1177/016555150202800601
8. Goldfinch S, Dale T, DeRouen K Jr. Science from the periphery: Collaboration, networks and 'periphery effects' in the citation of New Zealand Crown Research Institutes articles, 1995-2000. Scientometrics. 2003;37(3):321–337. http://dx.doi.org/10.1023/A:1025048516769
9. Chang H-W, Huang M-H. Prominent institutions in international collaboration network in astronomy and astrophysics. Scientometrics. 2013;97:443–460. http://dx.doi.org/10.1007/s11192-013-0976-x
10. Ynalvez M, Shrum W. Professional networks, scientific collaboration, and publication productivity in resource-constrained research institutions in a developing country. Res Pol. 2011;40:204–216. http://dx.doi.org/10.1016/j.respol.2010.10.004
11. Sooryamoorthy R, Shrum W. Does the internet promote collaboration and productivity? Evidence from the scientific community in South Africa. J Comput Mediat Commun. 2007;12(2):733–751. http://dx.doi.org/10.1111/j.1083-6101.2007.00347.x
12. Sooryamoorthy R. Publication productivity and collaboration of researchers in South Africa: New empirical evidence. Scientometrics. 2014;98:531–545. http://dx.doi.org/10.1007/s11192-013-0990-z
13. Sooryamoorthy R. Transforming science in South Africa: Development, collaboration and productivity. Basingstoke and New York: Palgrave Macmillan; 2015. http://dx.doi.org/10.1057/9781137493071
14. Sooryamoorthy R. Producing information: communication and collaboration in the South African scientific community. Inf Commun Soc. 2016;12(2):141–159. http://dx.doi.org/10.1080/1369118X.2015.1047392
15. Durbach IN, Naidoo D, Mouton J. Co-authorship networks in South African chemistry and mathematics. S Afr J Sci. 2008;104:487–492. http://dx.doi.org/10.1590/S0038-23532008000600024
16. Newman MEJ. Coauthorship networks and patterns of scientific collaboration. Proc Natl Acad Sci USA. 2004;101(suppl.1):5200–5205. http://dx.doi.org/10.1073/pnas.0307545100
17. Yoshikane F, Nozawa T, Shibul S, Suzuki T. An analysis of the connection between researchers' productivity and their co-authors' past attributions, including the importance in collaboration networks. Scientometrics. 2009;79(2):435–449. http://dx.doi.org/10.1007/s11192-008-0429-8
18. Rubí-Barceló A. Core/periphery scientific collaboration networks among very similar researchers. Theo Dec. 2012;72:463–483. http://dx.doi.org/10.1007/s11238-011-9252-9
19. Sebestyén T, Varga A. Research productivity and the quality of interregional knowledge networks. Ann Reg Sci. 2013;51(1):155–189. http://dx.doi.org/10.1007/s00168-012-0545-x
20. Cohen J. Computer mediated communication and publication productivity among faculty. Int Res. 1996;6(2/3):41–63. http://dx.doi.org/10.1108/10662249610127328
21. Defazio D, Lockett A, Wright M. Funding incentives, collaborative dynamics and scientific productivity: Evidence from the EU framework program. Res Pol. 2009;38(2):293–305. http://dx.doi.org/10.1016/j.respol.2008.11.008
22. Hou H, Kretschmer H, Liu Z. The structure of scientific collaboration networks in scientometrics. Scientometrics. 2006;75(2):189–202. http://dx.doi.org/10.1007/s11192-007-1771-3
23. Bozeman B, Dietz J, Gaughan M. Scientific and technical human capital: An alternative model for research evaluation. Int J Technol Manage. 2001;22(7–8):636–655. http://dx.doi.org/10.1504/IJTM.2001.002988
24. Ponomariova BL, Boardman PC. Influencing scientists' collaboration and productivity patterns through new institutions: University research centers and scientific and technical human capital. Res Pol. 2010;39(5):613–624. http://dx.doi.org/10.1016/j.respol.2010.02.013
25. Kyvik S, Larsen M. International contact and research performance. Scientometrics. 1994;29(1):161–172. http://dx.doi.org/10.1007/BF02018388
26. Walsh JP, Maloney NG. Computer network use, collaboration structures and productivity. In: Hinds P, Kiesler S, editors. Distributed work. Cambridge, MA: MIT Press; 2002. p. 433–458.
27. Pouris A. Science in South Africa: The dawn of a renaissance? S Afr J Sci. 2012;108(7/8), Art. #1018, 6 pages. http://dx.doi.org/10.4102/sajs.v108i7/8.1018

BOOK TITLE:
Run racist run: Journeys into the heart of racism

AUTHOR:
Eusebius McKaiser

ISBN:
9781928257158 (softcover)

PUBLISHER:
Bookstorm, Johannesburg; ZAR240

PUBLISHED:
2015

BOOK TITLE:
What if there were no whites in South Africa?

AUTHOR:
Ferial Haffajee

ISBN:
9781770104402 (softcover)

PUBLISHER:
Picador Africa, Cape Town; ZAR282

PUBLISHED:
2015

BOOK TITLE:
Raising the bar: Hope and renewal in South Africa

AUTHOR:
Songezo Zibi

ISBN:
9781770103689 (softcover)

PUBLISHER:
Picador, Cape Town; ZAR221

PUBLISHED:
2015

BOOK TITLE:
We have now begun our descent: How to stop South Africa losing its way

AUTHOR:
Justice Malala

ISBN:
9781868426799 (softcover)

PUBLISHER:
Jonathan Ball, Cape Town; ZAR246

PUBLISHED:
2015

REVIEWER:
Iain Edwards

EMAIL:
edwards.researchsa@gmail.com

AFFILIATION:
Gordon Institute of Business Science, University of Pretoria, Johannesburg, South Africa

POSTAL ADDRESS:
GIBS, PO Box 787602, Sandton 2146, South Africa

HOW TO CITE:
Edwards I. Rising expectations. S Afr J Sci. 2016;112(5/6), Art. #a0153, 3 pages. http://dx.doi.org/10.17159/sajs.2016/a0153

Rising expectations

In 1953 Arthur Keppel-Jones, the pioneer of the South African alternate, predictive and apocalyptic history genre, published a work entitled *When Smuts goes: A history of South Africa from 1952–2010, first published in 2015*[1]. In his Introduction, Keppel-Jones makes three telling points: South Africans resist asking hard and difficult questions; South Africans show irresponsible optimism in the future by choosing oft-tried but ill-tested views as future solutions; and that 'the salvation of the country can lie only in a reversal of historic[al] tendencies, a reversal so thorough as to constitute a revolution' (p. xi–xii). The four books reviewed here – all of which were published in 2015 – are important contributions to current public debates. All of the authors are well-known professional media figures. Ferial Haffajee is the editor of *City Press*, Songezo Zibi is the immediately past editor of *Business Day* and Justice Malala and Eusebius McKaiser are noted public analysts. Two of the books – those by Haffajee and McKaiser – concern identity politics and race. The books by Malala and Zibi focus on politics, leadership and public policy. The books differ in approach and analysis, yet raise common themes.

Run racist run: Journeys into the heart of racism

McKaiser's work is an anthology of essays, presented through the heuristic device of storytelling. Argument and analysis is through the logic of debate, with McKaiser – who, from what details he provides, is a first-generation graduate from a Grahamstown-based family – taking readers through his reasoning. The core argument is that whilst not all white people were perpetrators of 'anti-black racism', they all benefitted, as they still do. McKaiser refers to this as 'unearned privilege'. As white people evince 'degrees of racism' there are 'degrees of whiteness' and 'privilege'. White people are blind to their racism, and now rationalise their ignorance by viewing black people as 'race obsessed' (p.18, 63, 67, 69 & 126). Key people in the stories are young black university students and adult white liberals. McKaiser focuses on the apparent world views of these two identities, and notes how these two personas appear to need each other. McKaiser does not accept this situation and insists that white people work through their own racism. McKaiser is at his strongest noting the fallacy of the 'Rainbow Nation' (p. 124), and hypocrisy in public debates (p. 147 & 208–209). But there are problems. Neither of these points is new. McKaiser is skilled in unpicking public debates, but does not actually dissect and analyse – he asserts and notes, often to make a sweeping statement. McKaiser is asked by 'Kate', a 'world-class expert' studying organisational cultures in a large South African corporate, to offer professional advice on her research and findings. The study contains important insights into hierarchies of power, and the disjuncture between corporate policy culture and the actual dynamics of white male corporate power (p.129–147). McKaiser then asserts: 'The results of the research were not surprising at all. In fact she could be talking about *any company or institution in our country*' [my emphasis] (p. 130). McKaiser argues that black thinkers must express their anger in their writing. If so, where does violence come into public debate and behaviour? All too often McKaiser concentrates on exploring, revealing and confirming his own train of thought. Why is there not more hard evidence? Such evidence is easily available: media, particularly social media, provides abundant evidence for really getting to grips with race, identity, historical denial and revisionism, youth politics – including #RhodesMustFall and #FeesMustFall – and post-1994 South Africa. The book has few footnotes, no bibliography and no index.

In 2015, as McKaiser was writing this book, the South African Institute of Race Relations conducted extensive research into race relations in South Africa. Published in February 2016, the report[2] provides statistical evidence pointing to a vastly more complicated and positive picture than that offered by McKaiser. What is abundantly clear is that McKaiser is angry – and triumphal. He asserts that '(t)he year 2015 is spectacular testimony to active young citizenship, and the overshadowing of lazy elders whose time has expired' (p. 61).

What if there were no whites in South Africa?

Haffajee's enticingly entitled book is engaging, complex and thought provoking. At the outset it is important to note, from the title, what the book is not about. Because, curiously, this defines what the book is about. 'What if there were no whites in South Africa' is a popular refrain, with long provenance. On the one hand it is part of white South African triumphalist assertion: 'If it weren't for us…'. In post-1994 South Africa, such a refrain is used as racist self-confirming prophesy and as historical denial. From European settlement onwards, modernity came to southern Africa, but white people continually sought all means to deny black society access to the many tangible benefits of this modernity. The refrain is also part of a black refrain, of alternative history, imagining a country developing and progressing without imperialism and white people.

Haffajee, born and raised in the plebeian mixed-race Johannesburg suburb of Bosmont, is a first to graduate – from Wits – daughter in a Muslim family of tailors. She has an ambivalent attitude towards her alma mater. She studied at Wits during the politically charged late 1980s, when criss-crossing strands of politics flowed both to the United Democratic Front and Black Consciousness. She drew political sustenance from both.

Haffajee provides a complexly interwoven argument about how black professionals must actively fashion a better South Africa. Her work is a philosophically phrased love song for such a future *Mzanzi* – a popular term for South Africa, derived from the *isiXhosa* word for 'south'. Quite openly admitting that she is a daughter of the 'Rainbow Nation', with huge respect for Nelson Mandela, Haffajee also proudly proclaims that she is a benefactor of black economic empowerment. But then she looks around present-day South Africa and finds much wanting. As a good investigative journalist she listens to what is being said, and hears angry views, which are discordant to her own. She sets up 'meetings' which have the pedagogical virtues of tutorials, seminars and interviews, with Haffajee wonderfully cast as listener, reader of new information, interviewer, interlocutor and critic. These engagements are properly listed in the bibliographical references. She finds a new intellectual trend: that of whiteness studies and its associated term – 'white privilege'. She learns of its intellectual origins, in late 20th century America, and reads works by its major proponents. She cites much of this work. Curious and wishing to

understand, she asks versions of a single question: 'Why, in post-1994 South Africa is 'whiteness studies' being imported and transplanted?'; 'Why are academics and social activists in contemporary South Africa, where black people are in the majority, using social theories of importance and influence amongst black minorities in America?' As Haffajee realises, the importers feel angry and alienated still seeing themselves as white subjects. So they question what they see as fundamentals: 1652 (the arrival of Jan van Riebeeck), 1990 (President de Klerk's 2 February speech unbanning liberation politics, Mandela walking free, and the commencement of political negotiations) and 1994 (South Africa's first one-person one-vote general election and the formation of President Mandela's government of national unity); the liberation movement's elder's conceits, conferring, from upon high statuesque moral vantage points, a 'freedom' on the so-called 'born-frees'. She also finds that these young black people, from aspirational plebeian and middle-class families, are often first-time-in-family students at tertiary institutions. They are confronting what she experienced at Wits. This is the massive edifice of inherited white wealth, privilege, and the cultural assumptions and expectations made and asserted by the sons and daughters of such financial and professional solidity and certainty. She can see cogency in these arguments. But, she finds the whiteness thesis too all-assertive, all-encompassing, all-explaining, and thus *dis*-empowering. Again she surveys the South Africa around her, which includes white historical denial and a lack of genuine racial and gender diversity in private sector corporate worlds. But she also notes, cogently, that power in South Africa is in black hands. Black professionals have to be the agents of substantial positive change. Here is an honest, engagingly personal, philosophically reflective journey through contemporary South African social thought.

Raising the bar: Hope and renewal in South Africa

Zibi's work is precisely what Haffajee's book is not – Haffajee never mentions President Zuma – and precisely what Haffajee recommends: incisive reflection and analysis of post-1994 South Africa and its future. From the few biographical details given, Zibi may also be a first-in-family graduate, with a strong familial lineage in liberation politics; Zibi also has impressive professional experience in South Africa's corporate sector. Zibi's argument is cogent and his presentation – with footnotes, bibliography and index – is professionally faultless.

The central thesis is clear. South Africans have lost the capacity to go forward. No South African leader, or political party, has sought to address the country's past and how it lives on, often in their very own actions (p. 55). Old narratives of change and progress have run dry. South Africans have lost their faculties of reason and the ability and will to conduct rational and responsible public discussions, because of the knots of identity politics, moral relativism (p. 10 & 38) and race politics – all in the interests of fighting an old battle about to whom South Africa belongs. Zibi is very perceptive when presenting and analysing white and black fallacies of thought. He does this in an analysis of how popular informed analysis of private corporate business culture is bereft of reason, through moral relativism (p. 43–44). Similarly, Zibi, bringing to bear powerful scholarly comment, argues that 'racial reasoning should be replaced by moral reasoning'. Hence his trenchant critique of the #RhodesMustFall movement (p. 45, 48–50) and 'whiteness studies':

> *Excessive focus on whiteness detracts from the responsibility to do what is right in black communities, for black communities and in a way that takes into account their own peculiarities, historical and contemporary. This is because if left unchecked all it does is make every problem the fault of white people, and every solution the responsibility of white people. It relegates black people to being perpetual whiners where the efforts of white people are both demanded and resented at the same time, where nothing is ever right and no solution is good enough.* (p. 78)

Zibi provides concrete suggestions as to how South Africa must move forward and searing critique of the moral and political failings of South African leadership, across the spectrum (p. 82–105). Central to his understanding of a path forward is the notion of the greater common good (p. 16, 34, 37 & Chapter 3). This is the philosophical underpinning of what South Africa needs to become: an inclusive political and economic polity and society.

We have now begun our descent: How to stop South Africa losing its way

Malala's work is a powerful complement to Zibi's. As with Zibi, his argument and analysis is excellently expressed and backed up with extensive reading, all of which is properly set out. As with Haffajee and Zibi, Malala shows his depth of reading in a scholarly work, demanding attention.

Alarmed at the collapse of the liberation narrative, Malala honestly and importantly notes two key markers in his personal attitude towards contemporary South Africa: 'The Spear' incident and the Marikana massacre, both in 2012 (p. 7 & 102). Malala reflected in the aftermath of artist Brett Murray's 'Hail to the Thief' exhibition including 'The Spear' painting of President Zuma, penis exposed, at the Goodman Gallery in Johannesburg. At first he was angry, seeing this as yet another racist attack on the 'black body' – another in a long and cruel history dating to colonial times. As the public controversy drew on, Malala reflected further – on constitutionalism (p. 65 & 84–85). Then he found Zuma wanting (p. 60–66). For Malala, the ANC and South Africa have a 'Zuma problem'. President Zuma must either stand down, in the best interests of the country, or the ANC must recast its moral authority and dismiss him (p. 42). Not only is Zuma personally corrupt, but he scorns constitutional democracy, destroys its vital institutions – including the media – and has recast the ANC and MK security organs of the exiled armed struggle time to retain his control over both the ANC and the state (p. 91–98 & 107). Further, Zuma has surrounded himself with cronies, rent seekers and a predatory elite bent on state capture (p. 15–16, 23, 26, 30, 32–34, 113–115 & 145).

Malala is not saying that problems began with Zuma, and will end with Zuma's going. As with Zibi, he acknowledges the failings of Mandela's presidency (p. 48). He acknowledges huge 'unprocessed anger' and the chasm between this and the dreamers: Mandela and Tutu (p. 63–64). Nor is Zuma's removal as president the palliative. Indeed quite the opposite. For him, and Zibi, South Africa is dominated by a political elite – a point first made a few years back by then COSATU Secretary General Zwelizima Vavi (p. 138). And Malala acknowledges and sets out how manifold are South Africa's problems, and how 'bitter' the 'medicine' must be (p. 15–16). Yet there is optimism in much of Malala's understandings. Malala cites as evidence two issues. Firstly, he comprehends the ANC's tremendous 'moral authority', built over years of struggle, and how it managed, against all odds, to speak to and for ordinary people, of need and aspiration. That history must and can be recovered (p. 156–157). Secondly, Malala points to the period from 1990 to 1994 in which despite all odds and powerful hotheads – on all sides – political peace was fashioned, and from few instruments (p. 91 & 215). Indeed when dealing with the collapse of political trust in contemporary South Africa, Malala exhorts the reader to remember those desperately uncertain and violent times. In this context the #RhodesMustFall movement is found wanting for its simplistically linear understanding of politics, and its lack of rigorous historical understanding (p. 66).

Whilst seeing historical blame as toxic in present-day politics, Malala is cutting on continued apartheid denialism (p. 60 & 68). When Malala was a youngster, his family were forcibly relocated to a barren desolate area called Eersterus in the Hammanskraal area north of Pretoria. These families were amongst the millions of people who were the victims of grand apartheid's 'homeland consolidation' [sic] policy: for the creation of a 'white South Africa' and, in Malala's family's case, the Bophuthatswana homeland. The community hired their own teachers and were ever vigilant in ensuring their children received the very highest standards of schooling. Indeed, Malala, a graduate of UCT, may well also be a first-in-family graduate. In taut writing Malala tempers his anger

when telling contemptuously of how in 2012 F.W. de Klerk told CNN's Christiane Amanpour that 'blacks were not disenfranchised, they voted. They were not put in homelands, the homelands were historically there' (p. 46–49 & 62).

How to define future scenarios? These are provided, as are debating points (Chapters 7, 13 & 14 and Chapter 7 epilogue). Malala cautions against South African's often mythical 'exceptionalisms'. So too his assessment of the BEE balance sheet is less positive than is Haffajee's (p. 109–112). As with Haffajee and Zibi, Malala seeks reasoned public debate, without 'blame' (p. 67 & 72). Leaders will not emerge through miracles (p. 148); they will emerge through debates. As with Zibi, for Malala, the central themes in this public debate are standards of ethical leadership and the notion of a 'common good' (p. 39). He acknowledges his own failings in taking a line of least resistance against Zuma-inspired state threats to media freedom (p. 82–83 & 86). The ANC has lost its intelligentsia – they are now outside the movement – and probably a large proportion of its historical middle-class loyalists too (p. 33–34 & 137). This is also the case with the black consciousness movement. His vision is of an ANC confronting its 'elephant in the room' – Zuma and the crisis he has brought on – as other leaders, some of whom have yet to show themselves, clearly chart a new course (p. 43 & 98). These interesting scenarios are not particularly new, with two exceptions. For Malala, current Deputy President Cyril Ramaphosa is too compromised and does not display the strong visionary leadership now required (p. 141–142). And he sees soon to be stepping down Public Protector Thuli Madonsela as a vitally new, ethically and morally courageous leading force in future politics (Chapter 12).

Conclusion

The century-long era of the ever expanding South African patron state – from Milner's Reconstruction to Zuma's Developmental – is showing its finite limits. South Africa is now in the third phase of its post-1990 development. The first, from 1990 to 1996, produced political rapprochement, peaceful and legitimate elections, and a remarkable constitution. The ANC's long-held narratives on historical change and public policy as set in place in 1994 have yielded much, but successes and failures have born new dynamics. This third phase will test South Africa's ingenuity and resourcefulness. Dramatic policy changes are required. And there is urgency. In the genre pioneered by Keppel-Jones, analysts like R.W. Johnson have already incisively set out doom-laden narratives. In 2011, political analyst Moletsi Mbeki set a deadline: 2020 – then South Africa will experience its 'Tunisia Day' 'when the masses rise up against the powers that be'. This is when China's resources-driven economy changes, becoming less dependent upon commodities imported from, amongst others, South Africa (Malala, p. 213).

From the mid-1960s onwards, with the ANC and the PAC and other organisations outlawed and their members increasingly mostly in prisons or exile, politically aware and spirited cross-generational men, largely but not exclusively black, gradually formulated new questions. They had an international outlook, many were graduates – largely first-in-family, with, some having studied or travelled abroad. They were widely read: from and of Bonhoeffer, Fanon, Gandhi, Gutiêrrez, Martin Luther King Jnr, Marx and Nkrumah, and many others. They asked, in summary, two very non-linear questions of themselves: 'How to live in South Africa and challenge its very fundamentals?' and 'How to make the essence of oppression the seed of liberation?' From this milieu came the Black Consciousness Movement, the Christian Institute, the revived Natal Indian Congress, the Study Project on Christianity in Apartheid Society, the Worker Benefit Fund – the forerunner of the Federation of South African Workers – and South African liberation theology. By the time of the United Democratic Front and later difficult discussions on a possible negotiated political settlement, they used the same non-linear and transcendent approach. Here came a political strategy: accept intrinsic compromises with their dangers, and a personal principle: in thought and deed show yourself better than the racist. Some remember this still unacknowledged but vital part of South African political philosophy as 'rising expectations'. The title comes from economist and futurist Robert Theobald's[3] hugely influential book, published in 1960 and read by many, black and white, on segregated campuses throughout the country. However, in their usage these new political leaders gave the term an added behavioural quality, thereby benchmarking their own self-embraced understanding of an imperative defining anti-apartheid leadership.

In 1992 when formally welcoming the first sitting of one of the ANC's national commissions, then ANC Deputy Secretary General Cheryl Carolus and Commission Chair, and acclaimed novelist and exiled liberation cadre, Mongane Wally Serote spoke of how, in thought and deed, commission members had to rise above racists in order to show moral authority. Significantly, Joel Netshitenzhe – ANC NEC member and veteran, and head of the Mapungubwe Institute for Strategic Reflection – recently publicly re-stressed the importance of such an attitude: 'Black South Africans ought to set their sights a lot higher than equality with whites: they should work towards a new civilization'[4]. In decades to come, people may just reflect on this now still nascent present phase of South Africa's post-1990 history as being the moment of rising expectations, for the greater common good.

References

1. Keppel-Jones A. When Smuts goes: A history of South Africa from 1952–2010. First published in 2015. Pietermaritzburg: Shuter & Shooter; 1953.
2. South African Institute of Race Relations (SAIRR). Race relations in South Africa: Reasons for hope. Johannesburg: SAIRR; 2016. Available from: http://irr.org.za/reports-and-publications/occasional-reports/files/race-relations-in-sa-2013-reasons-for-hope-2013-29feb16.pdf
3. Theobald R. Rich and poor. A study of the economics of rising expectations. New York: C.N. Potter; 1960.
4. Netshitenzhe J. The black man's burden. Sunday Times. 2016 February 21; Opinion & Analysis. Available from: http://www.timeslive.co.za/sundaytimes/opinion/2016/02/21/The-black-mans-burden

A broader view of stewardship to achieve conservation and sustainability goals in South Africa

AUTHORS:
Jaco Barendse[1,2,3]
Dirk Roux[1,2]
Bianca Currie[1,4]
Natasha Wilson[5]
Christo Fabricius[1]

AFFILIATIONS:
[1]Sustainability Research Unit, Nelson Mandela Metropolitan University, George, South Africa
[2]South African National Parks – Scientific Services, George, South Africa
[3]Marine Stewardship Council – Science and Standards Team, London, United Kingdom
[4]School of Natural Resource Management, Nelson Mandela Metropolitan University, George, South Africa
[5]World Wide Fund for Nature – Land Programme, Cape Town, South Africa

CORRESPONDENCE TO:
Jaco Barendse

EMAIL:
jaco.barendse@gmail.com

POSTAL ADDRESS:
c/o Christo Fabricius, Sustainability Research Unit, Nelson Mandela Metropolitan University, Private Bag X6531, George 6530, South Africa

DATES:
Received: 10 Sep. 2015
Revised: 28 Nov. 2015
Accepted: 14 Dec. 2015

KEYWORDS:
earth stewardship; ecosystem services; natural resource management; biodiversity; eco-labels

HOW TO CITE:
Barendse J, Roux D, Currie B, Wilson N, Fabricius C. A broader view of stewardship to achieve conservation and sustainability goals in South Africa. S Afr J Sci. 2016;112(5/6), Art. #2015-0359, 15 pages. http://dx.doi.org/10.17159/sajs.2016/20150359

Stewardship is a popular term for the principles and actions aimed at improving sustainability and resilience of social-ecological systems at various scales and in different contexts. Participation in stewardship is voluntary, and is based on values of altruism and long-term benefits. At a global scale, 'earth stewardship' is viewed as a successor to earlier natural resource management systems. However, in South Africa, stewardship is narrowly applied to biodiversity conservation agreements on private land. Using a broader definition of stewardship, we identify all potentially related schemes that may contribute to sustainability and conservation outcomes. Stewardship schemes and actors are represented as a social network and placed in a simple typology based on objectives, mechanisms of action and operational scales. The predominant type was biodiversity stewardship programmes. The main actors were environmental non-governmental organisations participating in prominent bioregional landscape partnerships, together acting as important 'bridging organisations' within local stewardship networks. This bridging enables a high degree of collaboration between non-governmental and governmental bodies, especially provincial conservation agencies via mutual projects and conservation objectives. An unintended consequence may be that management accountability is relinquished or neglected by government because of inadequate implementation capacity. Other stewardship types, such as market-based and landscape initiatives, complemented primarily biodiversity ones, as part of national spatial conservation priorities. Not all schemes related to biodiversity, especially those involving common pool resources, markets and supply chains. Despite an apparent narrow biodiversity focus, there is evidence of diversification of scope to include more civic and community-level stewardship activities, in line with the earth stewardship metaphor.

Introduction

Over the past decade, 'stewardship' has become one of the dominant terms used to describe goals, principles and actions that aim to achieve sustainability in natural resource management, contribute to conservation priorities, and curb environmental degradation that threatens societal well-being.[1-3] Stewardship is not a new term[4], nor is it unique to a conservation perspective, e.g. in corporate management[5]. Even within an environmental context, its definition and interpretation varies greatly in its scale and application. At planetary scale, the terms 'ecosystem' and 'earth' stewardship are sometimes used interchangeably (e.g. by Chapin et al.[2,6]) to describe an overarching framework for dealing with social-ecological vulnerability and promoting general actions and systems that would enhance resilience in the light of global environmental change.[6-8]

On the other side of the spectrum, the stewardship tag is also applied in a much more focused manner, e.g. to describe market-linked incentives such as certification schemes for specific commodities such as the Marine Stewardship Council (MSC) for fisheries[9] and Forest Stewardship Council (FSC) for timber[10]. In some countries, stewardship is mostly associated with sustainability in agri-environmental systems (e.g. in the United Kingdom[11]) or with the adoption of better land and catchment management (e.g. 'Landcare' in Australia[12]), while elsewhere it may designate the management of formally protected or wilderness areas[13]. In South Africa today, 'stewardship' in a literal sense is understood to refer mainly to protecting biodiversity on privately owned land, under the banner of so-called Biodiversity Stewardship Programmes (BSPs).[14,15] Although such initiatives were identified over a decade earlier as a strategy to incentivise 'off-reserve' conservation[16], and well before adopting the term stewardship, it is now rarely used in any other context.

The broadscale, global interpretation of stewardship is based primarily on a developed country perspective, in which it is viewed by some as a possible 'successor' to earlier resource management regimes (namely steady-state and ecosystem management approaches – see Chapin et al.[2]). This view is embodied by a set of nine 'stewardship goals'[8] widely accepted as the guidelines for promoting earth stewardship[17]. These goals, grounded in the theory of social-ecological sustainability, include predominantly social aspects, e.g. equitable access to basic needs and opportunities, and sustaining ecosystem services.[3] The goals also include a number of other cross-cutting characteristics: (1) voluntary (as opposed to mandatory) participation[18]; (2) altruistic and moral-ethical connotations, sometimes associated with religion[19,20] that engender a sense of care[21] and shared responsibility, with consideration for the interests of human society, other species, and the natural world[22]; (3) an emphasis on inter-generational rather than short-term benefits[8]; (4) applicability across different spatial scales, i.e. from 'backyard to planet'[23]; and (5) the need for multiple partnerships, collaborations and linkages. Social networks and the stakeholder relationships that they represent are increasingly recognised as important features of natural resource management and conservation approaches.[24,25] Stewardship actions are often visualised as networks of actors with linkages within specific contexts, e.g. in urban ecosystems[26], at multiple scales[27], or across institutional and other divides[28].

Achieving these stewardship goals will require the implementation of practical mechanisms that could be viewed as the 'building blocks' of earth stewardship, preferably with metrics to indicate progress.[8] Such mechanisms could encompass decisions and actions at multiple scales (local, regional and global) based on the familiar principles of 'reduce, reuse and recycle'[29]. Other practical contributions may be the practice of 'civic' or 'urban' ecology: something as simple as planting a tree in the neighbourhood[28] or as complex as incorporating ecological principles into urban designs[30]. In practice, it is sometimes difficult to judge how stewardship differs from other environmental governance or natural resource management systems, or similar concepts like custodianship or trusteeship.[31] For example, co-management[32], like stewardship, is not necessarily driven purely by conservation objectives[33] but also by the need for benefit sharing[34]. In developing countries, stewardship principles are inherent to many community-based management systems[35] based on traditional and indigenous cultural values and beliefs (but for an opposing view see Fennell[36]).

One of the most compelling notions to emerge from proponents of the Earth Stewardship Initiative of the Ecological Society of America is the opportunity for less developed countries to 'leap frog' steps (e.g. steady-state resource management) on a typical Western resource management continuum directly to stewardship (see Figure 1 in Chapin et al.[2]), presumably avoiding the unsustainable practices of the past. How does this perspective relate to advancements in a developing or middle-income country context in which socio-economic disparities (e.g. developmental and income gaps) are far more pronounced and capacity to implement stewardship may be reduced? To assess this question we: (1) identified stewardship or stewardship-like mechanisms and their proponents or implementers in terrestrial, freshwater, and marine social-ecological systems in South Africa; (2) examined the relative influence and relationships between identified organisations or actors as a network; and (3) present the stewardship schemes as a simple typology, based on their objectives and operational scales. We discuss our findings relative to South African conservation and sustainability priorities, and in the wider context of Earth Stewardship Goals.[8] Our results not only provide a broader overview than the more traditional interpretation of stewardship in South Africa, but also allow us to reflect on whether stewardship has indeed emerged as a possible holistic or 'fast-track' option toward achieving conservation and sustainability goals.

Methods

Scoping and information retrieval

We anticipated considerable variation in the literal use of the term 'stewardship' and whether or how it is applied to initiatives that may be considered as stewardship activities. Therefore, during our review process, we did not take a purely systematic approach and adopted the following broad definition of stewardship:

> Any initiative, activity or voluntary involvement by an individual or organisation in the private, non-governmental or governmental sectors (including parastatal agencies), which seeks to contribute to, or promote, natural resource conservation or sustainability goals in social-ecological systems, both terrestrial and aquatic.

Further selection criteria were: (1) voluntary participation, i.e. not legislated (although a legal framework might apply); (2) non-commercial motivation, while acknowledging some operational costs (e.g. auditing costs for an eco-label); and (3) a natural resource or ecosystem management focus (as opposed to industrial processes).

We identified stewardship-related activities or initiatives ('schemes') and the most prominent organisations, individuals and other stakeholders involved in promoting and implementing these – collectively referred to as 'actors'. Importantly, compliance with our definition rather than explicit association with the term 'stewardship' was the main criterion for inclusion into our database, which was populated using both systematic and non-systematic search methods over a period of about 16 months (October 2012 – January 2014). We only considered schemes active within, but not necessarily restricted to, the Republic of South Africa. Actors could be based anywhere.

We did initial scoping through keyword searches on the Internet and in primary scientific indexing services using the terms 'stewardship' and 'Africa'. Next, we expanded our list of actors, schemes and associated terminology by a process of chain-referral (cf. snowball sampling[37]). We contacted or met with the most prominent actors, and asked about their own involvement in stewardship and for referrals to others, allowing us to identify more cryptic actors or schemes. Some referrals included suggestions to attend specific local and international meetings, including the Fynbos Forum (Cape St Francis, South Africa, July 2012), and the Symposium on Science & Stewardship to Protect & Sustain Wilderness Values at the 10th World Wilderness Congress (Salamanca, Spain, October 2013). Finally we conducted a more exhaustive round of searches based on two models of information retrieval: the 'berrypicking' model of Bates[38], and the 'Web moves' behavioural model of Choo[39]. The first is an 'evolving' search approach in which the cognitive response to results by the researcher may lead to on-the-fly modifications to the search process, e.g. by adding additional terms such as 'custodianship' or 'conservancies', or doing searches on a specific organisation, to broaden the sample. The latter model describes a progression of 'moves' whereby the researcher, starting on one website, follows links to other sites with relevant content ('chaining'), scans browsing results for most prominent returns, and differentiates between various results while bookmarking or capturing useful information. It includes an element of 'monitoring' whereby the sites are checked for updates and changes, and 'extraction' whereby a site is systematically searched for pertinent information (including type of scheme and its objectives, scale and mechanism of action).[40] These approaches enabled us to satisfy the objective of capturing the most readily available information on representative examples of stewardship within the region, including websites and primary scientific, academic and grey literature. As a minimum, for an initiative to be included, it had to comply with our definition and criteria, and we recorded additional information needed to identify the type of actor and scheme (detailed below).

Social network visualisation

We classified actors into five broad categories: (1) non-governmental organisations (NGOs), including registered charities or not-for-profit organisations; (2) funds – organisations that provide financing for schemes but do not normally undertake implementation; (3) governmental entities, including national and provincial ministries, departments or agencies; (4) private entities, including profit-driven companies and industry associations; and (5) partnerships – other groupings that do not fit the statutory entities described by (1)–(4), including collaborative networks, associations or programmes. All these actors represented the 'nodes' in our network. We then identified direct linkages ('edges') between pairs of nodes from stated collaborations on web pages and other documents, or implied through co-branding or logos on stewardship schemes, with each link assigned an arbitrary weight of one (i.e. multiple collaborations between the same actors would result in a weight greater than one). It is important to note that deriving linkages in this way did not allow us to assign directionality. We visualised relationships between actors as a social network using the software Gephi 0.8.2 beta[41]. We used the betweenness centrality – the number of times a node rests between two others which themselves are not linked – as a measure of relative prominence in stewardship (calculated with the algorithm of Brandes[42]). Actors with high betweenness centrality are considered important for long-term resource management planning, bringing together disconnected segments of a network.[43]

Typology

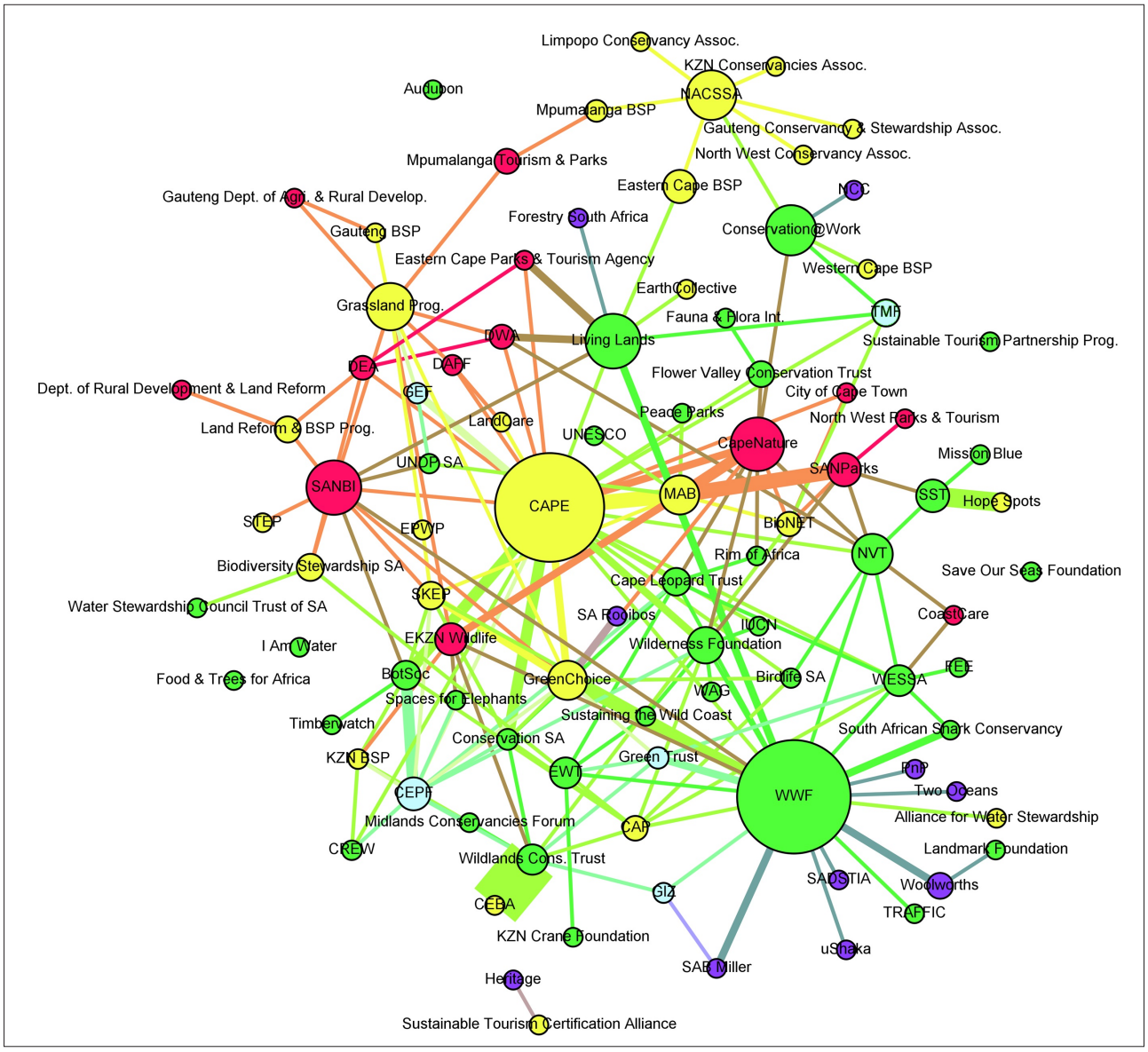
We identified broad types of stewardship schemes compliant with our definition, and based on information available about their objectives (e.g. focus on biodiversity or ecosystem services), mechanism of action (e.g. conservation on private land or market-based incentives), operational scale or footprint (global, national, sub-national or local). Despite some overlap between schemes and a lack of quantitative measures, we

could, during a workshop, agree on several major types, some of which had sub-types; we summarised these types and identified representative examples. Without consistent information available for each scheme, the types were subjectively placed on a conceptual plane based on our perceptions about: (1) the scale of their operational footprint (local to global); (2) their scope (extent of benefit or participation) ranging from individual to society; and (3) their 'tangibility' ranging from schemes that could be achieved through practical implementation to those needing more philosophical or ethical mind-shifts. Where available, we recorded any measurable indicators or documented achievements or challenges associated with specific types of schemes.

Results

Stewardship network and typology

Our final database included 38 NGOs, 14 governmental entities, 10 private entities, 5 funds and 27 partnerships (between any of the recorded actors). These 94 nodes and 180 edges between them formed the basis for the social network visualisation (Figure 1; see also Appendix 1 and the supplementary material). Among these there were seven global NGOs, and three global funds; most other NGOs were national ($n=18$) or sub-national ($n=10$), noting that international NGOs with South African branches were considered 'national', e.g. the World Wide Fund for Nature (WWF-SA) and Conservation South Africa. Partnerships were mostly national ($n=7$) or sub-national ($n=16$) with three examples of global and one local partnership. The network appears well-connected with many linkages but relatively few prominent actors, and some disconnected nodes. Most prominent with the highest betweenness centrality values were NGOs and partnerships that relate to biodiversity and landscape conservation initiatives, notably the Cape Action Plan for People and the Environment (CAPE) – a systematic conservation plan for the Cape Floristic Region, initiated in 1998 with funding from the Global Environmental Facility's Critical Ecosystem Partnership Fund (CEPF) and coordinated by WWF-SA[44]. National (e.g. South African National Biodiversity Institute) and sub-national governmental agencies (e.g. CapeNature) also feature as important bridging nodes.



Note: green = non-governmental oganisations; yellow = partnerships; red = government entities; blue = funds; lilac = private entities

Figure 1: Main actors involved in conservation and sustainability stewardship schemes in South Africa as a social network with 94 nodes and 180 edges (visualised in Gephi 0.8.2 beta). The thickness of links is relative to the number of collaborations and associations between nodes. The nodes are sized on an arbitrary scale (5–50) relative to their betweenness centrality as an indicator of relative prominence or involvement within the network. (See Appendix for full labels and types).

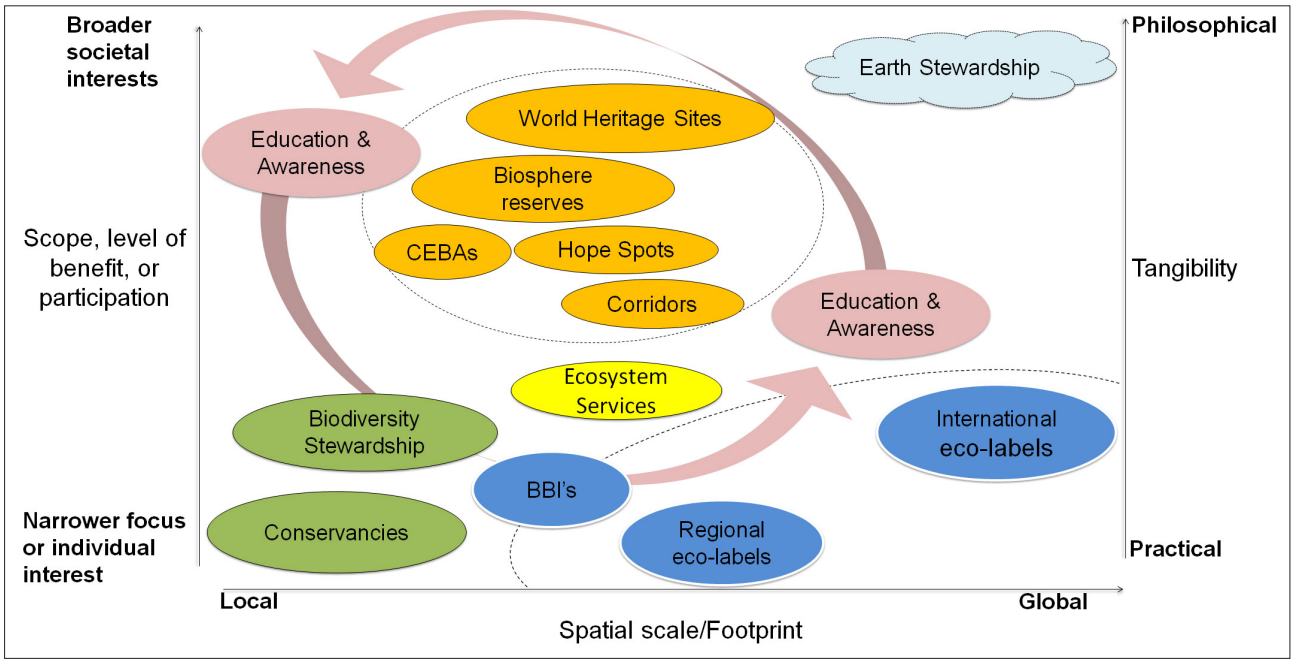
We distinguished six main types of stewardship schemes (defined in Table 1), some with sub-types or some degree of overlap. In terms of scale, 5 schemes were operational at a global scale, 2 at the African continent level, 13 at national, 38 at sub-national (i.e. coverage limited to one or more of South Africa's nine provinces), and 19 at local level (i.e. limited to smaller areas such as a city or catchment). Based on the above we positioned the main stewardship types or sub-types on a conceptual plane (Figure 2) and describe and present examples of each type in more detail below (also see Table 1).

Conservancies

Conservancies represent the oldest form of voluntary conservation on private land in South Africa – the first conservancy was established by a group of farmers in 1978 in the Balgowan District of KwaZulu-Natal (KZN), through encouragement by the former Natal Parks Board (now Ezemvelo KZN Wildlife). Although conservancies are required to be registered with the regional conservation authority, there is no binding agreement between these parties. Conservancies are viewed as the entry-level to more formal stewardship agreements, but are not

Table 1: Definitions of main types of conservation and sustainability stewardship schemes identified in South Africa

Type of scheme	Definition
Conservancies	Registered voluntary associations, established between like-minded landowners, residents, communities and other users, in a specified area with the shared aim of co-operative management of its natural resources in an environmentally sustainable manner, without necessarily changing the land use on the properties.
Biodiversity stewardship programmes	Mechanism to incentivise formal conservation on private lands with high biodiversity conservation value. Different participation levels are available but the ultimate aim is to proclaim such areas as formally protected by national laws.
Land- and seascape initiatives	Initiatives that focus at a land- or seascape level, often determined by unique or specific biophysical or other characteristics or features (e.g. geological or heritage), to promote resilience of protected areas through inclusion of buffer areas, or enhanced connectivity between formally protected areas through multiple mechanisms.
Market-linked schemes	Initiatives that focus on the production, management, or value chain of specific commodities or services and aim to promote sustainability by incentivising consumers to support such schemes, thus harnessing market forces to reward such producers.
Ecosystem services	Initiatives that broadly address issues around maintenance or restoration of ecological infrastructure or ecosystem services through practical or policy interventions.
Education and awareness initiatives	Initiatives aimed at education or raising awareness in specific or multiple sectors of society (e.g. the youth, or consumers and retailers) about particular or broader issues relating to sustainability or conservation, thus encouraging the voluntary adoption of behaviours and attitudes that contribute to such causes.



BBIs, Biodiversity and Business Initiatives; CEBAs, Community Ecosystems Based Adaptation sites of the Wildlands Conservation Trust

Note: Orange shapes (bound by dashed ellipse) are land- and seascape schemes and blue shapes (bound by dashed curve) are market-based schemes; curved arrows from and to education and awareness schemes (pink) indicate cross-cutting functions.

Figure 2: Conceptual representation of selected stewardship types and sub-types identified in South Africa based on their spatial extent and perceived focal scope of benefit, participation and tangibility.

generally recognised as components of BSPs (see below). Participation is often based on shared aims and a sense of identity (e.g. expressed through logos displayed in media forums and individual farm signage), which may enable members to access funds to implement conservation action, if they so choose. Each conservancy is made up of individual landowners, while individual conservancies are organised into provincial associations. Since 2003, provincial associations have unified under the National Association of Conservancies (and Stewardship) South Africa (NACSSA).[45,46] Conservancies are found in all nine South African provinces, with the number of those registered indicated: KZN (126), Free State (152), Western Cape (70), Northern Cape (8), Eastern Cape (34), Gauteng (50), Mpumalanga (39) and Limpopo (17) (Wesson J 2013, written communication, September 4). These conservancies include rural, urban, township and industrial sites (e.g. landfills) covering a total estimated 3 000 000 ha.[45,47] However, not all are registered and specific information such as conservancy names, GPS coordinates and areal extent are not readily available (Young A 2013, written communication, September 3).

Biodiversity stewardship

The conservation of biodiversity using so-called 'stewardship agreements' was conceived at national level by the South African National Biodiversity Institute (SANBI) but is implemented sub-nationally by provincial conservation agencies and NGOs. Biodiversity stewardship is most prominent in the Western Cape where it was piloted as the 'Conservation Stewardship Programme' in late 2002 by CapeNature and the Botanical Society[48] as part of the CAPE strategy[49-51]. The approach was later adopted in other provinces such as KZN.[52] The underlying objective of these BSPs is to improve protection of critical biodiversity and threatened ecosystems occurring on private and communal land as determined by national conservation plans and spatial assessments.[53,54] This protection is to be achieved by encouraging formal conservation agreements between the conservation agency and landowners through financial (e.g. tax relief[55,56]) and in-kind (extension services – habitat and land management advisory) incentives. The programme recognises various levels of participation, namely biodiversity agreements, protected environments and contract nature reserves (as defined in the *Protected Areas Act*[57]) that differ in degree of legal protection status, land-use restriction (on title deeds) and minimum duration of management tenure: 10 years for biodiversity agreements, 30 years for protected environments and 99 years for contract nature reserves[48,51]. By 2010 in the Western Cape there were 33 contract nature reserves (45 261 ha), 17 biodiversity agreements (11 336 ha) and 21 voluntary sites (20 446 ha).[58] By 2013 in KZN there were 9 contract nature reserves (35 953 ha), 1 protected environment (238 ha) and 3 biodiversity agreements (4274 ha), with a further 34 630 ha in the final stage of proclamation in the first two categories, and nearly 175 000 ha under negotiation (Martindale G 2013, written communication, August 29). In other provinces (e.g. Gauteng and Eastern Cape), BSPs are more recent (post-2009) and outcomes are not readily available. As the programme has developed, more local authorities and NGOs have expressed interest in adopting this model.

Land- and seascape initiatives

These schemes share a broad focus at land- (or sea-) scape level, usually determined by unique or exceptional biodiversity, geographical features or other characteristics, sometimes in combination. They aim to improve the protection of an area by raising awareness about the unique features or conservation profiles through special listings or other means of recognition. They vary in spatial scale from sub-national to regional, although some are international initiatives e.g. the United Nations Educational, Scientific, and Cultural Organisation's (UNESCO) Man and the Biosphere (MAB) Programme[59] and World Heritage Sites[60]. Landscape initiatives often correspond closely with national (e.g. CAPE) or international (e.g. IUCN 'Key Biodiversity Areas') bioregional programmes.

Corridors

One of the most diverse sub-types of landscape initiatives involve the concept of corridors. These include corridors that link formally protected areas primarily for conservation of biodiversity and processes (e.g. Gouritz River Initiative[61] and Eden to Addo[62]) to those that would enhance wildlife movement (e.g. for African elephants (*Loxodonta africana*) – the Lubombo Spine Corridor). Increasingly, corridors have evolved beyond biodiversity conservation tools to also include building resilience and adapting to climate change or provision of ecosystem services (e.g. freshwater stewardship or river catchment corridors). For example, the Climate Action Partnership between Conservation South Africa and other NGOs has identified 45 such corridors in KZN and 14 in the Eastern Cape (in collaboration with the Wildlands Conservation Trust). Finally, the corridor concept is also used in the context of eco-tourism routes to promote human–nature experiences (e.g. Segarona Heritage Park Hike between Pilanesberg and Madikwe Game Reserves[63]), to raise awareness about ecosystems and their conservation (e.g. Rim of Africa – a >600 km hiking trail in Western Cape mountains[64,65]) or even to conceptually link miscellaneous important 'heritage sites' at a continent scale, so furthering the notion of earth stewardship (see 'Africa Alive Corridors' concept[66]).

Sites with special features

The sites of many of the corridors (see above) are closely tied to existing biodiversity stewardship sites, protected areas and other landscape-level initiatives such as mega-reserves (e.g. Baviaanskloof). These sites also include MAB reserves of which there are eight, which together cover over 7 million ha: Kogelberg, 103 629 ha; Cape West Coast, 378 240 ha; Kruger to Canyons, 2 474 700 ha; Waterberg, 414 571 ha; Cape Winelands, 322 030 ha; Vhembe, 30 701 ha; Gouritz Cluster, 3 187 892 ha; and Magaliesberg, 357 870 ha in South Africa[59], the oldest being Kogelberg (designated in 1998), and the most recent the Gouritz Cluster and Magaliesberg MAB[65], both proclaimed in 2015. South Africa has eight listed World Heritage Sites under the World Heritage Convention of which four are for cultural features, three for natural features (Cape Floral Region, iSimangaliso Wetland Park and Vredefort Dome), and one mixed natural-cultural (Maloti-Drakensberg)[60]. Notably absent is marine and coastal coverage. However, the concept of International Hope Spots – part of Sylvia Earl's Mission Blue[67] – is being championed by the local NGO, Sustainable Seas Trust[68], with five proposed sites (Algoa Bay, Aliwal Shoal, Cape Whale Coast, Knysna and Plettenberg Bay).

Market-based schemes

These schemes focused on environmental sustainability objectives (we did not consider primarily social ones such as Fair Trade) at the resource production or ecosystem level of a value chain, by trying to influence consumers to reward more sustainable supplies of a product (e.g. seafood) or service (e.g. tourism) through their choices.[69] There were two sub-types: eco-labels and business and biodiversity initiatives.

Eco-labels

Eco-labels rely on a certification standard for a specified commodity or service; its adoption entitles the producer/service provider to use the eco-label mark as a marketing tool. These included leading international third party eco-labels: one MSC certified fishery (South African demersal 'Cape' hake (*Merluccius* spp.) trawl fishery of ca 120 000 t per year, first certified in 2004); 20 forestry management areas certified by the FSC covering >1.48 million ha; and the Blue Flag tourism eco-label for 36 beaches, 4 marinas, and 3 whale-watching boats implemented by the Wildlife and Environment Society of South Africa (WESSA)[70]. There were several national eco-labels addressing specific issues, e.g. badger-friendly honey, predator-friendly meat[71], sustainable golf courses (e.g. one in Audubon Cooperative Sanctuary Program) or tourism accommodation (e.g. Green Leaf). At a continental level, the African Eco-Labelling Mechanism[72] has developed draft standards for the agriculture, fishery and forestry sectors. It is important to note that national or regional eco-labels do not always make use of third-party verification or traceability mechanisms for certified products.

Business and biodiversity initiatives

Business and biodiversity initiatives (BBIs) focus on production systems for specific products by making a 'business case' for biodiversity conser-

vation and sustainable harvesting during production of specific products. It advocates voluntary adoption of better on-farm conservation practices and the setting aside of land for conservation (through BSPs), while the participant can use membership (including on-product information) as a potential marketing advantage. The first BBI was the Biodiversity & Wine Initiative, established in 2004 in the Western Cape winelands through a multi-stakeholder partnership.[73,74] The model is expanding to include other production sectors within the Cape Floristic Region and other bioregions, for example, rooibos (*Aspalathus linearis*) (note also three producers certified by the Rainforest Alliance[75]); 'Biodiversity &…Citrus, Red Meat, Rooibos, Ostrich, and Potato Initiatives; Grasslands Programme Red Meat; and Cape parrot (*Poicephalus robustus*) friendly pecan nuts[76]. In 2010, 280 BBI members had a total footprint of 250 153 ha of natural habitat, mainly in the Western Cape.[77]

Ecosystem services

These schemes focus on restoration of specific ecosystem services. They include primarily government-driven initiatives, but with more or less voluntary adoption (or in lieu of financial payments for ecosystem services[78]) by private landowners, such as the state-funded Expanded Public Works Programme of which the best known is Working for Water (WfW)[79]. Established in 1995 to provide low-skill employment opportunities for poor communities while restoring water run-off by clearing alien invasive plants from catchments[80], WfW treated over 1.3 million condensed hectares[81] between 2002 and 2008, mainly on public land (e.g. in National Parks). The model has been expanded to include Working for Wetlands (e.g. restoration through constructing gabions), Working for the Coast (e.g. coastal clean-ups and resource user monitoring) and Working on Fire (e.g. combating wild fires, or reducing fuel loads through invasive plant removal) programmes.[82]

Other schemes more specific to agricultural production include state initiatives such as the National LandCare Programme[83] which addresses, inter alia, soil management and erosion control on farms. Others are driven by NGO and private/corporate partnerships, like the Sustainable Sugarcane Farm Management System (known as SUSFARMS).[84] Some water stewardship initiatives focus strongly on the link between the supply chain and catchment management, e.g. WWF Water Futures Partnership with SABMiller on hops production[85], or the standards set by the Alliance for Water Stewardship that have now been adopted by South African producers of export stone fruits[86,87].

Education and awareness

These schemes either focus on a specific cause, e.g. sustainable seafood, or incorporate information about multiple causes into a 'basket' of sustainable options aimed at the general public. They also promote more sustainable living among specific sectors, e.g. scholars, through actions such as saving water and energy or recycling. For example, the Southern African Sustainable Seafood Initiative (SASSI) which encourages seafood consumers to consult a 'traffic-light' species list of sustainable seafood choices when buying fish[88,89]; through this market pressure its influence may extend into regulatory or policy areas[90]. Eco-Schools[91] is a sustainable schools programme from the international NGO Foundation for Environmental Education, but implemented in South Africa by WESSA with 1200 registered schools. Often, because education and awareness are ancillary functions to the main objectives of NGOs, such schemes were difficult to isolate, and tend to have a cross-cutting function (represented by the arrow in Figure 2) by linking multiple schemes, e.g. GreenChoice which markets a 'basket' of sustainable options from different schemes (including eco-labels and BBIs) to the general public[92].

Discussion

Our broad overview of stewardship schemes in South Africa is, to our knowledge, the first such at a countrywide scale. Our findings represent a much wider perspective on stewardship than has ever been used in any developing country. We present our findings under broad themes that aim to capture the key features of stewardship in South Africa, while maintaining a global context.

Biodiversity focus and the role of NGOs and partnerships

The strong focus on biodiversity conservation on private land over the past decade is perhaps not surprising, given that much of South Africa's globally recognised biodiversity and threatened environments, especially in the Cape Floristic Region, is located outside of formally protected areas.[54] This focus not only explains the prominence of CAPE (Figure 1), but also why many aspects of BSPs, especially within the CAPE planning domain, have been examined more in-depth: policy and governance frameworks[93,94]; perceptions and motivations for participation[95], e.g. tax incentives[55,56]; the relationship between biodiversity stewardship and social learning[96]; and evaluating the contribution of BSPs to national conservation goals[49]. The CAPE partnership, together with major 'global' NGOs (e.g. WWF-SA), form dominant elements of the stewardship network, in effect combining as a 'bridging organisation'[97]. Such organisations, on the one hand, leverage external resources or 'bridging ties' like international funding (e.g. from the CEPF), while on the other hand, connect and enable diverse local actors to utilise new 'possibilities for action'[98]. Although our data set did not allow an in-depth analysis or understanding of the stewardship network, it suggests that more social network analysis could be valuable in gaining a better understanding of stewardship at specific spatial scales, within specific groups of actors, and the links between international and local conservation priorities and actions.[99] The relative prominence of NGOs with national (e.g. Wildlands Conservation Trust) or sub-national footprints (e.g. Nature's Valley Trust) that act as implementing agencies, or that collaborate with state entities such as provincial conservation agencies on more diverse stewardship approaches, suggests examples of cross-scale and scale-bridging interactions[27].

Following CAPE's success, similar partnerships were started in other bioregions, notably the Succulent Karoo Ecosystem Plan (SKEP)[100] and the Subtropical Thicket Ecosystem Plan (STEP)[101]. After being piloted in the Western Cape and KZN Provinces, the BSP model has been expanded to other provinces, correlating with other globally recognised biodiversity priority areas, e.g. the Maputaland-Pondoland-Albany centre of endemism in the Eastern Cape, and grasslands in Mpumalanga.[102] Furthermore, the BSP sites have determined target areas for implementing market-based and landscape stewardship mechanisms like biosphere reserves (e.g. Kogelberg and West Coast), corridors (e.g. Greater Cederberg Biodiversity Corridor) and BBIs in the Cape Floristic Region (e.g. wine and flowers).

Non-biodiversity goals: Common pool, markets and ecosystem services

In contrast to the above, some stewardship schemes are not necessarily tied to biodiversity and bioregional focus. For example, the establishment of conservancies pre-dates spatial prioritisations. Conservancies are found in all provinces and motivations for their establishment are more diverse, sometimes tending toward self-interest (see below). Another exception to a singular biodiversity focus is stewardship schemes dealing with common pool resources, value chains and markets, or ecosystem services. Marine and coastal ecosystems present classical examples of common pool natural resources[103] held in 'public trust' by the state on behalf of its citizens[104]. Stewardship activities by citizens or interest groups in the marine environment thus present something of a conundrum: they are trying to be co-stewards of something already under government custodianship on their behalf (but see the concept of marine citizenship[105]). In the South African context, stewardship schemes in the marine environment are predominantly market-based or educational (e.g. MSC, Blue Flag and SASSI) with seascape-level schemes such as International Ocean Hope Spots only a recent development – not unexpected when bioregional planning and prioritisation has lagged in the marine environment. Surprisingly, co-management, which is generally considered conducive to sustainable harvesting and resource stewardship[106,107], has struggled to emerge within South Africa's current fisheries management regime[108]. Increasingly, the term 'stewardship' is adopted to describe collaborative governance approaches to manage global commons such as the deep ocean.[109] This trend may reflect growing recognition that, up to now, states have failed to adequately

manage oceans as a global commons by not viewing ocean ecosystems at a planetary scale. This failure may be a result of single species or regional foci, or by ignoring ethics and the interconnectedness of ecosystems and stakeholders.[110]

Stewardship based on value chains, markets or commodities can sometimes be at odds with biodiversity conservation; for example, the FSC, which in South Africa primarily certifies monoculture plantations of exotic (often invasive) tree species like pines and gums, located in biodiverse fynbos and grassland habitats. Ironically, the only exception to this contradiction is the FSC certificate held by South African National Parks for harvesting indigenous hardwoods in the Garden Route National Park. The fact that 'plantations are not forests' is strongly advocated by some lobby groups.[111] Similarly, many conservationists dispute that any bottom trawl fishery should be certified as sustainable.[112] Focus on specific ecosystem services or concepts like biodiversity offsets (or other mitigation measures) within the stewardship discourse is likely to remain uncomfortable, if not controversial, when there is evidence that non-biodiversity objectives are not always compatible with biodiversity ones.[113]

Motivations and mechanisms

Published sources suggest that intrinsic motivations to participate in stewardship include altruism and acting in societal interest.[114] Although environmental consciousness (cf. biophilia[115]) is an assumed prerequisite for private landowners to create conservancies in South Africa, a range of reasons are reported, ranging from nature conservation (primary) and security for domestic and wild animals to securing recreational or tourism opportunities including hunting, or sometimes to oppose development.[116] Some of these reasons may be equally applicable when entering into more formal BSP arrangements, but often it is up to the proponent (e.g. provincial conservation agency and NGOs) to 'sell' the concept to the potential steward. Incentives may include financial ones[55], but also 'extension services': specialist input and management assistance relating to land and biodiversity. The type of landowner (commercial versus lifestyle farmer), land size and opportunity costs can all impact on willingness to participate in conservation.[117] Recent work using the Biodiversity & Wine Initiative as example, suggests that both intrinsic and extrinsic factors are important for farmers to join this BBI, notably their own value systems.[118] The importance of issue 'champions' as a key driver for participation was also emphasised.

International eco-labels are seen to inadvertently encourage global stewardship by empowering mainly northern hemisphere consumers to take personal responsibility for the production of a commodity elsewhere, especially in the developing world.[119] Initially, adoption of both the FSC and MSC in South Africa was motivated by the demands of the export market, rather than local consumer choice.[120,121] Some argue that payment for ecosystem services is inherently easier to leverage from a business perspective than payment for biodiversity[122], hence diversification of stewardship mechanisms to include ecosystem services[78] or value chains. Market-based interventions, together with consumer awareness schemes (e.g. GreenChoice), contribute to making a 'business case' for biodiversity conservation.[76] The notion of a business case often finds resonance and expression in corporate stewardship 'sustainability journeys' of retailers[123], although there are possible weaknesses in using 'journey' as a metaphor for measuring progress in sustainability[124].

Successes, benefits and shortcomings

Although we could not directly measure the efficacy and drawbacks of stewardship schemes from our data, some published results offer indicators of their success. These indicators include participation levels in stewardship schemes, hectares of land in BSPs, or more tangible conservation outcomes, for example, significant reduction in seabird mortality in the hake trawl as a result of MSC certification[125]. Conversely, the withdrawal of Blue Flag status at Margate in KZN because of poor water quality has been equated to a substantial revenue loss.[126] For conservancies, the growing number of voluntary participants, their presence in all provinces, and a national alliance that includes 'community level stewardship' in its vision are all positive trends. There are several recognised benefits to game ranching of consolidated estates, including more profitable (from an eco-tourism perspective) and viable wildlife populations, especially for larger species with bigger ranges.[127]

Another apparent benefit of BSPs is achieving national conservation goals at much lower cost to the state (than land acquisition). While this may be so for provincial agencies, conservation on private land is sometimes viewed as an 'unfunded mandate' by national agencies (e.g. South African National Parks) – in other words, the budgetary and human resource requirements are not commensurate with the area to be managed. Thus, while participating in BSPs is considered 'voluntary', the underlying biodiversity objectives may confound the voluntary nature of participation, as land with 'low' conservation value is not wanted, given the financial and human capacity requirements for extension services and other management costs. There must remain serious concerns regarding the statutory security of conservancies and other forms of biodiversity stewardship. The dependence of conservancies on the personal values of the participant casts doubt on whether they should be included under national conservation targets.[116] For instance, an evaluation in 2010 of 280 BBI members indicated coverage of 250 153 ha of natural habitat; however, since 2006 there has been a loss of 2827 ha to habitat transformation and 892 ha to degradation,[128] bringing some doubt over the sensibility of 'banking' on a volunteer mechanism to achieve national mandates.

Inasmuch as governmental agencies tasked with biodiversity conservation have embraced these new governance arrangements[93] to achieve conservation targets, a single-minded focus on one mechanism may have additional drawbacks. For example, it may inadvertently cause neglect on other land with equally important biodiversity, such as 'escapee' pines invading state-controlled watersheds.[129] The position of the state may even appear 'schizophrenic', especially when the state defaults on its fiduciary duty as public biodiversity custodian.[130] At times, a government's action or inaction may pose a direct threat to biodiversity inside and outside protected areas, e.g. by assigning prospecting rights for shale gas across entire bioregions, or for benthic phosphate mining[131]; by permitting coal mining adjacent to nature reserves (e.g. at Hluhluwe-Imfolozi[132]); or by on-going political support to permit angling in Africa's oldest no-take Marine Protected Area, Tsitsikamma[133,134] (recently gazetted by the Department of Environmental Affairs[135]). In such instances, civic or special interest groups or industries may adopt stewardship as an anti-measure to such threats[136], evoking Section 24 of the South African Constitution – 'a right to a healthy environment and sustainable development and use of natural resources'. Examples of this type of adoption include the South African Deep Sea Trawling Industry Association (SADSTIA) 'forcing' management actions[137] through the possible forfeiture of the MSC certification of the South African hake trawl fishery (as a result of management authority's impasse on collecting annual stock assessment data) in an arena in which political agendas sometimes appear to trump conservation ones[138]; and, most recently, SADSTIA aligning with NGOs to oppose bulk marine sediment mining authorised by the Minister of Mineral Affairs[139]. The breakdown in trust (of citizens in the state) resulting from a government's neglect of its duties as steward can be difficult to restore[104], and can lead to instances of 'extreme' stewardship, e.g. the formation of vigilante groups[140] against the abalone poaching crisis in South Africa[141]. Conversely, disparate views and apparent disasters can result in cooperative learning between diverse stakeholders, thereby improving ecosystem stewardship.[142] In many African countries where the conservation priorities are clear, yet resources and capacities are genuinely lacking, management responsibility may be readily delegated to public–private partnerships driven by international NGOs.[143]

A contrasting scenario is presented by the primarily government driven schemes for ecosystem services restoration. For example, the cost-effectiveness of WfW has been assessed at local and national scales[81,144] and, although many regard it as overwhelmingly positive, there is a sense that its overall performance needs to be improved, inter alia, by better prioritisation of alien invasive species, more targeted actions, and less emphasis on social benefits as a measure of success. Even less successful has been the ability to stimulate stewardship actions among private landowners, i.e. by maintaining cleared areas to prevent re-invasion after initial WfW clearing.

This inability may be because of uncertainty around the extent of state versus private responsibilities, and differential attitudes towards incentives or disincentives to participate.[145] Nonetheless, negligence on private land erodes overall gains made by public programmes[144] emphasising the importance of public–private relations.

Relationship to other management approaches

The stewardship approaches found in South Africa are similar to other conservation and natural resource management strategies elsewhere. For example, conservation easements in the United States of America[146,147] are very similar to biodiversity stewardship agreements in South Africa: the term 'cooperative environmental governance systems' has been used to describe such arrangements[148]. Some believe that stewardship differs from other management systems by its recognition of 'embedded values' and preoccupation with conservation and sustainability.[31] However, these traits are common to community-based natural resource management (CBNRM) – another decentralised approach to achieving environmental, social and economic goals by balancing the exploitation and conservation of valued ecosystem components.[149] In CBNRM, voluntary local civic institutional arrangements are formed to manage natural resources, suggesting that it may be viewed as a form of stewardship[150] or as a mechanism for achieving stewardship of watersheds[151], and wildlife and forests[152,153], and for sustainable rural agriculture[154]. Further, developing social capital, collaborative partnerships and networks have been highlighted as key principles of CBNRM[155], which is echoed in the stewardship metaphor[156]. It is noteworthy that CBNRM predominantly focuses on common pool resources and is often underpinned by cultural and traditional values, e.g. the conservation of sacred landscapes[157], totemic species or culturally important natural features[158].

In practice, different conservation mechanisms are rarely applied in isolation and stewardship schemes may be seen as ancillary to tactics such as land acquisition (e.g. by the Nature Conservancy in the United States of America) that all form part of a modern strategic conservation approach[159] in areas of high biodiversity, as in South Africa[160]. For example, WWF-SA, in addition to facilitating private land stewardship, actively pursues the expansion of existing or establishment of new protected areas through land acquisitions (ca 400 000 ha or 5% of the national terrestrial protected area estate), predominantly financed through land trusts.

Earth stewardship in South Africa?

In retrospect, it is apparent that many of the described stewardship schemes could fit under the banner of 'earth stewardship': operational at multiple scales with diverse stakeholders, and recognising interconnectedness, ethics and indigenous knowledge (for a snapshot of examples see Sayre et al.[161] and papers in that volume). However, the dominance of contractual biodiversity conservation initiatives in the South African stewardship narrative has masked the emergence of a more holistic stewardship strategy as advancement on contemporary resource management approaches (as contemplated by Chapin et al.[2]).

The narrow association of stewardship with systematic biodiversity conservation plans and associated spatial priorities is perhaps not surprising, given that most post-colonial countries still develop within the 'constraints' of governance or management systems inherited from the North – much as 'global assemblages' can impact on poorer nations[162] – in effect inhibiting the emergence of earth stewardship. Furthermore, weaker governance in some developing countries has driven 'decentralised' mechanisms of environmental decision-making and policy implementation at community level.[163]

Our research approach was unlikely to provide adequate resolution to detect the emergence of local-level governance approaches, especially for tacit (e.g. CBNRM and co-management) and local grassroots level (e.g. urban greening initiatives) stewardship forms that were underrepresented in the information sources we consulted, and so more difficult to detect. In fact, the recent broadening of the stewardship narrative in South Africa – firstly by NGOs adopting the term 'earth stewardship'[164] and, secondly, through the fairly rapid diversification away from a strictly biodiversity focus to more holistic models – suggests that a shift is taking place. In countries with a legacy of post-colonial land ownership, this shift may reflect recognition of the need to acknowledge and incorporate local socio-political issues into any stewardship approaches. This new approach is typified by the Community Ecosystems Based Adaptation ('CEBA') sites of the Wildlands Conservation Trust which uses a 'basket of products' approach (including 'Green-preneurship' and restoration), with implementation strategies[165] that mirror many of the earth stewardship principles, while strongly emphasising involvement of local communities. There is thus a clear need to evolve Western-based concepts of stewardship and conservation to include indigenous values[157] or more collaborative management approaches[166].

Finally, while we believe that 'earth stewardship' may be an appropriate metaphoric term to describe the link between primarily conservation-driven schemes with more social and economic ones, it is unlikely to be an implementable 'catch-all' solution in countries with weak or ineffective governance systems (as suggested by Kinzig et al.[167]). Overuse of the term in a global or philosophical sense may eventually dilute its value and practicality in an implementation context. We contend that, to achieve sustainability or conservation outcomes, reliance on a single mechanism – whether voluntary (i.e. 'stewardship') and thus dependent on the social norms, ethics, values or behaviours of individuals, or as determined by government policies (i.e. mandatory or legislated) – is risky. This contention is important, as formalising any voluntary participation into binding agreements may result in issues similar to those faced by extant formal management systems, e.g. lack of capacity or 'non-compliance' by participants, and corruption. In a developing world context the need for complementarity between different management approaches[168] is key to achieving the desired conservation and sustainability outcomes.

Acknowledgements

We are extremely grateful to the following people for valuable discussions, inputs and other contributions: O. Huyser, D. Nel, I. Kotze, A. Venter, G. Martindale (unpublished data on KwaZulu-Natal stewardship sites), J. Wesson (unpublished data on national conservancies), B. Schuman, M. Hope, A. Jones, A. Young (KZN conservancies), J. Lindsay (national conservancy information), M. Alant, N. Diederichs-Mander, K. McCann, K. Coetzee, A. du Plessis (Western Cape stewardship sites). All work was conducted in our academic or professional capacities, and the views expressed do not necessarily reflect those of our various organisations. Two anonymous reviewers provided invaluable comments and suggested improvements. J.B. was supported by a postdoctoral research grant from NMMU and SANParks.

Authors' contributions

The concept was developed by J.B., C.F. and D.R. J.B. collected and analysed the data with inputs from D.R., C.F., B.C. and N.W. J.B. wrote the manuscript with contributions from D.R., C.F., B.C. and N.W.

References

1. Chapin III FS, Kofinas GP, Folke C, Chapin MC. Principles of ecosystem stewardship: Resilience-based natural resource management in a changing world. New York: Springer; 2009. http://dx.doi.org/10.1007/978-0-387-73033-2_15
2. Chapin III FS, Carpenter SR, Kofinas GP, Folke C, Abel N, Clark WC, et al. Ecosystem stewardship: Sustainability strategies for a rapidly changing planet. Trends Ecol Evol. 2010;25(4):241–249. http://dx.doi.org/10.1016/j.tree.2009.10.008
3. Raymond CM, Singh GG, Benessaiah K, Bernhardt JR, Levine J, Nelson H, et al. Ecosystem services and beyond: Using multiple metaphors to understand human–environment relationships. BioScience. 2013;63(7):536–546. http://dx.doi.org/10.1525/bio.2013.63.7.7
4. Odum E. Earth stewardship. Pan Ecology. 1991;6(5):1–5.
5. Hernandez M. Toward an understanding of the psychology of stewardship. Acad Manage Rev. 2012;37(2):172–193. http://dx.doi.org/10.5465/amr.2010.0363

6. Chapin III FS, Pickett SA, Power M, Jackson R, Carter D, Duke C. Earth stewardship: A strategy for social–ecological transformation to reverse planetary degradation. J Environ Stud Sci. 2011;1(1):44–53. http://dx.doi.org/10.1007/s13412-011-0010-7

7. Steffen W, Persson Å, Deutsch L, Zalasiewicz J, Williams M, Richardson K, et al. The Anthropocene: From global change to planetary stewardship. Ambio. 2011;40(7):739–761. http://dx.doi.org/10.1007/s13280-011-0185-x

8. Chapin III FS, Power ME, Pickett STA, Freitag A, Reynolds JA, Jackson RB, et al. Earth stewardship: Science for action to sustain the human–earth system. Ecosphere. 2011;2(8), Art. #89, 20 pages. http://dx.doi.org/10.1890/ES11-00166.1

9. Gulbrandsen LH. The emergence and effectiveness of the Marine Stewardship Council. Mar Policy. 2009;33(4):654–660. http://dx.doi.org/10.1016/j.marpol.2009.01.002

10. Pattberg PH. The Forest Stewardship Council: Risk and potential of private forest governance. J Env Dev. 2005;14(3):356–374. http://dx.doi.org/10.1177/1070496505280062

11. Dobbs TL, Pretty JN. Agri-environmental stewardship schemes and "multifunctionality". Appl Econ Persp Policy. 2004;26(2):220–237. http://dx.doi.org/10.1111/j.1467-9353.2004.00172.x

12. Curtis A, De Lacy T. Landcare, stewardship and sustainable agriculture in Australia. Environ Value. 1998;7(1):59–78. http://dx.doi.org/10.3197/096327198129341474

13. Hobbs RJ, Cole DN, Yung L, Zavaleta ES, Aplet GH, Chapin III FS, et al. Guiding concepts for park and wilderness stewardship in an era of global environmental change. Front Ecol Environ. 2009;8(9):483–490. http://dx.doi.org/10.1890/090089

14. Pence GQK, Botha MA, Turpie JK. Evaluating combinations of on-and off-reserve conservation strategies for the Agulhas Plain, South Africa: A financial perspective. Biol Conserv. 2003;112(1–2):253–273. http://dx.doi.org/10.1016/S0006-3207(02)00413-5

15. Gallo JA, Pasquini L, Reyers B, Cowling RM. The role of private conservation areas in biodiversity representation and target achievement within the Little Karoo region, South Africa. Biol Conserv. 2009;142(2):446–454. http://dx.doi.org/10.1016/j.biocon.2008.10.025

16. Pasquini L, Cowling RM, Twyman C, Wainwright J. Devising appropriate policies and instruments in support of private conservation areas: Lessons learned from the Klein Karoo, South Africa. Conserv Biol. 2010;24(2):470–478. http://dx.doi.org/10.1111/j.1523-1739.2009.01344.x

17. Colón-Rivera RJ, Marshall K, Soto-Santiago FJ, Ortiz-Torres D, Flower CE. Moving forward: Fostering the next generation of earth stewards in the STEM disciplines. Front Ecol Environ. 2013;11(7):383–391. http://dx.doi.org/10.1890/120307

18. Prakash A, Potoski M. Voluntary environmental programs: A comparative perspective. J Policy Anal Manag. 2012;31(1):123–138. http://dx.doi.org/10.1002/pam.20617

19. Hitzhusen GE, Tucker ME. The potential of religion for earth stewardship. Front Ecol Environ. 2013;11(7):368–376. http://dx.doi.org/10.1890/120322

20. Azizan MH, Wahid NA. A proposed model on environmental stewardship. Procedia Soc Behav Sci. 2012;65:587–592. http://dx.doi.org/10.1016/j.sbspro.2012.11.169

21. Nassauer JI. Care and stewardship: From home to planet. Landsc Urban Plan. 2011;100(4):321–323. http://dx.doi.org/10.1016/j.landurbplan.2011.02.022

22. Worrell R, Appleby M. Stewardship of natural resources: Definition, ethical and practical aspects. J Agric Environ Ethics. 2000;12(3):263–277. http://dx.doi.org/10.1023/A:1009534214698

23. Hobbie SE, Baker LA, Fissore C, King JY, McFadden JP, Nelson KC. Planetary stewardship begins at home. Bull Ecol Soc Am. 2011;92(4):389–391. http://dx.doi.org/10.1890/0012-9623-92.4.389

24. Crona B, Hubacek K. The right connections: How do social networks lubricate the machinery of natural resource governance? Ecol Soc. 2010;15(4), Art. #18. Available from: http://www.ecologyandsociety.org/vol15/iss4/art18/

25. Vance-Borland K, Holley J. Conservation stakeholder network mapping, analysis, and weaving. Conserv Lett. 2011;4(4):278–288. http://dx.doi.org/10.1111/j.1755-263X.2011.00176.x

26. Belaire JA, Dribin AK, Johnston DP, Lynch DJ, Minor ES. Mapping stewardship networks in urban ecosystems. Conserv Lett. 2011;4(6):464–473. http://dx.doi.org/10.1111/j.1755-263X.2011.00200.x

27. Guerrero AM, McAllister RRJ, Wilson KA. Achieving cross-scale collaboration for large scale conservation initiatives. Conserv Lett. 2015;8(2):107–117. http://dx.doi.org/10.1111/conl.12112

28. Krasny ME, Tidball KG. Civic ecology: A pathway for earth stewardship in cities. Front Ecol Environ. 2012;10(5):267–273. http://dx.doi.org/10.1890/110230

29. Peachey B. Environmental stewardship – What does it mean? Process Saf Environ. 2008;86(4):227–236. http://dx.doi.org/10.1016/j.psep.2008.02.006

30. Felson AJ, Bradford MA, Terway TM. Promoting earth stewardship through urban design experiments. Front Ecol Environ. 2013;11(7):362–367. http://dx.doi.org/10.1890/130061

31. Lertzman K. The paradigm of management, management systems, and resource stewardship. J Ethnobiol. 2009;29(2):339–358. http://dx.doi.org/10.2993/0278-0771-29.2.339

32. Olsson P, Folke C, Berkes F. Adaptive comanagement for building resilience in social-ecological systems. Environ Manage. 2004;34(1):75–90. http://dx.doi.org/10.1007/s00267-003-0101-7

33. Gray TS, Hatchard J. Environmental stewardship as a new form of fisheries governance. ICES J Mar Sci. 2007;64(4):786–792. http://dx.doi.org/10.1093/icesjms/fsl041

34. Nkhata BA, Mosimane A, Downsborough L, Breen C, Roux DJ. A typology of benefit sharing arrangements for the governance of social-ecological systems in developing countries. Ecol Soc. 2012;17(1), Art. #17. http://dx.doi.org/10.5751/ES-04662-170117

35. Kellert SR, Mehta JN, Ebbin SA, Lichtenfeld LL. Community natural resource management: Promise, rhetoric, and reality. Soc Natur Resour. 2000;13(8):705–715. http://dx.doi.org/10.1080/089419200750035575

36. Fennell DA. Ecotourism and the myth of indigenous stewardship. J Sustain Tour. 2008;16(2):129–149. http://dx.doi.org/10.2167/jost736.0

37. Atkinson R, Flint J. Accessing hidden and hard-to-reach populations: Snowball research strategies. Social Research Update. 2001;28(1):93–108.

38. Bates MJ. The design of browsing and berrypicking techniques for the online search interface. Online Info Rev. 1989;13(5):407–424. http://dx.doi.org/10.1108/eb024320

39. Choo CW, Marton C. Information seeking on the web by women in IT professions. Internet Res. 2003;13(4):267–280. http://dx.doi.org/10.1108/10662240310488951

40. Knight SA, Spink A. Toward a web search information behavior model. In: Spink A, Zimmer M, editors. Web search. Berlin: Springer; 2008. p. 209–334. http://dx.doi.org/10.1007/978-3-540-75829-7_12

41. Bastian M, Heymann S, Jacomy M. Gephi: An open source software for exploring and manipulating networks. In: Proceedings of the Third International ICWSM Conference; 2009 May 17–20; San Jose, CA, USA. Palo Alto, CA: Association for the Advancement of Artificial Intelligence; 2009. Available from: http://aaai.org/ocs/index.php/ICWSM/09/paper/view/154

42. Brandes U. A faster algorithm for betweenness centrality. J Math Soc. 2001;25(2):163–177. http://dx.doi.org/10.1080/0022250X.2001.9990249

43. Bodin Ö, Crona BI. The role of social networks in natural resource governance: What relational patterns make a difference? Glob Environ Chang. 2009;19(3):366–374. http://dx.doi.org/10.1016/j.gloenvcha.2009.05.002

44. Younge A, Fowkes S. The Cape Action Plan for the Environment: Overview of an ecoregional planning process. Biol Conserv. 2003;112(1–2):15–28. http://dx.doi.org/10.1016/S0006-3207(02)00393-2

45. Wesson J. Conservancies – The sustainable alternative. Environment. 2010:22–25.

46. National Association of Conservancies of South Africa (NACSA). Conservancies Handbook. Johannesburg: NACSA; 2003. Available from: http://www.conservancies.org/Downloads/Condensed%20Conservancy%20Handbook.pdf

47. Lindsey JD. Conserving our future. Thola. 2013:24–26.

48. CapeNature. Stewardship operational procedures manual: The wise and sustainable use of the land [document on the Internet]. c2009 [cited 2015 Nov 15]. Available from: http://www.sanbi.org/sites/default/files/documents/documents/stewardship-operational-procedures-manual.pdf.

49. Von Hase A, Rouget M, Cowling RM. Evaluating private land conservation in the Cape lowlands, South Africa. Conserv Biol. 2010;24(5):1182–1189. http://dx.doi.org/10.1111/j.1523-1739.2010.01561.x

50. Lochner P, Weaver A, Gelderblom C, Peart R, Sandwith T, Fowkes S. Aligning the diverse: The development of a biodiversity conservation strategy for the Cape Floristic Region. Biol Conserv. 2003;112(1–2):29–43. http://dx.doi.org/10.1016/S0006-3207(02)00394-4

51. Ashwell A, Sandwith T, Barnett M, Parker A, Wisani F. Fynbos Fynmense: People making biodiversity work. Pretoria: SANBI; 2006.

52. Burns A. Farmers take stewardship of their land. Veld Flora. 2007;93(1):9.

53. Reyers B, Rouget M, Jonas Z, Cowling RM, Driver A, Maze K, et al. Developing products for conservation decision-making: Lessons from a spatial biodiversity assessment for South Africa. Divers Distrib. 2007;13(5):608–619. http://dx.doi.org/10.1111/j.1472-4642.2007.00379.x

54. Rouget M, Richardson DM, Cowling RM. The current configuration of protected areas in the Cape Floristic Region, South Africa – Reservation bias and representation of biodiversity patterns and processes. Biol Conserv. 2003;112(1–2):129–145. http://dx.doi.org/10.1016/S0006-3207(02)00396-8

55. Paterson A. Tax incentives – Valuable tools for biodiversity conservation in South Africa. S Afr Law J. 2005;122:182.

56. Van Wyk E. Tax incentives for biodiversity conservation in the Western Cape. Meditari Account Res. 2010;18(1):58–75. http://dx.doi.org/10.1108/10222529201000005

57. National Environmental Management: Protected Areas Act 57 of 2003, South Africa. Gazette no. 26025, notice no. 181.393.

58. Mortimer G. Cape Action Plan for People and the Environment, CAPE Stewardship Project [document on the Internet]. No date [cited 2015 Nov 25]. Available from: http://www.capeaction.org.za/index.php/resources/effective-protection?view=document&id=61.

59. Pool-Stanvliet R. A history of the UNESCO Man and the Biosphere Programme in South Africa. S Afr J Sci. 2013;109(9/10), Art. #a0035, 6 pages. http://dx.doi.org/10.1590/sajs.2013/a0035

60. United Nations Educational, Scientific and Cultural Organization [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://whc.unesco.org/en/statesparties/za

61. Lombard AT, Cowling RM, Vlok JHJ, Fabricius C. Designing conservation corridors in production landscapes: Assessment methods, implementation issues, and lessons learned. Ecol Soc. 2010;15(3), Art. #7. Available from: http://www.ecologyandsociety.org/vol15/iss3/art7/

62. Markham R. Conserving biodiversity outside protected areas: The Eden to Addo case. Environment. 2013:20–24.

63. Open Africa. Segarona heritage experience [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.openafrica.org/experiences/route/103-segarona-heritage-experience

64. Rim of Africa. Cape Mountains and trail initiative [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.rimofafrica.co.za/Rim_of_Africa/Welcome.html

65. Gouritz Cluster Biosphere Reserve [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.gouritz.com

66. Felix Toteu S, Malcolm Anderson J, De Wit M. 'Africa Alive Corridors': Forging a new future for the people of Africa by the people of Africa. J Afr Earth Sci. 2010;58(4):692–715. http://dx.doi.org/10.1016/j.jafrearsci.2010.08.011

67. Mission Blue Sylvia Earle Alliance. Hope Spots [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://mission-blue.org/hope-spots-new/

68. Sustainable Seas Trust. Hope Spots [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.sst.org.za/hope-spots/

69. Blackman A, Rivera J. Producer-level benefits of sustainability certification. Conserv Biol. 2011;25(6):1176–1185. http://dx.doi.org/10.1111/j.1523-1739.2011.01774.x

70. Wildlife and Environment Society of South Africa. Blue Flag [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://blueflag.org.za/index.php/25-welcome-to-blue-flag

71. Landmark Foundation. Fair game – Wildlife friendly products [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.landmarkfoundation.org.za/fair-game/

72. Eco Mark Africa [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.ecomarkafrica.com/

73. Honig MBH. Towards understanding private conservation in the Cape Winelands of South Africa: Developing a theory of change. Budapest: Central European University; 2012.

74. McEwan C, Bek D. The political economy of alternative trade: Social and environmental certification in the South African wine industry. J Rural Stud. 2009;25(3):255–266. http://dx.doi.org/10.1016/j.jrurstud.2009.03.001

75. Rainforest Alliance. Red bush tea goes green [homepage on the Internet]. No date [cited 2015 Nov 25]. Available from: http://www.rainforest-alliance.org/newsroom/press-releases/rooibos-release

76. Petersen C. The business case for biodiversity and good biodiversity practice in the Republic of South Africa. Pretoria: South African National Biodiversity Institute; 2007. http://dx.doi.org/10.5962/bhl.title.66305

77. Pence GQK. Contribution of C.A.P.E. Business and Biodiversity Initiatives to conservation of critical biodiversity, landscape connectivity and ecological support areas. Cape Town: Conservation South Africa; 2011.

78. Turpie JK, Marais C, Blignaut JN. The Working for Water programme: Evolution of a payments for ecosystem services mechanism that addresses both poverty and ecosystem service delivery in South Africa. Ecol Econ. 2008;65(4):788–798. http://dx.doi.org/10.1016/j.ecolecon.2007.12.024

79. Hobbs RJ. The Working for Water programme in South Africa: The science behind the success. Divers Distrib. 2004;10(5–6):501–503. http://dx.doi.org/10.1111/j.1366-9516.2004.00115.x

80. Woodworth P. Working for Water in South Africa: Saving the world on a single budget? World Policy J. 2006;23(2):31–43.

81. Van Wilgen BW, Forsyth GG, Le Maitre DC, Wannenburgh A, Kotzé JDF, Van den Berg E, et al. An assessment of the effectiveness of a large, national-scale invasive alien plant control strategy in South Africa. Biol Conserv. 2012;148(1):28–38. http://dx.doi.org/10.1016/j.biocon.2011.12.035

82. Department of Environmental Affairs. Projects and programmes [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: https://www.environment.gov.za/projectsprogrammes#workingfor.

83. Department of Agriculture, Forestry and Fisheries. LandCare [homepage on the Internet]. No date [cited 2015 Sep 15]; Available from: http://www.daff.gov.za/daffweb3/Programmes/LandCare

84. Hurly KM. Sustainability for a sugarcane grower in the South African sugar industry –Can SUSFARMS® add value? In: Proceedings of the 19th International Farm Management Congress; 2013 July 21–26; Warsaw, Poland. Warsaw: International Farm Management Association; 2013. Available from: http://ifmaonline.org/proceedings/19th-vol3/

85. South African Breweries. Water stewardship in the hops industry [document on the Internet]. No date [cited 2015 Sep 15]. Available from: https://www.sabmiller.com/docs/default-source/sustainability-documents/water-stewardship-in-the-hops-industry

86. Whitehead K. M&S, Ecolab and General Mills adopt water stewardship standard. Green Futures [magazine on the Internet]. 2014 May 30. Available from: http://www.thefuturescentre.org/articles/1976/ms-ecolab-and-general-mills-adopt-water-stewardship-standard

87. WWF. Water stewardship experience in the Western Cape. Cape Town: WWF-SA; 2014. Available from: http://awsassets.wwf.org.za/downloads/water_stewardship_final.pdf

88. Basson J. Not all seafood is equal. S Afr J Sci. 2011;107(5/6):8–10. http://dx.doi.org/10.4102/sajs.v107i5/6.718

89. Sink KJ, Barendse J, Bürgener M, Nel DC. Driving changes for healthy oceans: Lessons from the Southern African Sustainable Seafood Initiative. Poster presented at: 5th Western Indian Ocean Marine Science Association (WIOMSA) Scientific Symposium; 2007 October 22–26; Durban, South Africa.

90. Barendse J, Francis J. Towards a standard nomenclature for seafood species to promote more sustainable seafood trade in South Africa. Mar Policy. 2015;53:180–187. http://dx.doi.org/10.1016/j.marpol.2014.12.007

91. Wildlife and Environment Society of South Africa. WESSA Eco-Schools [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://wessa.org.za/what-we-do/eco-schools.htm

92. Sherriffs P. Good food foodie. S Afr Food Rev. 2011;38(1):38–39. Available from: http://reference.sabinet.co.za/webx/access/electronic_journals/im_safr/im_safr_v38_n1_a27.pdf

93. Fourie A, Muller K. Innovations in governance for biodiversity conservation: The case of the Conservation Stewardship Programme, Western Cape. Administratio Publica. 2011;19(1):88–103.

94. Pasquini L, Twyman C, Wainwright J. Toward a conceptual framework for blending social and biophysical attributes in conservation planning: A case-study of privately-conserved lands. Environ Manage. 2010;46(5):659–670. http://dx.doi.org/10.1007/s00267-010-9548-5

95. Rossouw AS. Towards developing an understanding of biodiversity stewardship in the city of Cape Town [MEnvDev thesis]. Pietermaritzburg: University of Kwazulu-Natal; 2012.

96. Walker CI. Stewardship as an educational process of social learning and change: Two case studies conducted in the Western Cape [master's thesis]. Grahamstown: Rhodes University; 2011.

97. Crona BI, Parker JN. Learning in support of governance: Theories, methods, and a framework to assess how bridging organizations contribute to adaptive resource governance. Ecol Soc. 2012;17(1), Art. #32. http://dx.doi.org/10.5751/ES-04534-170132

98. Connolly JJ, Svendsen ES, Fisher DR, Campbell LK. Organizing urban ecosystem services through environmental stewardship governance in New York City. Landscape Urban Plan. 2013;109(1):76–84. http://dx.doi.org/10.1016/j.landurbplan.2012.07.001

99. Mills M, Álvarez-Romero JG, Vance-Borland K, Cohen P, Pressey RL, Guerrero AM, et al. Linking regional planning and local action: Towards using social network analysis in systematic conservation planning. Biol Conserv. 2014;169:6–13. http://dx.doi.org/10.1016/j.biocon.2013.10.015

100. Driver A, Maze K. The Succulent Karoo Ecosystem Plan [SKEP]. An introduction to SKEP. Veld Flora. 2002;88(1):12–13.

101. Knight AT, Cowling RM, Boshoff AF, Wilson SL, Pierce SM. Walking in STEP: Lessons for linking spatial prioritisations to implementation strategies. Biol Conserv. 2011;144(1):202–211. http://dx.doi.org/10.1016/j.biocon.2010.08.017

102. Stephens A. Making biodiversity stewardship work. BC Grasslands. 2009:32–34.

103. Ostrom E. The challenge of common-pool resources. Environ Sci Policy Sustain Develop. 2008;50(4):8–21. http://dx.doi.org/10.3200/ENVT.50.4.8-21

104. Turnipseed M, Sagarin R, Barnes P, Blumm MC, Parenteau P, Sand PH. Reinvigorating the public trust doctrine: Expert opinion on the potential of a public trust mandate in U.S. and international environmental law. Environ Sci Policy Sustain Develop. 2010;52(5):6–14. http://dx.doi.org/10.1080/00139157.2010.508666

105. McKinley E, Fletcher S. Improving marine environmental health through marine citizenship: A call for debate. Mar Policy. 2012;36(3):839–843. http://dx.doi.org/10.1016/j.marpol.2011.11.001

106. Gilmour PW, Day RW, Dwyer PD. Using private rights to manage natural resources: Is stewardship linked to ownership? Ecol Soc. 2012;17(3), Art. #1. http://dx.doi.org/10.5751/ES-04770-170301

107. Gutiérrez NL, Valencia SR, Branch TA, Agnew DJ, Baum JK, Bianchi PL, et al. Eco-label conveys reliable information on fish stock health to seafood consumers. PLoS One. 2012;7(8),e43765, 8 pages. http://dx.doi.org/10.1371/journal.pone.0043765

108. Branch GM, Clark BM. Fish stocks and their management: The changing face of fisheries in South Africa. Mar Policy. 2006;30(1):3–17. http://dx.doi.org/10.1016/j.marpol.2005.06.009

109. Mengerink KJ, Van Dover CL, Ardron J, Baker M, Escobar-Briones E, Gjerde K, et al. A call for deep-ocean stewardship. Science. 2014;344(6185):696. http://dx.doi.org/10.1126/science.1251458

110. Von Zharen WM. Ocean ecosystem stewardship. Wm Mary Envtl L Pol Rev. 1998;23(1):109.

111. Timberwatch Coalition [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.timberwatch.org.za/old_site/

112. Heupel E, Auster PJ. Eco-labeling seafood: Addressing impacts to vulnerable seafloor species, communities, habitats and ecosystems in data-poor regions. Mar Policy. 2013;38:8–15. http://dx.doi.org/10.1016/j.marpol.2012.05.014

113. Mason NWH, Ausseil A-GE, Dymond JR, Overton JM, Price R, Carswell FE. Will use of non-biodiversity objectives to select areas for ecological restoration always compromise biodiversity gains? Biol Conserv. 2012;155:157–168. http://dx.doi.org/10.1016/j.biocon.2012.05.019

114. Bramston P, Pretty G, Zammit C. Assessing environmental stewardship motivation. Environ Behav. 2010;43(6):776–788. http://dx.doi.org/10.1177/0013916510382875

115. Simaika JP, Samways MJ. Biophilia as a universal ethic for conserving biodiversity. Conserv Biol. 2010;24(3):903–906. http://dx.doi.org/10.1111/j.1523-1739.2010.01485.x

116. Downsborough L, Shackleton CM, Knight AT. The potential for voluntary instruments to achieve conservation planning goals: The case of conservancies in South Africa. Oryx. 2011;45(03):357–364. http://dx.doi.org/10.1017/S0030605310001559

117. Conradie B, Treurnicht M, Esler K, Gaertner M. Conservation begins after breakfast: The relative importance of opportunity cost and identity in shaping private landholder participation in conservation. Biol Conserv. 2013;158:334–341. http://dx.doi.org/10.1016/j.biocon.2012.08.028

118. Honig M, Petersen S, Shearing C, Pintér L, Kotze I. The conditions under which farmers are likely to adapt their behaviour: A case study of private land conservation in the Cape Winelands, South Africa. Land Use Policy. 2015;48:389–400. http://dx.doi.org/10.1016/j.landusepol.2015.06.016

119. Taylor PL. In the market but not of it: Fair Trade coffee and Forest Stewardship Council Certification as market-based social change. World Develop. 2005;33(1):129–147. http://dx.doi.org/10.1016/j.worlddev.2004.07.007

120. Morris M, Dunne N. Driving environmental certification: Its impact on the furniture and timber products value chain in South Africa. Geoforum. 2004;35(2):251–266. http://dx.doi.org/10.1016/j.geoforum.2003.09.006

121. Ponte S. The Marine Stewardship Council (MSC) and the making of a market for 'sustainable fish'. J Agr Change. 2012;12(2–3):300–315. http://dx.doi.org/10.1111/j.1471-0366.2011.00345.x

122. Turpie JK. The existence value of biodiversity in South Africa: How interest, experience, knowledge, income and perceived level of threat influence local willingness to pay. Ecol Econ. 2003;46(2):199–216. http://dx.doi.org/10.1016/S0921-8009(03)00122-8

123. Dos Santos MAO. Minimizing the business impact on the natural environment: A case study of Woolworths South Africa. Eur Bus Rev. 2011;23(4):384–391. http://dx.doi.org/10.1108/09555341111145762

124. Milne MJ, Kearins K, Walton S. Creating adventures in wonderland: The journey metaphor and environmental sustainability. Organization. 2006;13(6):801–839. http://dx.doi.org/10.1177/1350508406068506

125. Maree BA, Wanless RM, Fairweather TP, Sullivan BJ, Yates O. Significant reductions in mortality of threatened seabirds in a South African trawl fishery. Anim Conserv. 2014;17(6):520–529. http://dx.doi.org/10.1111/acv.12126

126. Nahman A, Rigby D. Valuing blue flag status and estuarine water quality in Margate, South Africa. S Afr J Econ. 2008;76(4):721–737. http://dx.doi.org/10.1111/j.1813-6982.2008.00208.x

127. Lindsey PA, Romañach SS, Davies-Mostert HT. The importance of conservancies for enhancing the value of game ranch land for large mammal conservation in southern Africa. J Zool. 2009;277(2):99–105. http://dx.doi.org/10.1111/j.1469-7998.2008.00529.x

128. Pence GQK. Contribution of C.A.P.E. Business and Biodiversity Initiatives to conservation of critical biodiversity, landscape connectivity and ecological support areas: Post-baseline assessment 2010. Cape Town: Conservation South Africa; 2012.

129. Cowling R, Van Wilgen B, Kraaij T, Britton J. How no-man's-land is now everyone's problem. Veld Flora. 2009;95(3):147–149.

130. Van der Schyff E. Stewardship doctrines of public trust: Has the eagle of public trust landed on South African soil? S Afr Law J. 2013;130(2):369–389.

131. Benkenstein A. Seabed mining: Lessons from the Namibian experience [homepage on the Internet]. c2014 [cited 2014 Oct 20]. Available from http://www.saiia.org.za/policy-briefings/seabed-mining-lessons-from-the-namibian-experience/

132. Barbee J, Smith D. Mining poses new threat to world's greatest rhino sanctuary. The Guardian. 2014 July 17; Conservation. Available from: http://www.theguardian.com/environment/2014/jul/17/-sp-mining-threat-south-africa-rhino-sanctuary-poaching

133. Tsitsikamma MPA shock. Fishing Industry News Southern Africa. 2007;8(2):8–10.

134. Gosling M. Mixed feelings over Tsitsikamma fishing. IOL. 2015 December 02; SciTech/Science/Environment. Available from: http://www.iol.co.za/scitech/science/environment/mixed-feelings-over-tsitsikamma-fishing-1954235

135. Department of Environmental Affairs. Government Gazette no. 39423. 2015 November 19. Available from: http://www.gov.za/sites/www.gov.za/files/39423_gon1145.pdf

136. Centre for Environmental Rights [cited 2015 Sep 15]. Available from: http://cer.org.za/

137. Field JG, Attwood CG, Jarre A, Sink K, Atkinson LJ, Petersen S. Cooperation between scientists, NGOs and industry in support of sustainable fisheries: The South African hake *Merluccius* spp. trawl fishery experience. J Fish Biol. 2013;83:1019–1034.

138. Planting S. Slippery business. State of the fishing industry. Financial Mail. 2010 June 18. Available from: http://www.financialmail.co.za/fm/2010/06/17/state-of-the-fishing-industry

139. Safeguard our Seabed Coalition. Bulk marine sediment mining in South Africa's marine environment [document on the Internet]. c2015 [cited 2015 Nov 25]. Available from: http://sadstia.co.za/images/SOSC_Letter_to_Minister_Mineral_Resources_-_30-07-2015.pdf

140. Labuschagne R. Independent group tackle abalone poaching in the Eastern Cape, South Africa [video on the Internet]. c2013 [cited 2015 Nov 25]. Available from: www.youtube.com/watch?v=HxlFVvyuDjs

141. Raemaekers S, Hauck M, Bürgener M, Mackenzie A, Maharaj G, Plagányi ÉE, et al. Review of the causes of the rise of the illegal South African abalone fishery and consequent closure of the rights-based fishery. Ocean Coast Manage. 2011;54(6):433–445. http://dx.doi.org/10.1016/j.ocecoaman.2011.02.001

142. White I, Melville M, Macdonald B, Quirk R, Hawken R, Tunks M, et al. From conflicts to wise practice agreement and national strategy: Cooperative learning and coastal stewardship in estuarine floodplain management, Tweed River, eastern Australia. J Cleaner Product. 2007;15(16):1545–1558. http://dx.doi.org/10.1016/j.jclepro.2006.07.049

143. Hatchwell M. Public–private partnerships as a management option for protected areas. Anim Conserv. 2014;17(1):3–4. http://dx.doi.org/10.1111/acv.12098

144. McConnachie MM, Cowling RM, Van Wilgen BW, McConnachie DA. Evaluating the cost-effectiveness of invasive alien plant clearing: A case study from South Africa. Biol Conserv. 2012;155:128–135. http://dx.doi.org/10.1016/j.biocon.2012.06.006

145. Urgenson LS, Prozesky HE, Esler KJ. Stakeholder perceptions of an ecosystem services approach to clearing invasive alien plants on private land. Ecol Soc. 2013;18(1), Art. #26. http://dx.doi.org/10.5751/ES-05259-180126

146. Rissman AR, Lozier L, Comendant T, Kareiva P, Kiesecker JM, Shaw MR, et al. Conservation easements: Biodiversity protection and private use. Conserv Biol. 2007;21(3):709–718. http://dx.doi.org/10.1111/j.1523-1739.2007.00660.x

147. Newburn D, Reed S, Berck P, Merenlender A. Economics and land-use change in prioritizing private land conservation. Conserv Biol. 2005;19(5):1411–1420. http://dx.doi.org/10.1111/j.1523-1739.2005.00199.x

148. Müller JJ. Assessing cooperative environmental governance systems: The cases of the Kogelberg Biosphere Reserve and the Olifants-Doorn Catchment Management Agency. Politeia. 2008;27(1):86–104.

149. Armitage D. Adaptive capacity and community-based natural resource management. Environ Manage. 2005;35(6):703–715. http://dx.doi.org/10.1007/s00267-004-0076-z

150. Breen CM, editor. Community based natural resource management in southern Africa: An introduction. Bloomington, IN: AuthorHouse; 2013.

151. Hibbard M, Lurie S. Creating socio-economic measures for community-based natural resource management: A case from watershed stewardship organisations. J Environ Plan Manage. 2011;55(4):525–544. http://dx.doi.org/10.1080/09640568.2011.614093

152. Ayoo C. Community-based natural resource management in Kenya. Manag Environ Qual. 2007;18(5):531–541. http://dx.doi.org/10.1108/14777830710778292

153. Musumali MM, Larsen TS, Kaltenborn BP. An impasse in community based natural resource management implementation: The case of Zambia and Botswana. Oryx. 2007;41(3):306–313.

154. Agriculture Investment Sourcebook. Washington DC: World Bank; 2005.

155. Gruber JS. Perspectives of effective and sustainable community-based natural resource management: An application of Q Methodology to forest projects. Conserv Soc. 2011;9(2):159–171. http://dx.doi.org/10.4103/0972-4923.83725

156. Svendsen E, Campbell LK. Urban ecological stewardship: Understanding the structure, function and network of community-based urban land management. Cities & The Environment. 2008;1(1):4.

157. Cocks ML, Dold T, Vetter S. 'God is my forest': Xhosa cultural values provide untapped opportunities for conservation. S Afr J Sci. 2012;108(5–6):52–59. http://dx.doi.org/10.4102/sajs.v108i5/6.880

158. Bernard PS, Kumalo S. Community-based natural resource management, traditional governance and spiritual ecology in southern Africa: The case of chiefs, diviners and spirit mediums. In: Fabricius C, Koch E, Magome H, Turner S, editors. Rights, resources and rural development: Community-based natural resource management in southern Africa. London: Earthscan; 2004. p. 115–126.

159. Groves CR, Jensen DB, Valutis LL, Redford KH, Shaffer ML, Scott JM, et al. Planning for biodiversity conservation: Putting conservation science into practice. BioScience. 2002;52(6):499–512. http://dx.doi.org/10.1641/0006-3568(2002)052[0499:PFBCPC]2.0.CO;2

160. Heydenrych BJ, Cowling RM, Lombard AT. Strategic conservation interventions in a region of high biodiversity and high vulnerability: A case study from the Agulhas Plain at the southern tip of Africa. Oryx. 1999;33(3):256–269. http://dx.doi.org/10.1017/S003060530003060X

161. Sayre NF, Kelty R, Simmons M, Clayton S, Kassam K-A, Pickett STA, et al. Invitation to earth stewardship. Front Ecol Environ. 2013;11(7):339. http://dx.doi.org/10.1890/1540-9295-11.7.339

162. Ogden L, Heynen N, Oslender U, West P, Kassam K-A, Robbins P. Global assemblages, resilience, and earth stewardship in the Anthropocene. Front Ecol Environ. 2013;11(7):341–347. http://dx.doi.org/10.1890/120327

163. Agrawal A, Ostrom E. Collective action, property rights, and decentralization in resource use in India and Nepal. Polit Soc. 2001;29(4):485–514. http://dx.doi.org/10.1177/0032329201029004002

164. Galliers C, Barnes G. Earth stewardship: A new approach to the protection of South Africa's natural capital. Environment. 2013:62–64.

165. Wildlands Conservation Trust [homepage on the Internet]. No date [cited 2015 Sep 15]. Available from: http://www.wildlands.co.za

166. Hoole A, Berkes F. Breaking down fences: Recoupling social–ecological systems for biodiversity conservation in Namibia. Geoforum. 2010;41(2):304–317. http://dx.doi.org/10.1016/j.geoforum.2009.10.009

167. Kinzig AP, Ehrlich PR, Alston LJ, Arrow K, Barrett S, Buchman TG, et al. Social norms and global environmental challenges: The complex interaction of behaviors, values, and policy. BioScience. 2013;63(3):164–175. http://dx.doi.org/10.1525/bio.2013.63.3.5

168. Leménager T, King D, Elliott J, Gibbons H, King A. Greater than the sum of their parts: Exploring the environmental complementarity of state, private and community protected areas. Global Ecol Conserv. 2014;2(0):238–247. http://dx.doi.org/10.1016/j.gecco.2014.09.009

Appendix: Types of stewardship nodes and their labels shown in Figure 1

Node label	Full name	Type
Alliance for Water Stewardship	Alliance for Water Stewardship	Partnership
Audubon	Audubon International	Non-governmental organisation (NGO)
Biodiversity Stewardship SA	Biodiversity Stewardship South Africa	Partnership
BioNET	BioNET	Partnership
Birdlife SA	Birdlife South Africa	NGO
BotSoc	Botanical Society of South Africa	NGO
CAPE	Cape Action Plan for People and the Environment	Partnership
CAP	Climate Action Partnership	Partnership
Cape Leopard Trust	Cape Leopard Trust	NGO
CapeNature	CapeNature	Government
CEBA	Community Ecosystems Based Adaptation	Partnership
CEPF	Critical Ecosystem Partnership Fund	Fund
City of Cape Town	City of Cape Town	Government
CoastCare	CoastCare	Government
Conservation@Work	Conservation at Work	NGO
Conservation SA	Conservation South Africa	NGO
CREW	Custodians of Rare and Endangered Wildflowers	NGO
DAFF	Department of Agriculture, Forestry & Fisheries	Government
DWA	Department of Water Affairs	Government
DEA	Department of Environmental Affairs	Government
Dept. of Rural Development & Land Reform	Department of Rural Development and Land Reform	Government
EarthCollective	EarthCollective	Partnership
Eastern Cape BSP	Eastern Cape Biodiversity Stewardship Programme	Partnership
Eastern Cape Parks & Tourism Agency	Eastern Cape Parks and Tourism Agency	Government
EKZN Wildlife	Ezemvelo KwaZulu-Natal Wildlife	Government
EWT	Endangered Wildlife Trust	NGO
EPWP	Expanded Public Works Programme	Partnership
Fauna & Flora Int.	Fauna and Flora International	NGO
FEE	Foundation for Environmental Education	NGO
Flower Valley Conservation Trust	Flower Valley Conservation Trust	NGO
Food & Trees for Africa	Food and Trees for Africa	NGO
Forestry South Africa	Forestry South Africa	Private
Gauteng BSP	Gauteng Biodiversity Stewardship Programme	Partnership
Gauteng Conservancy & Stewardship Assoc.	Gauteng Conservancy and Stewardship Association	Partnership
Gauteng Dept. of Agri. & Rural Develop.	Gauteng Department of Agriculture and Rural Development	Government
GEF	Global Environmental Facility	Fund
GIZ	Deutsche Gesellschaft for Internationale Zusammenarbeit	Fund
Grassland Prog.	Grassland Programme	Partnership
Green Trust	Green Trust	Fund

Node label	Full name	Type
GreenChoice	GreenChoice Alliance	Partnership
Heritage	Heritage Environmental Management Company	Private
Hope Spots	International Hope Spots	Partnership
I Am Water	I Am Water Ocean Conservation Trust	NGO
IUCN	World Conservation Union	NGO
KZN BSP	KZN Biodiversity Stewardship Programme	Partnership
KZN Conservancies Assoc.	KwaZulu-Natal Conservancies Association	Partnership
KZN Crane Foundation	KwaZulu-Natal Crane Foundation	NGO
Land Reform & BSP Prog.	Land Reform and Biodiversity Stewardship Programme	Partnership
LandCare	National LandCare Programme	Partnership
Landmark Foundation	Landmark Foundation	NGO
Limpopo Conservancy Assoc.	Limpopo Conservancy Association	Partnership
Living Lands	Living Lands	NGO
MAB	Man and the Biosphere Programme	Partnership
Midlands Conservancies Forum	Midlands Conservancies Forum	NGO
Mission Blue	Mission Blue	NGO
Mpumalanga BSP	Mpumalanga Biodiversity Stewardship	Partnership
Mpumalanga Tourism & Parks	Mpumalanga Tourism and Parks Agency	Government
NACSSA	National Association of Conservancies/Stewardship South Africa	Partnership
NCC	Nature Conservation Corporation	Private
North West Conservancy Assoc.	North West Conservancy Association	Partnership
North West Parks & Tourism	North West Parks and Tourism	Government
NVT	Nature's Valley Trust	NGO
Peace Parks	Peace Parks Foundation	NGO
PnP	Pick n Pay	Private
Rim of Africa	Rim of Africa Initiative	NGO
SA Rooibos	SA Rooibos Council	Private
SAB Miller	SAB Miller	Private
SADSTIA	SA Deep Sea Trawling Industry Association	Private
SANParks	South African National Parks	Government
Save Our Seas Foundation	Save Our Seas Foundation	NGO
SKEP	Succulent Karoo Ecosystem Plan	Partnership
SANBI	South African National Biodiversity Institute	Government
South African Shark Conservancy	South African Shark Conservancy	NGO
Spaces for Elephants	Spaces for Elephants Foundation	NGO
SST	Sustainable Seas Trust	NGO
STEP	Succulent Thicket Ecosystem Plan	Partnership
Sustainable Tourism Certification Alliance	Sustainable Tourism Certification Alliance	Partnership
Sustainable Tourism Partnership Prog.	Sustainable Tourism Partnership Programme	NGO
Sustaining the Wild Coast	Sustaining the Wild Coast	NGO

Node label	Full name	Type
Timberwatch	Timberwatch	NGO
TMF	Table Mountain Fund	Fund
TRAFFIC	TRAFFIC East/Southern Africa	NGO
Two Oceans	Two Oceans Aquarium	Private
UNDP SA	United Nations Development Program South Africa	NGO
UNESCO	United Nations Educational, Scientific and Cultural Organisation	NGO
uShaka	uShaka Marine World	Private
WAG	Wilderness Action Group	NGO
Water Stewardship Council Trust of SA	Water Stewardship Council Trust of South Africa	NGO
WESSA	Wildlife and Environment Society of South Africa	NGO
Western Cape BSP	Western Cape Biodiversity Stewardship Programme	Partnership
Wilderness Foundation	Wilderness Foundation	NGO
Wildlands Cons. Trust	Wildlands Conservation Trust	NGO
Woolworths	Woolworths	Private
WWF	World Wide Fund for Nature (South Africa)	NGO

Note: This articles includes supplementary material.

Plagiarism and ghostwriting: The rise in academic misconduct

AUTHORS:
Shawren Singh[1]
Dan Remenyi[1]

AFFILIATION:
[1]School of Computing, University of South Africa, Roodepoort, South Africa

CORRESPONDENCE TO:
Shawren Singh

EMAIL:
singhs@unisa.ac.za

POSTAL ADDRESS:
School of Computing, University of South Africa, Cnr Christiaan de Wet Road and Pioneer Ave, Florida 1710, South Africa

DATES:
Received: 12 Aug. 2015
Revised: 12 Oct. 2015
Accepted: 21 Dec. 2015

KEYWORDS:
academic integrity; contract cheating; essay mills; university cheating

HOW TO CITE:
Singh S, Remenyi D. Plagiarism and ghostwriting: The rise in academic misconduct. S Afr J Sci. 2016;112(5/6), Art. #2015-0300, 7 pages. http://dx.doi.org/10.17159/sajs.2016/20150300

The aim of this paper is to review the current situation regarding plagiarism and ghostwriting, and to stimulate debate about how universities should respond to the rise in these forms of academic misconduct. The apparent upsurge in academic misconduct means that universities today face one of the greatest challenges to academic integrity they have had to deal with ever since the university system came into existence some 800 years ago. Plagiarism and ghostwriting are undermining the integrity of university degrees to an extent not seen before. Academia and fraud are not strangers. Universities have a long history of cheating of one sort or another, often associated with examinations, but also with research. In the past this cheating involved activities such as smuggling notes (commonly called 'crib sheets') into examinations, and consulting them even under the watchful eyes of invigilators. It also involved students obtaining sight of an examination paper in advance. The fraudulent creation of research results has also been an issue. However, in the 21st century, the opportunities for cheating have exploded. This has resulted in universities becoming more concerned about ensuring the integrity of their examination processes and the degrees they award. Our paper focuses on cheating in the writing of dissertations or theses required at undergraduate or postgraduate level, with an emphasis on plagiarism and ghostwriting. We do not propose a simple solution to these problems, as preventing or stopping cheating is not just a matter of catching the wrongdoers. Cheating is endogenous to the current university education system, and needs to be addressed in terms of not only prevention and detection but also how people who are found to engage in such misconduct are treated. We suggest that creative ways of promoting learning would help to minimise cheating at universities. It is also important to ensure that the issue is discussed openly among students and faculty staff.

Introduction

There have been some dramatic instances of academic fraud at universities.[1-3] Professors have been found to have no credible academic credentials, having either exaggerated or outright lied in their curriculum vitae.[4-6] Vice-chancellors have been accused of plagiarism[7,8] or been found to have plagiarised their theses.[9] Laboratory directors have been involved in the falsification of research findings.[10-12] However, such dramatic events, although interesting and often newsworthy, are fortunately quite rare. What is more concerning is the mundane matter of fraud or cheating at the routine examination level. The main form of cheating at university is inappropriate assistance in examinations, or in the preparation of written work submitted for evaluation.[13]

The extent of cheating at universities is hard to gauge.[14-17] This is largely because the most common reaction once cheating is exposed is that the institution becomes secretive. Over the years, students have been found consulting crib sheets or notes written on their skin, or hidden in pencil boxes or even sandwiches – to mention only a few hiding places. Students struggling to answer a question might also try to glance at the answer sheet of a fellow student.

Another form of cheating occurs when students have been informed about examination questions in advance. Perhaps more seriously, students have sometimes employed other people to actually sit their examinations. The most famous culprit of this offence was the late Senator Edward Kennedy, who paid a co-student and friend to sit a Spanish examination in his place. The university spotted the substitution and both students were expelled.[18]

However, despite the wide range of nefarious ways in which students have been known to cheat the system, academics often argue that traditionally there has been a low occurrence of this type of behaviour. An accurate estimate of the number of such incidents is almost impossible, as the only cases reported are those in which individuals are caught.

Cheating in examinations is difficult. It is hard to smuggle items into an examination hall and consult them unseen – despite the watchful eye of an invigilator. Nonetheless, some students try to do this. To counter the problem, at certain universities some subjects are examined in 'open book' exams, where the student may bring into the examination room any texts he or she might like to consult. This practice allows the student to refer to any material he or she wishes during the examination. In such exams, the application and interpretation of knowledge is tested, and for this reason open-book exams are often regarded as a superior form of test.

In many universities there has been a substantial increase in the use of term papers, assignments, and dissertations to evaluate the progress or knowledge development of students. Such documents are produced by the student outside of an examination environment. One reason for this system is the now widespread belief in the value of ongoing assessment instead of simply an end-of-term examination. However, when this type of written work is used for assessment, the system is especially vulnerable to cheating.[19] Because of this vulnerability, some older academics insist that formal examinations remain the only reliable method of student evaluation.

All forms of academic cheating are highly detrimental to any university. Academic cheating undermines the good name of the institution and calls into question the integrity of both the faculty and students. There is every reason for a university to take all forms of cheating seriously, and to eliminate it wherever possible.

This paper focuses on two specific types of academic misconduct, namely plagiarism and ghostwriting.

Inappropriate assistance in preparing written work

There are several ways in which students can cheat in the preparation of written work. The two offences addressed in this paper are plagiarism and ghostwriting. Plagiarism has affected the academic community for centuries,[20] and allegations of plagiarism have been made against certain famous academics, including Galileo and Newton.[21] By contrast, ghostwriting in academia appears to be a relatively modern phenomenon, perhaps only a few hundred years old.

Stavisky[22] states that there has been a long tradition of plagiarism in American universities, fostered by fraternity groups and dating back to the 19th century. Fraternity files have been used to recycle written academic work. Stavisky also states that in the 1940s, advertisements appeared weekly in a prestigious New York newspaper, advertising ghostwriting services – which included producing dissertations, theses, and term papers. Stavisky goes on to describe how this practice proliferated in the 1960s and 1970s. With the arrival of the Internet, ghostwriting has become a global industry. The terms 'paper mill' and 'essay mill' are often used to describe this industry.

Plagiarism and ghostwriting, although different, have the same outcome: the student presents fraudulent academic work that is purportedly his or her own. In fact, it is the creation of another person or persons. The sections below discuss plagiarism and ghostwriting separately and in greater detail.

Plagiarism

Plagiarism refers to the use of other people's ideas and words without giving the original author appropriate acknowledgement.[23,24] If ideas are used in an essay or dissertation that have been found in the published work of another author, it is academic misconduct not to specifically acknowledge the original source. The acknowledgement must follow the rules of the referencing system employed in the work. Although the use of ideas without acknowledging them is an offence, it is even worse if the actual words of other authors are copied without acknowledgement. This principle sometimes leads to debate about how many words can be cited without incurring an accusation of plagiarism. Guidelines such as three, four, or five words are sometimes quoted. However, there is no simple answer to this question, and it is generally agreed that even a small number of words reproduced from another text need to be attributed.

There are several reasons why plagiarism is unacceptable in academic writing. The Penn State University website[25] lists a number of reasons, which include:

- Plagiarism committed intentionally is an act of deceit and may even constitute fraud.
- The plagiarist denies him or herself 'the opportunity to learn and practice' the skills of academic research.
- A plagiarist does not avail him or herself of the 'opportunity to receive honest feedback' on his or her academic skills.
- The plagiarist opens him or herself to future enquiry into his or her 'integrity and performance in general'.

Clearly plagiarism is unacceptable. Universities generally state in their regulations that plagiarism is a disciplinary offence.[26] However, it is not always easy to ascertain what sort of penalty will be imposed on authors who are found to have plagiarised.

In recent decades, the Internet has become a common tool for academic research, and this has enabled plagiarism to flourish on a large scale.[27-29] It is impossible to estimate the exact extent of online plagiarism, but there are regular reports of students' work being found to contain large passages copied from other people's works by cut-and-paste methods, without any attribution. In 2015, Adams reported that in the United Kingdom, 'in the past 4 years more than 58 000 undergraduates have been investigated by their universities for plagiarism'[30]. Some of these cases involve whole essays being copied and fraudulently presented under the name of the student being assessed.

Because of this situation, computer-based plagiarism detection methods and tools have become extremely popular with university faculty and administrative staff.[31] The market leader, Turnitin.com, claims to have 10 000 clients working in 135 countries, and this product alone is estimated as being used to check 40 million academic papers each year. Using anti-plagiarism software makes it relatively easy to detect and quantify how much plagiarism appears in a piece of academic work. However, this type of checking may be considered to be a 'band aid' or 'sticky plaster' placed on a wound when perhaps a more medical and surgical intervention is actually required.[32]

When plagiarism is suspected and this suspicion is supported by the software results, the issue arises of how to deal with the offence. University regulations specifically forbid plagiarism and may prescribe the need to refer such behaviour to a disciplinary procedure. However, such disciplinary action does not always happen.[33] Sometimes plagiarism is treated by giving the student a mark of zero for the work submitted. Sometimes the student is required to resubmit a new version of the work. In other words, formal disciplinary action is not automatically invoked.[34] The process of university disciplinary action can be a long and costly procedure, and universities are often reluctant to follow this course on the grounds of the resources a formal plagiarism enquiry would require. Mathews[35] quotes an academic, who wished to remain anonymous, as saying, 'I'm ashamed to admit it but you simply don't have the time to launch a plagiarism case.'

If this is indeed the case, is the university not implicitly condoning plagiarism? In any event, plagiarism is frequently regarded more as a misdemeanour than a felony. Nonetheless, the practice of plagiarism undermines the integrity of the academic process, and calls into question the quality of education that is said to be evident in the holding of a degree.

It has been suggested that plagiarism is rife in certain cultures, and in developing countries in particular.[36] This may be true, but the challenge is certainly not absent from western culture or universities.

Ghostwriting

Ghostwriting is the practice of hiring a writer (or writers) to produce a piece of work that follows a predefined style, and none of the original writing credit is attributed to the ghostwriter. The practice of ghostwriting has long existed in the field of literature. Related practices occur in other forms of the arts, including music composition, singing, and the visual arts. In the university context, all subjects are to some extent vulnerable to plagiarism. But with regard to ghostwriting, computational sciences are particularly vulnerable. Specialist websites can produce programming code for computer science students.[37] Undergraduates have been known to pass on, or indeed sell, their laboratory journals to other students – which in some cases have been copied verbatim the following year.

Ghostwriting has traditionally been associated with famous individuals who have contracted someone else to produce a work in the field of literature, perhaps an autobiography, because they do not have the time or skills to complete the task themselves.[38] The ghostwriter produces the work for an agreed fee. This is a legal and sustainable business, and many ghostwriting agencies have a substantial history of success.[39]

In pre-Internet days, a student struggling to complete an essay or assignment might have asked a friend or family member (or another individual) to help write a piece of the work. Sometimes money changed hands, but not always. It would generally have been difficult to find someone to do this type of work and the incidence of such collusion would have been low. The Internet has changed things, and a large number of essay-writing services are now offered on the web.[40]

Some providers offer a full range of services, from writing a simple essay to producing a doctoral thesis. Furthermore, the purchaser is able to specify if he or she wants an essay to be written to a particular standard of excellence, such as being good enough to obtain a first class or a second class grade. Some of these services respond quickly, offering a 24-hour turnaround time for an essay. Of course, the fee asked for the

production of one of these essays depends on the standard required and the time-frames for delivery.[27] The academic networks these organisations purport to have established suggest that a large number of competent academics are willing to earn money from the process of defrauding, or at least undermining, the examination process.

When ghostwriting services first appeared on the Internet, they were relatively unsophisticated and a number of ghostwritten essays were caught by anti-plagiarism software. However, today the suppliers of these essays claim they can produce work that will not be detected by such software. Although there are programs that claim to be able to identify authors by their style using the principles of stylometry, a competently-produced piece of work by a ghostwriter would be original and thus would not be detected by software alone.

The range of organisations offering ghostwriting services is impressive. A recent Google search produced over 4.6 million references to these services in less than half a second. A number of these organisations claim that they employ graduates and faculty members from the best universities. It has been estimated that in the United Kingdom alone, more than GBP200 million is spent annually on these services.[41]

The classification of this type of academic misconduct is not always easy to define. Ghostwriting differs from plagiarism, although this point is not universally agreed on. For example, University College London (UCL) states that plagiarism includes 'turning in someone else's work as your own'.[42] Plagiarism is sometimes defined as theft, and the word 'plagiarism' comes from the Latin word for kidnapping. But ghostwriting is different; there is no direct theft involved. It is rather a question of misrepresentation or lying about the authorship of the work. However, this type of offence could be considered an extreme form of plagiarism at least. Some academics feel that ghostwriting is considerably more serious than plagiarism with regard to the degree of violation of academic trust.

Tomard[39] suggests that there are three distinct categories of students who employ ghostwriters. The first are students whose command of the English language is not sufficient to be able to write a competent research report. The second group includes students who have not been able to grasp the detail of the processes involved in academic research methodology, and therefore need an expert in the field of study to write up the research. The third group consists of students who are both uninterested in their studies and sufficiently well-funded to be able to afford the high fees asked by ghostwriting agencies. Being able to identify these groups should facilitate the university in creating policies to counter this type of academic misconduct.

When a dissertation is presented to a university, traditionally it has to be accompanied by a Certificate of Own Work.[43] This certificate is sometimes a simple one-sentence statement, signed by the student. If it transpires that the statement was false, that renders the work null and void and a degree would not be awarded. If a degree has already been awarded and it is subsequently found out that the student did not perform the work, the degree may be withdrawn. However, certification of authenticity has generally not been required for other work submitted, such as essays or term papers. Perhaps such certification should now become a routine requirement.

If a piece of academic work has been ghostwritten, this can generally be detected only if the evaluator is personally acquainted with the student's level of subject knowledge and his or her natural writing style.

Extent of cheating

As noted earlier, it is difficult to give an estimation of the extent to which cheating occurs. However, it is clear that the ghostwriting industry is highly active and appears to be expanding. This trend is certainly cause for concern.

Universities are often quite secretive about issues concerning discipline, or in fact any legal action in which they are involved. Documentary filmmakers have suggested that there are a growing number of students taking action against universities on a wide range of issues. Most of these cases have to do with universities not living up to promises they made to students before the students registered. However, in at least one case on record, a student sued a university after he was accused of plagiarism.[44] He argued that the university had not appropriately informed him that plagiarism would be regarded as academic misconduct, and therefore he could not be held responsible for an activity he did not realise was regarded as misconduct. It is hard to imagine how anyone could have come through schooling in the western world, including the early years of university, without having the issue of plagiarism fully explained to them.

The special case of dissertations

Formerly, dissertations and theses were an academic activity that was required only at the level of masters and doctoral degrees. However, this has changed, and dissertations are now often required at undergraduate level.[45,46] As a material piece of work that is researched and written almost wholly by the student, sometimes without much direct assistance by academic staff, the dissertation has become an important part of many degrees.

Students often struggle with their dissertations, which can be a significant challenge. Ranging in length from perhaps as few as 5000 words at undergraduate level to as many as 50 000 words at masters level and 80 000 words at doctorate level, the dissertation requires a material amount of focused worked over an extended period of time.

Sometimes at master's level, and even more so at doctoral level, many universities experience a high rate of non-completion of this type of work.[47-49] For a number of reasons, students are not able to complete all the work required for a dissertation. This happens more frequently among part-time students, especially at doctoral level. For this reason it has been suggested by some academics that a doctoral degree candidate could outsource some of the work required. However, this leads to some challenging questions and issues.

Attitudes towards outsourcing differ considerably from university to university. Some universities take the stance that virtually nothing should be outsourced, whereas others are far more relaxed. The reality is that students have long outsourced certain aspects of their dissertations. For many decades and probably for most of the 20th century, students have had their dissertations typed by others, normally professional typists. When word processors became commonplace and desktop typesetting became available, students had their work professionally produced without any questions being raised.

However, there are a number of other aspects of the research work that outsiders now offer to undertake for students, and which are not as acceptable as typing and typesetting. For example, it has been proposed that compiling the literature review, collecting the data, and analysing the results are all activities that could be outsourced. These are important doctoral-level activities, and as such are central to the intellectual development of the student. Having these tasks performed by anyone other than the student (i.e. the degree candidate) undermines the objective of acquiring the degree. It is difficult to see how this level of outsourcing could be acceptable to the academic world.

Another aspect of dissertations also presents a major problem, although it might not be regarded as so obviously problematic as the issues described above. That is the actual writing of the dissertation. Frequently students are not accomplished academic writers, and the way they attempt to present the arguments behind their research can be difficult to understand. Traditionally a supervisor would give the student guidance with regard to academic writing style, and in some cases supervisors would actually edit the text of the dissertation. However, some universities explicitly forbid supervisors to do this, and the result has been that students now tend to hire freelance editors. In some cases, the amount of work undertaken by these editors has amounted to having the dissertation ghostwritten.

At the same time, as mentioned earlier, the ghostwriting industry now offers the writing of an entire dissertation as part of its product range. Some of the claims are startling, as they suggest an entire dissertation can be produced within a matter of weeks – if not days. It is hard to imagine what calibre of university would accept a student who presents a dissertation, even at undergraduate level, without having had a number

of consultations with his or her supervisor. In the case of masters or doctoral degrees, universities require a number of years of supervised research before a final dissertation can be submitted or considered. Nonetheless, the fact that dissertations are being offered for sale by ghostwriting enterprises strongly suggests that there are indeed universities that accept this type of written work and award degrees on the basis of ghostwritten dissertations.

The ethics of the ghostwriter

It does not appear that ghostwriters have much, if any, ethical concern about the work they do. Ghostwriting agencies boast that they hire only writers with the highest qualifications from the best universities, and judging from the apparent satisfaction of their clients this does seem to be the case. One ghostwriter who decided to write anonymously for *The Times Higher Education*[50] stated:

> *I don't justify the work I'm doing on ethical grounds. While what I do is not illegal, it does enable others to break rules and suffer the consequences if they are caught. The agencies maintain the image of legitimate businesses: many do not even refer to 'cheating'. You are simply 'helping' with an assignment (making up, as one agency argues, for the university's failure to provide adequate tuition). While I'm happy to acknowledge that I am dependent on clients' continued cheating, this doesn't mean I am not conscious that my job is a symptom of an illness, a fracture, in our universities.*

It is challenging not to sympathise with the argument that the university system is 'fractured'. However, the argument used by many essay mills – that they are only 'helping' students – is at best disingenuous.[36] This is a for-profit industry that operates within the law but exhibits little concern about the morality of its activity.[51]

Interestingly, the same ghostwriter quoted above[50] also commented that:

> *I stay away from applied fields – it is my only ethical standard as a ghost writer. I will not help a nurse to qualify on false pretences: who knows, it might be my parents who find themselves in their care.*

This is an interesting admission of the impropriety involved in the act of ghostwriting.

The reaction of the universities

The issue of plagiarism and ghostwriting is of critical importance to universities for at least two reasons. The first is that these types of misconduct discredit the degrees that are awarded. If plagiarism and ghostwriting are perceived as being rife at a particular university, this is a disincentive to anyone who desires a robust qualification to attend that institution. The second reason is that it is unfair for a student to obtain credit for work he or she did not actually do. Having the money to buy completed academic work does not enhance the intellectual capability of the student – as the holding of a degree is supposed to indicate. Furthermore, the consequence of such misconduct is that honest students are placed at a disadvantage.

With regard to plagiarism, in general universities have reacted rather slowly and with some trepidation. A few years ago, when plagiarism was suspected in a master's degree dissertation by an examiner and the use of anti-plagiarism software was suggested, the student's supervisor might have exclaimed, 'Are you impugning the integrity of my student?' Fortunately, those days are past and most universities now require dissertations to be submitted both electronically and in hard copy (i.e. on paper). Nonetheless, the question of what action to take against people who are found to be plagiarists has not been answered.

Ghostwriting is, in a sense, a more difficult issue than plagiarism. The outright purchasing of essays, term papers, and dissertations is clearly an act of fraud. It would be unwise for a university to do anything less than take the most severe disciplinary action. But in some cases, obtaining help from an informal mentor could come quite close to ghostwriting, and penalising this activity could present difficulties. Perhaps the real issue is that universities have been, and still are, focused on catching misconduct after it has occurred instead of preventing it. Is there really any way of preventing these types of cheating? The answer is an unreserved 'yes'. How this can be done is addressed below, under the subheading 'Prevention and detection'.

The university as a fractured institution

In discussing how universities have become fractured, it is necessary to bear in mind the exceptional pressures these institutions have had to face in recent years. Since the mid-1990s many universities have been expected to deliver what is sometimes referred to as 'mass education'.[52,53] This term is not well defined but its general meaning is clear: education has to be made available to large numbers of students.

At the same time, the resources made available to universities have not correspondingly increased. Class sizes have increased, with a concomitant unfavourable shift in the ratio of students to lecturers. In some cases, lecturers no longer grade the work of their students and this task is sub-contracted to either teaching assistants or even to outside contractors. In addition, a greater number of students from other countries, who may have inadequate command of English, are being admitted to degree courses at English-medium universities. These factors have made the relationship between faculty and students more challenging, as illustrated by the following remarkable statement made by the anonymous ghostwriter[50]:

> *I operate on the assumption that the student I'm working for will have little or no personal interaction with academic staff. This means there is only a small likelihood that the lecturer who sets and marks the questions will be familiar with the student's style of writing.*

If this is a correct assessment of the situation, and there is *prima facie* evidence to suggest that it is, then the universities for which this is true are not performing their expected function.

The American publication *The Chronicle of Higher Education* interviewed a ghostwriter who remarked as follows[54]:

> *You've never heard of me, but there's a good chance that you've read some of my work. I'm a hired gun, a doctor of everything, an academic mercenary. My customers are your students. I promise you that. Somebody in your classroom uses a service that you can't detect, that you can't defend against, that you may not even know exists.*

Matthews similarly reflected on ghostwriters' services as follows[35]:

> *There are also concerns that in an age of mass higher education and high student-to-staff ratios, lecturers are less able to get to know their students' work, making this form of cheating more difficult to detect. And there are fears that the pressures of the job might encourage some academics to turn a blind eye to the practice. But perhaps the most important question is whether it is possible to prevent this form of cheating in the first place.*

It is clear that universities are not addressing this subject with the energy or commitment one might expect when such important issues are at stake.

The function of universities

The purpose of a university is not only to communicate and test students' knowledge but also to inspire them to become lifelong learners.[55,56] This means there should be an onus on the university to help students realise that engaging with the subject matter is interesting, enjoyable and rewarding. If these positive experiences are achieved, learning should occur naturally and students should become well-informed in their fields of study. If this were the case, there would be little (or less) motivation for anyone to cheat.

Although some academics would argue that they already take this approach to teaching, in reality most lecturers present fairly routine material, and students are expected to learn and to reproduce it in rather unimaginative ways – either during examinations or in essays, term papers, and even dissertations. This means the evaluation processes in universities are often a test of one's memory of material that has been offered for the purposes of learning in a relatively structured fashion. When this approach is combined with large numbers of students, the temptation to engage in plagiarism or use ghostwriters appears to offer a solution to some students.

Prevention and detection

Some universities have taken the view that plagiarism and ghostwriting can be prevented by adequate detection methods and the imposition of appropriate penalties. For this reason, anti-plagiarism software has become a large business sector. The penalties imposed by universities range from requiring the student to resubmit the piece of work to suspending or even expelling him or her from the university. However, universities do not easily impose suspensions and expulsions. Students sometimes face only a rebuke and have to resubmit their work. Countering plagiarism in this half-hearted manner may, however, produce an 'arms-race' mentality, so that those who facilitate students' cheating will try to create increasingly clever ways of avoiding anti-plagiarism software. A better approach is to reassess the university's attitude to teaching and learning.

The New Zealand Government has produced an interesting set of guidelines for the effective prevention and detection of academic fraud.[57] The approach is intended to create awareness of the potential problems of academic fraud, and to continually remind students about how unacceptable the practice is. It is then necessary for staff to engage continually with their students and to be on the lookout for any surprising changes in their performance. This means getting to know the students well. All of these suggestions are welcome, but they are time-consuming to put into practice. In general, academics regard themselves as having a full workload without taking on any extra engagements or responsibilities.

Implementing the types of policies suggested in the New Zealand guidelines represents a significant move away from present general practice. The cost of such a transformation would be regarded by many to be a heavy burden on the financial resources of educational institutions. In addition, not all academics would necessarily welcome such a change. There is little doubt that many academics are largely comfortable with the present system.

With regard to ensuring the integrity of a dissertation, the issue is one of adequate supervision. If a programme of careful supervision is in place, there should be relatively little opportunity for plagiarism to escape unnoticed in the writing of a dissertation. To a large extent, a dissertation should be almost co-created by a student and his or her supervisor, with the student doing the work and the supervisor keeping a close eye on what is happening step-by-step.

With regard to ghostwriting, there should be almost no opportunity for a student to pass off a piece of work produced in this way. If the supervisor does not know the student well enough to be able to immediately detect that a written submission is inauthentic, the supervisor is not actually doing the job adequately. Unfortunately, some universities do not allocate sufficient time to supervisors for them to be able to get to know their students well enough; consequently, they may indeed be unable to detect if the work submitted has been written by someone else.

Summary and conclusion

This paper has reviewed the current situation with regard to plagiarism and ghostwriting at university level. The objective of the discussion was to stimulate debate as to how universities should react to these types of academic misconduct. (The paper has deliberately avoided addressing how these problems may be exaggerated in distance learning or e-Learning university programmes.)

The number of students who plagiarise or use ghostwriters appears to be on the increase. Although as a percentage of the entire student population, the number of those who engage in academic misconduct is believed to be small, even that small percentage represents a large number of students in absolute terms. This is cause for concern. If cheaters manage to 'beat the system' and obtain a degree they have not earned through their own academic performance, these fraudsters would represent a significant threat to the integrity of the relevant department, faculty, and university – indeed, to the whole notion of higher education.

Plagiarism and ghostwriting must be eliminated to the fullest extent possible, because these practices are fundamentally unfair to honest students who rely on their own intellectual abilities to create the academic work required of them.

In general, plagiarism is a substantial and unwelcome misconduct, but it tends to be relatively easily identifiable. However, as it is regarded a form of academic fraud, there should be a material penalty paid by those who are found to engage in this behaviour. Although universities are in a position to impose such penalties, they do not appear to do so adequately and effectively, or often enough.

The use of a ghostwriter is an offence for which a greater penalty should be paid. Although ghostwriting is not illegal, lying about the authorship of a piece of work is potentially fraudulent. Ghostwriting cannot be easily detected by software in the same way as plagiarism can. There are some products that employ stylometry, and these may be of help and could be utilised by universities. However, ghostwriting is best detected by lecturers having personal knowledge of the capabilities of their students. Introducing a greater number of oral examinations could quickly and easily eliminate the entire issue of ghostwriting. Unfortunately, it would also create a substantial workload, which the current university examination system is ill-equipped to deal with.

Nonetheless, the overall approach of identifying academic misconduct and imposing penalties is unlikely in itself to solve the problem. What is really required is a new attitude to prevention. This would involve creating learning environments in universities that would invite students to become highly engaged with their subject material, and to express their creativity in such a way that it would be apparent who has succeeded in the learning process and who has not. Students who have succeeded should be valued and rewarded with an appropriate degree, whereas those who do not master the process will be deemed to have failed.

With regard to misconduct at the dissertation level, it should not be a major problem for supervisors who are sufficiently engaged with their students to be able to detect misconduct. Most supervisors have access to anti-plagiarism software, and it is a simple matter to pass submitted work through such an analysis. On the question of ghostwriting, although a supervisor should expect the writing style of a student to improve significantly during the course of producing a dissertation, an extraordinary level of improvement should become the subject of an enquiry and maybe even an investigation. There needs to be a severe penalty for those who engage in this type of misconduct; the treatment of serious academic misconduct must be sufficiently firm to tackle the problem effectively.

Finally, it is also important for the issue of plagiarism and ghostwriting to be discussed more openly and regularly within universities. There might even be certain proven or known incidents where the transgressors should be named and shamed.

Authors' contributions

Both authors contributed equally to the manuscript.

Acknowledgement

The authors would like to acknowledge the assistance of Dr Filistea Naude from the Unisa Science Campus Library.

References

1. Retraction Watch [homepage on the Internet]. No date [cited 2015 Jul 28]. Available from: http://retractionwatch.com/

2. Bhattacharjee Y. The mind of a con man. New York Times. 2013 April 26 [cited 2015 July 28]. Available from: http://www.nytimes.com/2013/04/28/magazine/diederik-stapels-audacious-academic-fraud.html?pagewanted=all&_r=1

3. Honig B, Bedi A. The fox in the hen house: A critical examination of plagiarism among members of the academy of management. Acad Manag Learn Edu. 2012;11(1):101–123. http://dx.doi.org/10.5465/amle.2010.0084

4. Lewin T. Dean at M.I.T. resigns, ending a 28-year lie. New York Times. 2007 April 27 [cited 2015 July 28]. Available from: http://www.nytimes.com/2007/04/27/us/27mit.html

5. Farrell EF. California poet laureate admits to lie on résumé. Chronicle of Higher Education. 2002;49(12).

6. Flynn NS. UCD inquiry to study vetting practices. The Irish Times. 2001 November 13 [cited 2015 Aug 04]. Available from: http://www.irishtimes.com/news/ucd-inquiry-to-study-vetting-practices-1.336913

7. Guha A. Vice-Chancellors are not angels. Frontier. 2011;43(45); May 22–28 [cited 2016 May 17]. Available from: http://frontierweekly.com/archive/vol-number/vol/vol-43-2010-11/vol-43-45/vicechanellors-43-45.pdf

8. Monash head leaves over plagiarism charge. The Guardian. 2002 July 18 [cited 2015 Jul 28]. Available from: http://www.theguardian.com/education/2002/jul/18/highereducation.uk

9. Rodan P. A modest victory for academic values: The demise of David Robinson. Australian Universities' Review. 2004;46(2):16–22.

10. Columbia University. The cloning scandal of Hwang Woo-Suk [article on the Internet]. c2015 [cited 2015 Jul 28]. Available from: http://stemcellbioethics.wikischolars.columbia.edu/The+Cloning+Scandal+of+Hwang+Woo-Suk

11. McNeill D. In Japan, research scandal prompts questions. The Chronicle of Higher Education. 2014 June 30 [cited 2015 Jul 28]. Available from: http://chronicle.com/article/In-Japan-Research-Scandal/147417/?cid=at&utm_source=at&utm_medium=en

12. Office of Research Integrity. Case summaries [article on the Internet]. c2015 [cited 2015 Jul 28]. Available from: https://ori.hhs.gov/case_summary

13. Papp R, Wertz MH. Academic integrity violations: Isolated problems or epidemic? In: SAIS 2005 Proceedings; 2005 February 25–26; Savannah, GA, USA. Available from: http://aisel.aisnet.org/sais2005/30/?utm_source=aisel.aisnet.org%2Fsais2005%2F30&utm_medium=PDF&utm_campaign=PDFCoverPages

14. Sheard J, Dick M, Markham S, Macdonald I, Walsh M. Cheating and plagiarism: Perceptions and practices of first year IT students. ACM SIGCSE Bulletin. 2002;34(3):183–187. http://dx.doi.org/10.1145/637610.544468

15. Thomas A, De Bruin GP. Plagiarism in South African management journals. S Afr J Sci. 2015;111(1/2), Art. #2014-0017, 3 pages. http://dx.doi.org/10.17159/sajs.2015/20140017

16. Goveneder P. Universities battle a rising tide of cheating. Time Live. 2014 August 17. Available from: http://www.timeslive.co.za/local/2014/08/17/universities-battle-a-rising-tide-of-cheating1

17. The good, the bad, and the ugly. S Afr J Sci. 2003;99(9/10):402–403.

18. LeBlanc S. Harvard to honor Ted Kennedy, once expelled for cheating. CNSNews.com. 2008 December 01. Available from: http://cnsnews.com/news/article/harvard-honor-ted-kennedy-once-expelled-cheating

19. Wallace MJ, Newton PM. Turnaround time and market capacity in contract cheating. Educ Stud. 2014;40(2):233–236. http://dx.doi.org/10.1080/03055698.2014.889597

20. George JF, Davison RM, Heales J, Munro M. Report of the ad hoc committee on member misconduct to the AIS Council. Communications of the Association for Information Systems. 2003;11(1):2.

21. Reston J. Galileo: A life. Washington: Beard Books; 2000.

22. Stavisky LP. Term paper 'mills', academic plagiarism, and state regulation. Pol Sci Quart. 1973:445–461. http://dx.doi.org/10.2307/2148993

23. Clarke R. Plagiarism by academics: More complex than it seems. J Assoc Inf Systems. 2006;7(2):5.

24. Randall M. Pragmatic plagiarism: Authorship, profit, and power. Toronto: University of Toronto Press; 2001.

25. PennState. Why plagiarism is wrong [article on the Internet]. c2015 [cited 2015 Sep 22]. Available from: http://tlt.psu.edu/plagiarism/student-tutorial/why-plagiarism-is-wrong/

26. University of Oxford. Plagiarism [homepage on the Internet]. No date [cited 2015 Sep 22]. Available from: http://www.ox.ac.uk/students/academic/guidance/skills/plagiarism

27. Singh S. The zombie doctorate. In: Proceedings of the 14th European Conference on Research Methodology for Business and Management Studies; 2015 June 11–12; Valletta, Malta. Reading, UK: Academic Conferences and Publishing International Ltd; 2015.

28. Hallak J, Poisson M. Corrupt schools, corrupt universities: What can be done? Paris: UNESCO International Institute for Education Planning; 2007.

29. Richards D, Busch P, Germanou D, editors. Elearning promoting plagiarism or honesty? In: PACIS 2011 Proceedings; 2011 July 7–11; Brisbane, Australia. Available from: http://aisel.aisnet.org/pacis2011/

30. Adams R. Cheating found to be rife in British schools and universities. The Guardian. 2015 June 15. Available from: http://www.theguardian.com/education/2015/jun/15/cheating-rife-in-uk-education-system-dispatches-investigation-shows

31. Roberts J, Wasieleski D. The contemporary plagiarist: The roles of technology and moral development. In: ICIS 2011 Proceedings; 2011 Dec 4–7; Shanghai, China. Available from: http://aisel.aisnet.org/icis2011/proceedings/IScurriculum/9

32. Warn J. Plagiarism software: No magic bullet! High Educ Res Dev. 2006;25(02):195–208. http://dx.doi.org/10.1080/07294360600610438

33. Thomas A, De Bruin GP. Student academic dishonesty: What do academics think and do, and what are the barriers to action? Afr J Bus Ethics. 2012;6(1):13–24. http://dx.doi.org/10.4103/1817-7417.104698

34. Fuzile B. Plagiarism cases rock Fort Hare. Times Live. 2013 April 09. Available from: http://www.timeslive.co.za/thetimes/2013/04/09/plagiarism-cases-rock-fort-hare

35. Matthews D. Essay mills: University course work to order. Times Higher Education. 2013 October 10. Available from: https://www.timeshighereducation.co.uk/features/essay-mills-university-course-work-to-order/2007934.article?page=0%2C1

36. Matthews D. East Asia has 'toxic academic culture'. Times Higher Education. 2015 September 07. Available from: https://www.timeshighereducation.com/news/east-asia-has-%E2%80%98toxic-academic-culture%E2%80%99

37. BESTGhostWriters. Programming ghostwriter [homepage on the Internet]. No date [cited 2015 Sep 22]. Available from: http://www.bestghostwriters.net/programming-ghostwriter/

38. PLoS. Ghostwriting collection [homepage on the Internet]. No date [cited 2015 Jul 28]. Available from: http://collections.plos.org/ghostwriting

39. Tomar DA. Detecting and deterring ghostwritten papers: A guide to best practices [article on the Internet]. c2015 [cited 2015 Aug 04]. Available from: http://www.thebestschools.org/resources/detecting-deterring-ghostwritten-papers-best-practices/

40. Auger GA. Missing citations, bulking biographies, and unethical collaboration: Types of cheating among public relations majors. Journal Mass Comm Educat. 2013;68(2):150–165. http://dx.doi.org/10.1177/1077695813476953

41. Taylor M, Butt R. How do you make £1.6m a year and drive a Ferrari A: Sell essays for £400. The Guardian. 2006 July 29. Available from: http://www.theguardian.com/uk/2006/jul/29/highereducation.education

42. University College London. Plagiarism [homepage on the Internet]. No date [updated 2013 Sep 13; cited 2016 May 8]. Available from: http://www.ucl.ac.uk/current-students/guidelines/plagiarism

43. Remenyi D. Writing up your research. 2nd ed. Reading, UK: Academic Conferences and Publishing International Ltd; 2013. p. 127.

44. 'Plagiarist' to sue university. BBC News. 2004 May 27. Available from: http://news.bbc.co.uk/1/hi/education/3753065.stm

45. Collis J, Hussey R. Business research: A practical guide for undergraduate and postgraduate students. 4th ed. Basingstoke: Palgrave Macmillan; 2013.

46. Cook MC. The role of the academic supervisor for undergraduate dissertations in science and science-related subjects. Stud High Educ. 1980;5(2):173–175. http://dx.doi.org/10.1080/03075078012331377206

47. Herman C. Obstacles to success – doctoral student attrition in South Africa. Perspect Educ. 2011;29(3):40–52.

48. Mouton J. Doctoral production in South Africa: Statistics, challenges and responses. Perspect Educ. 2011;29(3):13–29.

49. Smith RL, Maroney K, Nelson KW, Abel AL, Abel HS. Doctoral programs: Changing high rates of attrition. J Humanist Educ Dev. 2006;45(1):17. http://dx.doi.org/10.1002/j.2161-1939.2006.tb00002.x

50. Why I write for an essay mill. Times Higher Education. 2013 Aug 01. Available from: https://www.timeshighereducation.co.uk/comment/opinion/why-i-write-for-an-essay-mill/2006074.article

51. Bosch X, Ross JS. Ghostwriting: Research misconduct, plagiarism, or fool's gold? Am J Med. 2012;125(4):324–326. http://dx.doi.org/10.1016/j.amjmed.2011.07.015

52. Scott P. The meanings of mass higher education: Buckingham: Open University Press; 1995.

53. Houle CO. The design of education. 2nd ed. San Francisco, CA: Jossey-Bass Higher & Adult Education Series; 1996.

54. Dante E [pseudonym]. The shadow scholar: The man who writes your students' papers tells his story. The Chronicle of Higher Education. 2010 November 12. Available from: http://chronicle.com/article/The-Shadow-Scholar/125329/

55. Goodchild LF, Wechsler HS. The history of higher education. ASHE Reader Series. Boston, MA: Pearson Custom Publishing; 1997.

56. Newman JH, Turner FM. The idea of a university. New Haven, CT: Yale University Press; 1996.

57. New Zealand Qualifications Authority. Effective practice in preventing and detecting academic fraud [document on the Internet]. c2013 [cited 2015 Aug 04]. Available from: http://www.nzqa.govt.nz/assets/Providers-and-partners/NZQA-Effective-practice-guide.pdf

DNA-based identification of aquatic invertebrates – useful in the South African context?

AUTHORS:
Hermoine J. Venter[1]
Cornelius C. Bezuidenhout[1]

AFFILIATION:
[1]Unit for Environmental Sciences and Management, School of Biological Sciences, North-West University, Potchefstroom, South Africa

CORRESPONDENCE TO:
Hermoine Venter

EMAIL:
hermoine.venter@outlook.com

POSTAL ADDRESS:
Unit for Environmental Sciences and Management, School of Biological Sciences, North-West University, Private Bag X6001, Potchefstroom 2520, South Africa

DATES:
Received: 27 Nov. 2015
Revised: 15 Jan. 2016
Accepted: 18 Jan. 2016

KEYWORDS:
aquatic macroinvertebrates; DNA barcoding; eDNA; metabarcoding; biomonitoring

HOW TO CITE:
Venter HJ, Bezuidenhout CC. DNA-based identification of aquatic invertebrates – useful in the South African context? S Afr J Sci. 2016;112(5/6), Art. #2015-0444, 4 pages. http://dx.doi.org/10.17159/sajs.2016/20150444

The concept of using specific regions of DNA to identify organisms – processes such as DNA barcoding – is not new to South African biologists. The African Centre for DNA Barcoding reports that 12 548 plant species and 1493 animal species had been barcoded in South Africa by July 2013, while the Barcode of Life Database (BOLD) contains 62 926 records for South Africa, 11 392 of which had species names (representing 4541 species). In light of this, it is surprising that aquatic macroinvertebrates of South Africa have not received much attention as potential barcoding projects thus far – barcoding of aquatic species has tended to focus on invasive species and fishes. Perusal of the BOLD records for South Africa indicates a noticeable absence of aquatic macroinvertebrates, including families used for biomonitoring strategies such as the South African Scoring System. Meanwhile, the approach of collecting specimens and isolating their DNA individually in order to identify them (as in the case of DNA barcoding), has been shifting towards making use of the DNA which organisms naturally shed into their environments (eDNA). Coupling environmental and bulk sample DNA with high-throughput sequencing technology has given rise to metabarcoding, which has the potential to characterise the whole community of organisms present in an environment. Harnessing barcoding and metabarcoding approaches with environmental DNA (eDNA) potentially offers a non-invasive means of measuring the biodiversity in an environment and has great potential for biomonitoring. Aquatic ecosystems are well suited to these approaches – but could they be useful in a South African context?

The rise of DNA-based identification for biomonitoring

Several authors have pointed out that conservation of natural resources and ecosystems hinges on the provision of data regarding the presence and distribution of species within an environment[1-3] – data which usually are supplied through biomonitoring initiatives. This principle is at the core of monitoring programmes, including the South African Scoring System or SASS (one of the programmes that forms part of the current National Aquatic Ecosystem Health Monitoring Programme: River Health Programme[4]), which analyses macroinvertebrate communities as a measure of stream ecosystem health. SASS, like similar indices, uses morphology to make identifications, which are then assigned weight and values which are then used to gauge the relative well-being of the system.[4]

The limitations of morphology-based indices for the identification of macroinvertebrates are discussed at length in other publications,[5,6] so will only be mentioned briefly here. DNA-based identification methods provide an alternative to morphological identifications, and have been useful in addressing several of these problems. For example, in morphology-based identification of macroinvertebrates it often is difficult to differentiate between cryptic (morphologically indistinguishable) species, several of which have come to light after DNA-based approaches were used.[7,8] Larval stages of aquatic insects are often extremely difficult to identify morphologically, even for experts. In addition, morphological identification keys for aquatic macroinvertebrates tend to focus on adult stages, compounding the difficulties of juvenile identification. Furthermore, linking of life stages (especially juveniles) and female and phenotypically variant individuals to identified representative/voucher specimens is incomplete for many macroinvertebrate species. DNA-based methods have proven useful in resolving such issues, because they rely on genetic loci which are applicable regardless of sex, life stage or appearence.[6,9,10]

Morphological identification is time-consuming and requires a great deal of taxonomic knowledge and skill to be successful. However, once the initial effort has been made to identify a species using DNA, the expertise needed to identify subsequent specimens of the same species is drastically reduced.[5] Another advantage of using DNA-based approaches is the potential for identification of specimens to species level, in contrast to many morphological indices which often stop at family level. Identification to species level can bring a sensitivity and depth of knowledge to biomonitoring that coarser identifications are not able to.[3]

The standardisation of DNA-based identification methods gave rise to the Consortium for the Barcode of Life. This initiative was set up with the goal of promoting the use of specific regions of DNA (in the COI gene for animals, *rbcl* and *matK* for plants) to determine the sequence of those regions which was particular to each species, for use in identification. The Barcode of Life Database (BOLD) is a centralised database in which such barcode sequences, as well as specimen collection and species distribution details, are available to any interested person. The database contains over 3.7 million entries from all over the world – including 62 926 from South Africa, 11 392 of which have species names (representing 4541 species).[11]

However, despite the optimism and success stories associated with this approach, there are a number of limitations attached to using the COI region as a marker, not least of which is that this method relies on a DNA sequence which is about 650 base pairs long. This effectively eliminates the possibility of using this approach on a large body of damaged DNA fragments as they are too short – including many museum voucher specimens (potentially a source of reference material[12]), as well as most environmental DNA (eDNA – see Box 1), which is typically damaged. And

so the method has continued to evolve, with the advent of the 'mini-barcode'[13], as well as several suggestions for alternative DNA regions to be used.

Box 1: Definition of terms

eDNA (environmental DNA)	Although some authors define eDNA simply as DNA obtained directly from environmental samples (soil, sediment, water, etc.) as opposed to individual specimens, a more specific definition would include the corollary that there should be no obvious signs of biological source material.[2]
Metabarcoding	The use of DNA markers (metabarcodes) to perform analyses on total DNA isolated from a bulk/environmental sample, in order to characterise the assemblage of organisms present in that environment.[14,15]
BINS (barcode index numbers)	Operational taxonomic units of genetically identical taxa indexed under a common identifier.[16]

Being able to work with eDNA precipitated from samples from aquatic ecosystems opens up a range of monitoring and research opportunities.[17,18] Thus other primers and markers have been developed for eDNA, which are often species-specific (in contrast to the broad range of COI), and which have been applied to the detection of indicator species as well as rare, invasive or pathogenic species.[17] Mächler et al.[18] demonstrated the potential of using eDNA and specific primers to detect macroinvertebrate species in both river and lake systems. Using standard polymerase chain reaction or PCR (which is cheaper than next generation sequencing), they were able to detect both indicator and non-native species using their own primer design. Specially designed primers and probes used in conjunction with qPCR were successfully used to survey the population of European weather loach in Denmark. During this study, this near-extinct fish was detected at sites where its presence had been observed recently, as well as one location where it had not been observed since 1995. In addition to successful detection, the authors report that this approach is less costly, both economically and in terms of effort (person-hours).[19]

DNA metabarcoding is an approach which uses bulk DNA collections (such as faeces[20], sediment meiofaunal communities[21] and eDNA[2]) coupled with next generation sequencing to obtain an overview of the organisms which are present in an environment as a whole. Instead of targeting one or a few species, this approach aims to give a more holistic view of ecosystem composition.[4,22] The reads which result from such high throughput sequencing are clustered into operational taxonomic units (OTUs). Depending on the genetic loci used, OTUs can be used to match to sequences in databases such as BOLD (if COI was used) or GenBank, in order to identify the organisms whose DNA was in the sample.[14,22,23] In other cases, in which primer sets other than COI are used or when no reference sequences are available for the OTUs produced, taxonomic assignation is not possible. However, that does not mean the data are not useful. Molecular taxonomic units (MOTUs) refer to representative sequence clusters which have been grouped together using particular algorithms. MOTU data can be utilised in lieu of 'true' species data, by comparing MOTU profiles of different environments or time periods.[24] Although the metabarcoding approach (see Box 1) still is being refined, it potentially allows monitoring of community-level responses to change, including responses to remediation strategies and climate change.[14,15]

eDNA in aquatic ecosystems

The fact that DNA has a relatively short turnover time in aquatic systems means that aqueous eDNA is likely to represent a 'real-time' view of species present within a relatively small window of time.[25] Strickler et al.[26] investigated the effects of temperature, pH and UVB radiation on eDNA in water and found that it degraded faster in warmer water, with a neutral pH and a moderate UVB level. Because these conditions are also amenable to microbial growth, the authors speculated that eDNA breakdown was at least partially facilitated by microbial action. During investigations into DNA persistence in both laboratory and field conditions (ponds), species could be detected using eDNA for 25 days and 21 days after removal of the organisms, respectively.[27] However, the fact that DNA may be concentrated and survive much longer in sediments may be a complicating factor.[25]

DNA dispersal in flowing streams and rivers is also a concern, as it may give false positive results downstream where the organism in question is not found. DNA dispersal was investigated by Laramie et al.[28] who traced eDNA of Chinook salmon (*Oncorhynchus tshawytscha*), and Deiner et al.[29] who studied eDNA of a daphnid (*Daphnia longispina*) and swollen river mussel (*Unio tumidus*). Both studies found that eDNA signals tend to decrease as distance from the source increases.[28,29] In Deiner et al.'s study, DNA from lake-dwelling invertebrates was detected 12 km downstream from the lake inhabited by the target organisms. The authors suggest that when using eDNA to estimate biodiversity in such ecosystems, sample sites should be 5–10 km apart, and follow the stream hierarchy.[29]

Other challenges

Just as a person cannot be identified by their fingerprints unless a record of their fingerprints exists as a reference, if the barcode databases do not contain a matching record of barcoded specimens of a species with which to compare a query sequence, DNA cannot be used to make species level identifications (although specimens may still be placed within families or genera). We encountered this problem when attempts were made to use macroinvertebrate COI sequences to provide further resolution to morphological identifications. Only weak matches could be found with sequences in the GenBank database. These matches often did not agree with the morphological identifications, or corresponded to species found in countries in the northern hemisphere or in Australia. In addition, as mentioned earlier, BOLD Systems Database records for South Africa indicate a noticeable dearth of aquatic macroinvertebrates (with greater focus on fish and invasive species[30]). For example, there are zero entries referring to families such as Baetidae and Ephemeridae, and only 43 records for the order Ephemeroptera. Similarly, Plecoptera had only 13 entries, while Odonata had 66 and Trichoptera 138. When compared to the 3621 records for Coleoptera and 3150 for Hemiptera,[16,31] it is clear that there is considerable room for improvement for aquatic organisms.

This challenge can be overcome by building up sequence libraries. A possible starting point may be natural history museums. The addition of sequence data to curated specimen records could be invaluable. The Fresh Water Invertebrates collection of the Albany Museum was reported to contain 67 000 specimens in 2009.[32] Although it would be a boon if this collection could act as a starting point for aquatic invertebrate barcoding initiatives, the storage conditions and age of museum specimens tend to lead to DNA degradation, and have been known to impede barcoding efforts.[12] Thus, although some success has been achieved using such specimens, it is not possible to escape collection and identification of fresh specimens entirely.

Because better quality DNA may be obtained from fresh specimens, or those which have been stored correctly (in 95% ethanol and then at -10 °C)[12], this practice should be encouraged among those who sample and collect aquatic invertebrates. Although they may not be familiar with the techniques necessary to isolate and process the DNA for barcoding, if the experts who are able to identify aquatic invertebrates do so and then store the specimens correctly, the molecular work can be done at a later stage. Alternatively, specimens can be barcoded first and clustered into barcode index numbers or BINs[16] (see Box 1) according to barcode similarity. Representative specimens from such BINs can be selected for morphological identification and description, especially if potential cryptic species come to light. Building up a library of aquatic invertebrate DNA may thus lead to interdisciplinary cooperation and collaboration. Establishing an identifying sequence for a species is not

the limit of information which can be gained from DNA. Additionally, if other DNA regions need to be selected in future applications (such as metabarcoding and genome skimming), then the DNA which has already been isolated can be used to characterise the organism from a different perspective.[33,34]

A challenge which is less of a problem during traditional barcoding – in which organisms are identified one specimen at a time – but is an obstacle for bulk samples, is PCR bias. During the initial enrichment steps, PCR bias can create a number of problems for metabarcoding and eDNA analyses. For example, COI primers used in DNA barcoding have been found not to operate with the same efficiency for all organisms' DNA. In a mixed sample, bias towards certain organisms may cause their presence to be overstated, while others are underrepresented or missed entirely.[35] To overcome this bias, Taberlet et al.[33] suggest that metabarcodes and primers be tailored to the needs of each project. In order to do this, they propose that the DNA and barcodes collected and placed in curated collections during barcoding efforts could be used to develop this technique further. Although there are advantages to designing metabarcodes from within the COI barcoding region – such as access to the vast amounts of already identified sequences – these advantages must be weighed against potential biases.[14,22,36]

Competence in bioinformatics and molecular biology techniques, particularly those involving high-throughput sequencing, will have to be developed and encouraged in order to take full advantage of the huge data sets which such techniques generate. So too, expertise in traditional taxonomy and morphological identification – far from rendering such skills obsolete, as some fear[37], these initiatives cannot be accomplished without robust morphological identifications. Utilising the three approaches of taxonomy, barcoding and metabarcoding in tandem will allow researchers to link nearly three centuries of taxonomic research to modern data sets and community structures.

Conclusion

By understanding the biodiversity of South Africa better, we may be better able to protect it. By learning more about biota and their interaction with the environment, predictions can be made regarding how ecosystems will respond to change, and what can be done to preserve them. The recent report by Dallas and Rivers-Moores[1] both highlights the possible changes which may be wrought by climate change, and calls for more proactive monitoring. Clearly, barcoding and metabarcoding could be advantageous for biologists working with aquatic macroinvertebrates and aquatic ecosystem monitoring, particularly for those who do not have a background in taxonomy. However, in order to harness the usefulness of these techniques, an effort has to be made to collect the necessary data. For this reason, we advocate the establishment of regional collections which link identified aquatic species with their DNA sequences, which can be used to develop primer sets and standard methods for the use of eDNA in biomonitoring. Furthermore, we recommend that the establishment of collections be done in conjunction with a SASS approach, so that DNA-based approaches can be made relatable to previous work.

Acknowledgements

We gratefully acknowledge Biosafety South Africa as well the National Research Foundation of South Africa for funding. Views expressed are those of the authors and not of the funding entities. We also thank the anonymous reviewers for their helpful comments and suggestions which improved the manuscript, particularly one reviewer for practical suggestions regarding the binning of specimens.

Authors' contributions

C.C.B. supervised and made conceptual contributions; H.J.V. wrote the manuscript and made conceptual contributions.

References

1. Dallas HF, Rivers-Moore N. Ecological consequences of global climate change for freshwater ecosystems in South Africa. S Afr J Sci. 2014;110(5–6), Art. #2013-0274, 11 pages. http://dx.doi.org/10.1590/sajs.2014/20130274
2. Thomsen PF, Willerslev E. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biol Conserv. 2015;183:4–18. http://dx.doi.org/10.1016/j.biocon.2014.11.019
3. Baird DJ, Hajibabaei M. Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Mol Ecol. 2012;21(8):2039–2044. http://dx.doi.org/10.1111/j.1365-294X.2012.05519.x
4. South African Department of Water Affairs and Forestry. National Aquatic Ecosystem Health Monitoring Programme (NAEHMP): River Health Programme (RHP) Implementation Manual. Pretoria: Department of Water Affairs and Forestry; 2008.
5. Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R. Towards writing the encyclopaedia of life: An introduction to DNA barcoding. Philos T Roy Soc B. 2005;360(1462):1805–1811. http://dx.doi.org/10.1098/rstb.2005.1730
6. Hebert PD, Cywinska A, Ball SL. Biological identifications through DNA barcodes. P Roy Soc Lond B Bio. 2003;270(1512):313–321.
7. Janion C, Bedos A, Bengtsson J, Deharveng L, Jansen van Vuuren B, Leinaas HP, et al. Springtail diversity in South Africa. S Afr J Sci. 2011;107(11–12), Art. #582, 7 pages. http://dx.doi.org/10.4102/sajs.v107i11/12.582
8. Nadel RL, Slippers B, Scholes M, Lawson S, Noack A, Wilcken C, et al. DNA barcoding reveals source and patterns of *Thaumastocoris peregrinus* invasions in South Africa and South America. Biol Invasions. 2010;12(5):1067–1077. http://dx.doi.org/10.1007/s10530-009-9524-2
9. Park D, Foottit R, Maw E, Hebert PD. Barcoding bugs: DNA-based identification of the true bugs (Insecta: Hemiptera: Heteroptera). PLoS ONE. 2011;6(4):e18749, 9 pages. http://dx.doi.org/10.1371/journal.pone.0018749
10. Ekrem T, Stur E, Hebert PD. Females do count: Documenting Chironomidae (Diptera) species diversity using DNA barcoding. Org Divers Evol. 2010;10(5):397–408. http://dx.doi.org/10.1007/s13127-010-0034-y
11. Ratnasingham S, Hebert PD. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Mol Ecol Notes. 2007;7(3):355–364. http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x
12. Baird DJ, Pascoe TJ, Zhou X, Hajibabaei M. Building freshwater macroinvertebrate DNA-barcode libraries from reference collection material: Formalin preservation vs specimen age. Freshwater Sci. 2011;30(1):125–130. http://dx.doi.org/10.1899/10-013.1
13. Hajibabaei M, Smith M, Janzen DH, Rodriguez JJ, Whitfield JB, Hebert PD. A minimalist barcode can identify a specimen whose DNA is degraded. Mol Ecol Notes. 2006;6(4):959–964. http://dx.doi.org/10.1111/j.1471-8286.2006.01470.x
14. Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, et al. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. Ecol Lett. 2013;16(10):1245–1257. http://dx.doi.org/10.1111/ele.12162
15. Epp LS, Boessenkool S, Bellemain EP, Haile J, Esposito A, Riaz T, et al. New environmental metabarcodes for analysing soil DNA: Potential for studying past and present ecosystems. Mol Ecol. 2012;21(8):1821–1833. http://dx.doi.org/10.1111/j.1365-294X.2012.05537.x
16. BOLD Systems. BOLD Systems Databases. c2014 [cited 2015 Jun 20]. Available from: http://www.boldsystems.org/index.php/databases
17. Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, et al. Environmental DNA for wildlife biology and biodiversity monitoring. Trends Ecol Evol. 2014;29(6):358–367. http://dx.doi.org/10.1016/j.tree.2014.04.003
18. Mächler E, Deiner K, Steinmann P, Altermatt F. Utility of environmental DNA for monitoring rare and indicator macroinvertebrate species. Freshwater Sci. 2014;33(4):1174–1183. http://dx.doi.org/10.1086/678128
19. Sigsgaard EE, Carl H, Møller PR, Thomsen PF. Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. Biol Conserv. 2015;183:46–52. http://dx.doi.org/10.1016/j.biocon.2014.11.023
20. De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, et al. DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. Mol Ecol Resour. 2014;14(2):306–323. http://dx.doi.org/10.1111/1755-0998.12188

21. Lallias D, Hiddink JG, Fonseca VG, Gaspar JM, Sung W, Neill SP, et al. Environmental metabarcoding reveals heterogeneous drivers of microbial eukaryote diversity in contrasting estuarine ecosystems. ISME J. 2015;9(5):1208–1221. http://dx.doi.org/10.1038/ismej.2014.213

22. Cristescu ME. From barcoding single individuals to metabarcoding biological communities: Towards an integrative approach to the study of global biodiversity. Trends Ecol Evol. 2014;29(10):566–571. http://dx.doi.org/10.1016/j.tree.2014.08.001

23. Leray M, Knowlton N. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proc Natl Acad Sci USA. 2015;112(7):2076–2081. http://dx.doi.org/10.1073/pnas.1424997112

24. Jones M, Ghoorah A, Blaxter M. jMOTU and Taxonerator: Turning DNA barcode sequences into annotated operational taxonomic units. PLoS ONE. 2011;6(4):e19259, 10 pages. http://dx.doi.org/10.1371/journal.pone.0019259

25. Turner CR, Uy KL, Everhart RC. Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biol Conserv. 2015;183:93–102. http://dx.doi.org/10.1016/j.biocon.2014.11.017

26. Strickler KM, Fremier AK, Goldberg CS. Quantifying effects of UV-B, temperature, and pH on eDNA degradation in aquatic microcosms. Biol Conserv. 2015;183:85–92. http://dx.doi.org/10.1016/j.biocon.2014.11.038

27. Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, Taberlet P, et al. Persistence of environmental DNA in freshwater ecosystems. PLoS ONE. 2011;6(8):e23398, 4 pages. http://dx.doi.org/10.1371/journal.pone.0023398

28. Laramie MB, Pilliod DS, Goldberg CS. Characterizing the distribution of an endangered salmonid using environmental DNA analysis. Biol Conserv. 2015;183:29–37. http://dx.doi.org/10.1016/j.biocon.2014.11.025

29. Deiner K, Altermatt F. Transport distance of invertebrate environmental DNA in a natural river. PLoS ONE. 2014;9(2):e88786, 8 pages. http://dx.doi.org/10.1371/journal.pone.0088786

30. Swartz ER, Mwale M, Hanner R. A role for barcoding in the study of African fish diversity and conservation. S Afr J Sci. 2008;104(7–8):293–298.

31. African Centre for DNA Barcoding [homepage on the Internet]. c2013 [cited 2016 Jan 14]. Available from: http://www.acdb.co.za/

32. Coetzer W. A new era for specimen databases and biodiversity information management in South Africa. Biodivers Inform. 2012;8(1), Art. #4263, 11 pages. http://dx.doi.org/10.17161/bi.v8i1.4263

33. Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol. 2012;21(8):2045–2050. http://dx.doi.org/10.1111/j.1365-294X.2012.05470.x

34. Coissac E, Hollingsworth PM, Lavergne S, Taberlet P. From barcodes to genomes: Extending the concept of DNA barcoding. Mol Ecol. 2016;25(7):1423–1428. http://dx.doi.org/10.1111/mec.13549

35. Pauls SU, Alp M, Bálint M, Bernabò P, Čiampor F, Čiamporová-Zat'ovičová Z, et al. Integrating molecular tools into freshwater ecology: Developments and opportunities. Freshwater Biol. 2014;59(8):1559–1576. http://dx.doi.org/10.1111/fwb.12381

36. Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P. DNA metabarcoding and the cytochrome *c* oxidase subunit I marker: Not a perfect match. Biol Lett. 2014;10(9), Art. #20140562, 4 pages. http://dx.doi.org/10.1098/rsbl.2014.0562

37. Hebert PDN, Gregory TR. The promise of DNA barcoding for taxonomy. Syst Biol. 2005;54(5):852–859. http://dx.doi.org/10.1080/10635150500354886

From research excellence to brand relevance: A model for higher education reputation building

AUTHORS:
Nina Overton-de Klerk[1]
Marilet Sienaert[2]

AFFILIATIONS:
[1]Department of Strategic Communication, School of Communication, University of Johannesburg, Johannesburg, South Africa
[2]Research Office, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Nina Overton-de Klerk

EMAIL:
ndeklerk@uj.ac.za

POSTAL ADDRESS:
Dept of Strategic Communication, University of Johannesburg, APK Campus, PO Box 524, Auckland Park 2006, South Africa

DATES:
Received: 14 Sep. 2015
Revised: 04 Nov. 2015
Accepted: 28 Jan. 2016

KEYWORDS:
social impact; stakeholders; strategic communication; strategic research

HOW TO CITE:
Overton-de Klerk N, Sienaert M. From research excellence to brand relevance: A model for higher education reputation building. S Afr J Sci. 2016;112(5/6), Art. #2015-0365, 8 pages. http://dx.doi.org/10.17159/sajs.2016/20150365

In this article we propose a novel approach to reputation development at higher education institutions. Global reputation development at higher education institutions is largely driven by research excellence, is predominantly measured by research output, and is predominantly reflected in hierarchical university rankings. The ranking becomes equated with brand equity. We argue that the current approach to reputation development in higher education institutions is modernist and linear. This is strangely out-of-kilter with the complexities of a transforming society in flux, the demands of a diversity of stakeholders, and the drive towards transdisciplinarity, laterality, reflexivity and relevance in science. Good research clearly remains an important ingredient of a university's brand value. However, a case can be made for brand relevance, co-created in collaboration with stakeholders, as an alternative and non-linear way of differentiation. This approach is appropriate in light of challenges in strategic science globally as well as trends and shifts in the emerging paradigm of strategic communication. In applying strategic communication principles to current trends and issues in strategic science and the communication thereof, an alternative model for strategic reputation building at higher education institutions is developed.

Introduction

The rise of the post-industrial knowledge economy has placed a strong value on higher education institutions as 'engines of development'.[1] This has dramatically changed the role of higher education institutions from being an elite system to a more universal system. The previous elite system was characterised by a gross enrolment ratio of 15% of the population, whereas current ratios are up to 70% in some advanced economies.[2]

Governments increasingly see investment in higher education and in research and development as essential for ensuring the knowledge base necessary for economic growth. Knowledge production is regarded as more important than capital or labour.[3] As a result, competition in the higher education sector has become intense. How to differentiate each institution from the rest has become the main challenge. It has become evident that the most powerful differentiator is research excellence, as research is a critical function in the production of new knowledge and the enhancement of society.

Reputation building at higher education institutions has therefore to a large degree become premised on the construct 'research excellence'. While it is acknowledged that other factors contribute to the reputation of higher education institutions, such as teaching and learning, academic freedom, tradition, facilities, and student experience, research excellence is widely regarded as the key reputation builder.[2-4]

At the same time, research itself has come under intense scrutiny as new challenges and issues affecting the relationship between science and society have arisen. Growing tensions about the democratisation and re-contextualisation of science have forced scientists to become more reflexive about their own role in, and effect on, a complex society that engages with a multiplicity of stakeholders. Even the concept of research excellence has come under scrutiny. These dynamics have an influence on higher education institutions as knowledge producers in society, although such forces also extend beyond those institutions.

In addition, higher education institutions in South Africa are struggling with their own set of challenges. They need to remain responsive to local problems while striving towards global competitiveness, and must maintain a balance between accessibility and global reputation. Resultant tensions between excellence and transformation imperatives were brought to the world's attention when the #RhodesMustFall campaign swept across the country, resulting not only in the removal of the Rhodes statue at the University of Cape Town but also a name change of Rhodes University, and the Open Stellenbosch campaign. All these issues culminated in the #FeesMustFall student protests, which have significantly changed the national scenario and have irrevocably changed the climate in which higher education institutions operate and are funded. High on the agenda are issues such as the decolonisation of knowledge, free education, and other challenges in a society fissured by educational inequalities.

While mindful of competing binaries such as global vs. local, excellence vs. relevance, and excellence vs. transformation, we argue that the current approach to reputation building at higher education institutions is steeped in modernist, fixed and linear notions. The current approach focuses strongly on excellence. However, the underlying modernist and linear assumptions are being increasingly challenged, not only with regard to the role of science in society, but also with regard to the emerging multi-paradigmatic approach to strategic communication and reputation building. We offer a novel concept for reputation building at higher education institutions, that of brand relevance. Although higher education institutions historically have been hesitant to embrace brand and business principles, and profess to be uncomfortable with the premise of university brand building, branding in the higher education sector is not an entirely new concept. Higher education branding in the past was, however, built around generic constructs or linear rankings, which in turn led to similarity instead of differentiation among higher education brands in the reputation race.

Research objectives

This paper is conceptual and analytical in nature. One of our key research objectives is to reflect on the coherence that emerges from the diverse and fragmented academic, meta-scientific and industry-related literature. Such literature is relevant for higher education reputation building from a transdisciplinary perspective. Through the identification of trends and issues in strategic research, and trends and shifts in strategic communication, we adopt a multi-paradigmatic approach to reputation building and propose a new model for reputation building. The purpose is to contribute to the much-needed high-level understanding of the complex nature of higher education reputation building not only in Africa, but globally, and to raise the possibility of further debate and research.

The approach is structured in the following ways:

1. To consider key challenges and trends emerging in strategic research that affect the relationship between science and society, including the communication of science to society.
2. To compare these trends with key trends and shifts in the emerging paradigm of strategic communication, and to reflect on how challenges in science communication can be addressed by strategic communication.
3. To consider the implications of strategic communication for strategic branding and brand relevance, from a multi-stakeholder perspective.
4. To examine research excellence as a global reputation builder at research-intensive higher education institutions.
5. To build a case for changing the discourse from 'brand excellence' to 'brand relevance'.
6. To propose an alternative model for reputation building at higher education institutions, and for building a purposeful higher education brand.

Strategic research: Challenges and issues

Strategic research denotes 'applied research with a long term perspective'.[5,6] It gained momentum in the 1980s because of wide interest in the knowledge economy and in scientific technologies as an engine for economic growth.[5]

Strategic research combines two principles, namely *excellence* and *relevance*, which are not regarded as contradictory. The spread of theme-based, problem-oriented centres of research excellence and relevance across the globe bears evidence of how important strategic research has become. At the same time, a new set of challenges is emerging that has forced strategic researchers to reconsider the role of science in society, as discussed below.

Research uptake

The European Commission MASIS report (EC MASIS report[5]) for monitoring activities of science in society expresses reservations as to whether strategic research has not perhaps evolved into a type of basic research. Equally questionable is whether the gap between research and its eventual uptake has not become larger. Research results circulate mainly among researchers themselves, contributing to a reservoir of scientific knowledge, visible in the contents of scientific journals. Other researchers 'fish' in the reservoir and create new combinations of knowledge. This scenario may result in a somewhat incestuous cycle where the greatest impact of research is on researchers themselves, and other people who might have benefited from the results remain out of the loop.

Biswas and Kircherr[7] argue that 'the impact of most peer-reviewed publications even within the scientific community is miniscule'. They state that scholars' publications in the popular media must count as well, as these are far more likely to shape public debate and influence policies.

Globally and on the continent of Africa, public and private sector funders of research, as well as research chairs at higher education institutions, are increasingly engaged in evaluating the socioeconomic effects of research.[8] The San Francisco Declaration on Research Assessment, which was signed by a wide and diverse group of individuals and interest groups across the global North, calls for improvement in the way research output is evaluated. The Declaration supports assessing research on its own merits, eliminating journal-based metrics, and promoting practices that focus on the 'value' and 'influence' of specific research outputs.[9] Similarly, Vale, in an article titled 'Evaluating how we evaluate', calls for the evaluation of research *quality* instead of *quantity*, and stresses the need for researchers to contribute to community, society and education itself.[10]

Reflexive research

Reflexive research is a continual, evolving process of observing and reflecting upon knowledge itself, and about its value and applicability to new, complex emerging contexts.[11]

Instead of a *linear* model of innovation, strong arguments are being made for a *lateral* model, in which the transformative effect of research is emphasised. This transformative potential is especially relevant in developing societies and may be guided by leadership that spans traditional boundaries.[5,8] Of particular relevance here is the reciprocal relationship between research and a variety of stakeholders.

Stakeholder theory, as described in the influential work of Freeman[12] and other scholars,[13] denotes the influence of multiple stakeholders that do not have a direct stake in the institution but are viewed as active 'influencers' who can affect the actions of the institution.[14] The stakeholder concept has clearly taken root in research fields. The influence of a multiplicity of stakeholders at all levels of research is such that science has been required to become much more reflexive regarding its nature, the contexts in which it operates, and the stakeholders who affect – and can be affected by – research.[5] Public scrutiny of research has become a fact of life and provides a critical evaluation of expertise. Questions about how research can enhance society highlight this trend, as do calls for the decolonisation of knowledge and the development of 'Africa-rooted' evaluation models.[15,16] Reflexive research has now become the mediator between stakeholder positions and scientific interests, via expert narratives linked to evidence that is robust enough for stakeholders in society.[5]

Re-contextualisation of science

Proposals for the re-contextualisation of science began with scholars[17,18] who argued that a number of changes have led to a 'Mode 2' of knowledge production. This new mode is characterised by the following ideals:

- fluidity and changing research teams
- a more general distribution of research
- contextualisation of application
- transdisciplinarity
- new forms of quality control and social robustness.

Similarly, Cilliers[19] argues that complexity must connect with contextualised information, leading to the integration of the observer with the observed. The transdisciplinary approach has gained considerable ground, guided as it is by the argument that humanity requires a research approach that transcends narrow disciplines. This broader approach would assist us to engage with complex and interlinked problems, such as the logic and ethics of natural sciences, climate change, poverty, systemic unemployment, and so on – particularly with respect to the continent of Africa.[20]

The debate has escalated with a growing tension between Western science and the rest of society, also referred to in the EC MASIS report.[5] In some quarters, Western science and notions of universality are regarded as ideology in itself.[21] According to Max-Neef[22],

> *The growing rupture in communication is, to a large extent the product of the exacerbation of rational thought, which manifests itself through the predominance of reductionism and of a binary and linear logic that, among other shortcomings, separates the observer from the observed.*

Thus, the transdisciplinary paradigm is an attempt at formulating an integrative, holistic process of knowledge production that goes beyond a multidisciplinary or interdisciplinary approach. Yet it is also partly a reaction against the 20th century features of 'undeserved deference to authority', 'stifling disciplinary specialisation' and 'methodological commodification'.[20] The emerging trend is thus the need for a 're-contextualisation of science'[5] and the integration of knowledge paradigms that involve a multiplicity of stakeholders.

Is 'excellence' still relevant?

'Research excellence' has experienced a major revival since 2000, with the establishment of research councils around the world, particularly in the global North. These councils' mission is to support excellence, and they include the German Excellenz Initiative, Australia's Group of Eight (CO8), the United States' Ivy League, the United Kingdom's Russell Group, and China's C9 League. This revival gained momentum through the continued emphasis on measurable research output (in ISI journals) in assessments and evaluations.

Excellence and relevance are not necessarily mutually exclusive. However, the question remains whether a one-sided emphasis on excellence, or the choice of indicators for measuring excellence, might hamper the pursuit of relevance.[5] According to the EC MASIS report,[5]

> *A case in point is the increasing importance of the ISI impact factor system which favours decontextualized and globalised science while context-related and more local research, dedicated to specific problem solving, is disadvantaged. Sciences could lose their link to practice resulting from the pressure to publish in international journals instead of engaging in local environments and problem solving. Thus there is a (perhaps unintended) tendency to bring science back to a more separated, perhaps isolated and more autonomous activity, following its own rules and hunting for impacts in the ISI system rather than in the 'real world'.*

That relevant research must be good research is not contested. However, the idea of excellence can be interpreted as 'being better than others in some competition, rather than being good'.[5] The notion of excellence has thus become the basis on which universities differentiate themselves from one another, as evident in university rankings worldwide. The result is that diversity is 'often … seen through the lens of superior or inferior status, a phenomenon which is aggravated by the halo effect of global rankings'.[23]

This scenario in turn relates to the linear, hierarchical approach to university reputation rankings around the globe, which is largely based on research output in accredited international journals. Mouton[24] warns against this vertical approach as 'a scientometric discourse' that largely occurs on the descriptive level and depicts what universities look like in terms of absolute scientific output. Instead, Mouton claims that higher education differentiation is embedded in widely different discourses. Is it a discourse about allocation of national resources, or redress and transformation, or competitiveness, or strategic positioning? Each discourse gives priority to specific criteria when measuring differentiation.

Communicating research

One of the biggest challenges emerging in research is how to communicate research, and how to bridge the gaps between knowledge production, policy and application. The issue is not the supply of knowledge (which is abundant), but how to make this knowledge accessible to society and engage in communication and dialogue about it. According to the EC MASIS report,[5] although researchers and research councils or agencies are trying to engage the public, the practice of knowledge communication still seems to be unidirectional. It is also steeped in ideological debate, with accusations that the global evolution of knowledge – primarily flowing from the North – has influenced local knowledge paradigms, and that 'neo-liberal' concepts such as a 'knowledge economy' primarily serve the interests of economies in the global North.[25]

Compounding the issue of research communication is the age of interactivity. The development of Web 2.0 technology and the rise of e-science have created new opportunities for improving public understanding of science. However, these opportunities bring their own challenges. Certain scholars believe that scientists are still not realising the potential of the Worldwide Web, which was developed as a scientific collaborative workspace by Berners-Lee in 1989.[5] Advocates of open access systems argue that the rise of e-science requires methods that enable the open, immediate and free sharing of knowledge and peer-reviewed literature on the Web. They argue that open access is not a luxury but rather a necessity, particularly in societies where educational inequality exists.[26] Not only do open access systems improve the speed at which science reaches the rest of society, but they also help researchers to communicate online more rapidly and collaborate more effectively.[27]

The principle of open access is still resisted by many researchers. This is despite the new policies that higher education institutions have either put in place already or are in the process of doing so, in response to funder requirements. Many people still regard the primary means of scientific communication as formal scientific publications, and individual career and institutional assessments are based on such publications.[5,10] Internet content is regarded as volatile and perishable, whereas scholarly journals produced by prominent publishers are seen as more prestigious and lasting. Also, open access can provide public access to knowledge that might be misinterpreted and awaken the 'irrational masses'.[5]

In the following section, we discuss the manner in which the above issues and challenges in strategic research are reflected in the emerging paradigm of strategic communication.

Strategic communication: Trends and shifts

To the extent that research has become strategic and reflexive, communication has also become strategic and reflexive. Overton-de Klerk and Verwey[28] discuss the trends and shifts leading to the emerging paradigm of strategic communication against the backdrop of four key epistemological tenets that underlie postmodern knowledge and communication. These are *emergence*, *reflexivity*, *difference* and *resistance*.[29-32]

Strategic communication, simply defined as 'purposeful communication to achieve a mission',[28,33] is essentially the result of the digital communication revolution. This revolution has taken the control of information out of the hands of a limited elite and made it available to many people. Abundant literature exists on the shifts that led to strategic communication,[28] which are summarised in Table 1. The table also shows how these shifts have been reflected in strategic science communication.

The shifts in strategic communication, as discussed in Table 1, all essentially indicate that in a digitally interactive era, power has shifted from institutional communicators to individual recipients. Top-down, unidirectional transmission of information aimed at achieving consensus is no longer a sustainable model. Strategic communication must allow creative solutions to evolve spontaneously, bottom-up through active participation in dialogue.

These shifts in strategic communication are not necessarily equally reflected in the communication of strategic science, even if they feature in the scientific debate. Although strategic science has become more collaborative and reflexive, science communication – according to many scientists themselves – needs to open up and encourage more socially accountable, transparent and participatory modes of transaction between science and stakeholders in society.

Purposeful brands: The case for brand relevance

Brands are made up of a complex set of tangibles and intangibles, where the whole is bigger than the sum of the parts. 'Tangibles' refers to the product itself (inherent value), whereas 'intangibles' refers to the added value, including associations with the brand (perceived value). The latter component is the most vulnerable to risk, and can – in a matter of days, if not hours – affect public perceptions of the whole. Several South African higher education institutions experienced this phenomenon during the #FeesMustFall protests.

Table 1: Shifts in strategic communication

Shifts	As they manifest in strategic communication	As they manifest in strategic science communication
From top-down to bottom-up	Linear one-way transmission has progressed to two-way transaction.[28,34] An 'evolving and emerging process of discourse and negotiations'.[35]	Remains stuck in transmission mode.[5,18,36] Prevailing model is top-down: science makes discoveries and makes them available to society.[37] Hidden agenda (according to some) to promote fascination with natural sciences and engineering.[5] 'Public understanding of science' nothing but marketing to promote economic and innovation interests.[5]
From monologue to dialogue	Communication flow has moved from one-way monologue delivered by top management (aimed at ensuring compliance and agreement) to an inclusive and unpredictable dialogue among institutional stakeholders at all levels[28] – e.g. #FeesMustFall. Insistence on transparency, particularly among under-30 generation (more than 50% of world's population.[38])	Increasing calls to produce research communication suitable for dialogue (e.g. transdisciplinary approach[20]), as a unidirectional flow is no longer sustainable.[5,36] Increasing demands for open access to science.[26,27]
From consensus to dissent	Communication not intended to achieve consensus but to hear multiple voices and dissent – which is not only tolerated but encouraged.[26] Meaning is created through influence, not power. Emphasis on process rather than outcomes of discourse.[39]	Allows debate in early stages, but deliberations are closed when consensus is reached. Conflicts are 'managed'.[5] Insistence that deliberations must remain open, particularly when a diversity of stakeholders may be affected[5] (e.g. nanotechnology, fracking, Homo Naledi).
From control to self-organisation	Institutions no longer control outcomes but allow creative solutions to spontaneously evolve bottom-up, through active participation in dialogue[28,35,40] Emphasis on sense-giving and sense-making activities.[28]	Resonates with Gibbons' call for transparent, participative and self-organising contract between science and society, based on the 'joint production of knowledge by science and society'.[37] Must generate its own accountability and audit systems.
From social responsibility to accountability	Institutions no longer merely responsible, but now also accountable, for contributions to society and environment.[41] Stakeholders active and activist.[42,43] Accountability is not owned by institution but granted by stakeholders – earned through transparency and congruency between words and deeds.[44]	Increasing calls for 'social accountability of science'[36] and 'accountable systems of knowledge production'.[37] Accountability now a measure of sustainability in HE institutions, as the harbinger of transformation in developing societies.[42]
From integration to co-creation	From one voice (integration[45]) to diverse voices (co-creation[28]), particularly raised in consumer-generated media.[46] Institutions no longer control messaging and content[38] but rather allow for strategic co-creation in collaboration with stakeholders (important in institutional branding). Whole is bigger than sum of the parts.	Same shift has occurred in some areas of research,[20] where multidisciplinary (additive) research has evolved into interdisciplinary (integrative) research, or into transdisciplinary (holistic, co-creative) research that produces a single –sometimes complex – multidimensional result. The whole differs from the parts.[20]

The shifts in strategic communication have also exerted a profound influence on brands. The most important shift is that brands can no longer, through the mass media, control what stakeholders think of them. Stakeholders can now rewrite the script, often in consumer-generated media and hash-tag campaigns. This can place the brand reputation, however carefully constructed, at risk. As a result, institutions are now forced to engage in dialogue to minimise their reputational risk. The need for such dialogue – both online and offline – is the strongest among the so-called 'millennials',[47] many of whom are enrolled at higher education institutions.

In other words, brands can no longer rest on their laurels and rely on a reputation of excellence alone. The very foundations on which those reputations are built are increasingly being challenged. Instead, brands are required to constantly reflect upon the values they represent, the value they add, and their purpose.[48] The questions that need to be asked include: Who are we? What is it that we do? What difference do we make, and to whom? In short, are we relevant?

Although other definitions of 'brand relevance' exist,[49,50] in this context the term refers to brands that have a purpose which matters to all of their stakeholders. Constantly reflecting upon a brand's relevance has become the mantra for building a sustainable brand.

Current discourse: Research excellence and higher education reputation

A growing number of higher education institutions worldwide aspire to research excellence, and claim to be research-intensive institutions. How 'research excellence' is best defined, achieved and measured remains an open question. The concept often presumes the exclusion of multi-stakeholder collaboration for the common good. This view is enforced by widely recognised global ranking systems of higher education institutions, such as Quacquarelli Symonds (QS), the Academic Ranking of World Universities (ARWU) of Shanghai Jiao Tong University, and the Times Higher Education (THE) rankings. Although softer indicators – such as peer and employer reputation – are sometimes included in the rankings, it is indisputable that all top-ranked universities hold that position mainly due to their international published and measurable research output.[2,3,24]

However, when research excellence is understood too narrowly and is defined solely by rankings, there is considerable slippage between the value of research excellence as a measurable indictor and its worth to society. By consensus, the rankings themselves are biased and flawed.[51,52] In this paper we focus on the formulaic approach to excellence, in order to clearly juxtapose this one-sided view against the need to understand 'excellence' more broadly.

In the narrow rank-focused approach to 'research excellence', research-intensive universities usually show several main features. Firstly, they emphasise postgraduate training and research, hence shifting their attention away from basic undergraduate education to the more high-powered stakes of postgraduate activity as the machine for innovation and engine-room of research. As a result, innovative models of supervision are in place, over and above the classic one-on-one training. These new models capitalise on both supervisory capacity and the peer-learning capabilities of students. Committee-based supervision and team-based hubs, where several PhD students work on a common theme under one supervisor with the help of a mentor and postdoctoral fellow, are good examples. These models are used even in disciplines that do not have a tradition of team-based research.

The postgraduate enterprise is valorised and care is taken to deliver an optimal training experience. Students present papers at conferences and spend time at the universities of their supervisors' international research partners. The result is a next-generation of potential academics who emerge from their studies already well networked and embedded in the global community of scholars. Such graduates are able to hold their own in any research environment, and are ready to take up junior academic posts.

Secondly, strategic partnerships, particularly those of global reach, are a key ingredient. It is well established[24] that publications that are co-authored by international collaborators are more visible, and have a higher impact, than single-author papers that are published in local journals. Research-intensive universities seek out the most advantageous partnership agreements and prestigious research excellence networks. Agreements for research collaboration, co-supervision, and the exchange of staff and students with Ivy League Universities across the world are obvious examples. So is the membership by invitation from networks such as the Worldwide Universities Network (WUN) and the International Alliance of Research Universities (IARU). The Alliance of Research Universities in Africa (ARUA) that was launched in March 2015 in Dakar is a more recent example.

In addition to prestige, the obvious value of such membership is the gearing effect of structured relations and dedicated funding. These networks often create a 'virtuous circle' of enhanced postgraduate experience through joint training and exposure to multiple laboratory or fieldwork sites, as well as strengthened relations between principal investigators. In turn, these factors lead to a greater number of co-authored publications and joint funding proposals.

A third key ingredient of universities that rely on research excellence as a defining factor is the clustering of expertise and critical mass to create theme-based, problem-focused centres of excellence. These hubs drive large-scale interdisciplinary projects around a common theme of global significance. Examples of such topics are climate change, poverty alleviation and sustainable environments. The research hubs have strong academic leadership and a team of researchers who spend at least 60% of their research time on related projects. As vehicles of collaboration, they drive international partnerships and leverage the biggest and best grants available globally. They train postgraduates and serve as knowledge incubators, thus creating a self-sustaining cycle of excellence and productivity.

Fourthly, staff recruitment at institutions that pursue the rankings is shaped by a passion for excellence. There is no time to nurture and slowly grow promising young scholars if that might slow down the generation of funding grants. The environment is far too competitive and pressurised to favour candidates who merely have potential over those who are already well established. In the South African context, where transformation of the science cohort is a priority, this scenario creates tensions that some say are impossible to resolve.[53]

Lastly, visibility is essential in a globally connected world. This has led to dramatically increased investment from institutions, over the past decade or so, in web-based platforms and open-access institutional repositories. Pockets of excellence and increased outputs are foregrounded, and researcher profiles and the institution's infrastructure capabilities are promoted. Such online visibility stimulates the imagination about the scope and reach of research as an activity per se. Interactive web-based portals and 'brag' sites display researcher profiles and tell stories about paradigm-shifting research findings. Monthly or weekly e-research newsletters are widely distributed, and a range of glossy promotional material is produced. All project an image of world-changing excellence. In the 'dating game' of international research partnerships, such visibility is a key requirement to assert a research-intensive identity.

Indeed, the *Webometrics* ranking of higher education institutions has elevated the visibility of research excellence, thereby increasing competition in this area and making investment in visibility inseparable from investment in research itself. This is sometimes to the despair of university planners.

Claiming research excellence, as defined by rankings, is therefore an expensive enterprise. Wealthy institutions in developed economies are in a position to pick and choose their students and research partners, while institutions in the developing world vie for recognition in research excellence. The recently established ARUA is a case in point, as research excellence is the suggested common denominator to make this a preferred network for global engagement. That all the member institutions are not necessarily the strongest research universities on the African continent seems immaterial, in this context of building a Pan-African knowledge economy in a highly politicised environment. No wonder then that social responsiveness and the common good are sometimes sacrificed in the quest to strengthen the image of research excellence.

Simply scorning the concept of excellence as being exclusive and elitist carries a great risk. It is essential, instead, to bring Southern perspectives to global challenges from a position of equity and strength, and so to bring insights from Africa to bear on research questions of global reach. It is dangerously simplistic to equate research excellence with the 'capitalist project' and to ignore the essential role research can play in improving the quality of life for ordinary citizens. It remains necessary, but not sufficient, to retain research excellence and to move beyond that to the notion of brand relevance.

Changing the reputation discourse: Brand relevance

From the previous discussion, it is evident that a strong link exists between research output and reputation building. Indeed, research excellence – as defined by the ranking systems – is widely regarded as the most important driver of a university's national and global stature. As a result, universities tend to direct their resources into areas of high measureable output, such as research. Rankings in turn are perceived as an indication of a university's brand equity,[2,54] and this is especially true of universities with global aspirations. The basic argument is as follows:

Research output = research excellence = ranking = global reputation = brand equity.

Responding to the intense demand for higher education worldwide and the rise of the 'global university shopper', ranking systems are therefore used as a shortcut to assess the brand value of a university. Universities are arranged in a highly structured manner, giving each an absolute position in a 'hierarchical order of things'[2]. Strategic planning at universities therefore focuses on directing the organisation's energies and resources in line with the requirements of ranking systems, and increasing research output and visibility where it matters.

It can be argued that the pursuit of research excellence remains critically important for higher education institutions, and represents to a large degree their *raison d'etre* in a knowledge economy. It can also be argued that rankings as such are important for all brands, as they have a substantial effect on what the market thinks of those brands. But there are a few problems with the argument that ranking equals brand value. It is also highly questionable whether research excellence can be the basis for brand differentiation or brand excellence. This is *product thinking*, not *brand thinking*, and will lead to 'sameness' – where ultimately few will survive. According to Gibbon et al.[2]:

> *... while rankings may give the impression of something fairly fixed, the jostling for places, by universities with aspirations, will make for some dramatic disturbances in ranking, and will make*

> *it increasingly difficult for a university to even remain in its current rank position.*

Most higher education institutions tend to build their reputations around a cluster of generic constructs, such as research excellence, academic reputation and tradition. This approach leads those institutions into a conformity trap – and lately into other forms of resistance. For example, Teferra[52] describes all of Africa's flagship universities as 'identical twins.' Muller argues that if universities are left to their own devices, they 'will tend to converge, because by competing with each other, they naturally tend to imitate the institutions perceived to be of higher status'.[23] This convergence is aggravated, according to Muller,[23] by traditional academic values: the stronger those values are, the stronger the tendency towards imitation.

Indeed, it is surprising to note how modernist higher education institutions' approach to reputation and brand building has persisted, and how fixed that approach still is in linear notions of top-down rankings. According to Rensburg[55], the linear top-down system signals elitism instead of relevance. A growing number of voices are calling for a more lateral, horizontal mode of thinking where 'apples are compared with apples'.[5,24,55] The recent announcement that plans are underway to differentiate South African higher education institutions into universities, university colleges and tertiary colleges is therefore to be welcomed, if only as a first step. It now becomes even more important for each of these institutions to find their brand purpose and craft their mission niche in collaboration with stakeholders.

In the final analysis, the challenge facing South African higher education institutions should not be trivialised. According to Mouton[24],

> *Scientific research and knowledge production are complex enterprises and its measurement cannot be reduced to single indicators or even very limited ranking systems. We are currently still at the conceptual stage where the challenge is to develop a more refined set of measures that will be valid and reliable... as well as sensitive to the different policy and normative/evaluative discourses where these measures will be applied and used.*

Missing, still, from the current equations is the concept of purposeful university branding. Brand reputation is the perceived value added to the intrinsic value. Although the intrinsic qualities of product excellence may be a minimum requirement to become a brand leader, these remain brand *inputs*. To become brand leaders, they must have purpose and relevance that resonate with their stakeholders and transcend product qualities. Brand relevance is the great intangible, built on tangibles, that sets a leader apart. In other words, brand relevance is a way of re-contextualising research excellence in praxis.

To use some examples from the leading brands of our time – such as Apple, Coke or Nike, brand relevance should begin with the creation of a brand manifesto.[56] This involves crafting a mission niche that can be tied to specific areas of focus. The mission niche is a social contract with stakeholders, for which the brand is held accountable. Given the needs of a society in transformation, such a brand manifesto should be co-created with a diversity of stakeholders, alert and responsive to the complex demands of a nascent democracy and continent at large.

Higher education branding can never be based upon research or peer reputation alone. This feeds into a somewhat incestuous cycle of reputation building, which places relevance and uptake into society at risk. Scientists are not the only ones to fish from the pond of knowledge; society needs to be involved in the fishing too. Branding requires *local* input so that the difference the brand makes to its immediate surroundings is of key importance. For these reasons, it may be important to include end-users (students) and other stakeholders (such as communities) in the measurement of universities' reputations. Measuring *social impact* has become the new imperative.

As shown in Figure 1, research publications, PhD training, global networks, centres of excellence, and excellence among staff create measurable and necessary inputs towards research excellence in a higher education institution. However, for an institution to differentiate itself and find a clear positioning in a highly competitive environment, the concept of research excellence must be broadened and combined with the concept of brand relevance. Brand relevance is reflexive, inclusive and co-created from the bottom up, in collaboration with a number of internal and external stakeholders. The outputs of these interactive processes should be the crafting of a social contract (brand manifesto), including a mission niche, against which the institution can be held accountable. The contract would also enable social impact to be measured by qualitative methods. The result is bigger than the sum of its parts: a purposeful higher education brand that instils a sense of belonging with its stakeholders, especially those that are closest to it.

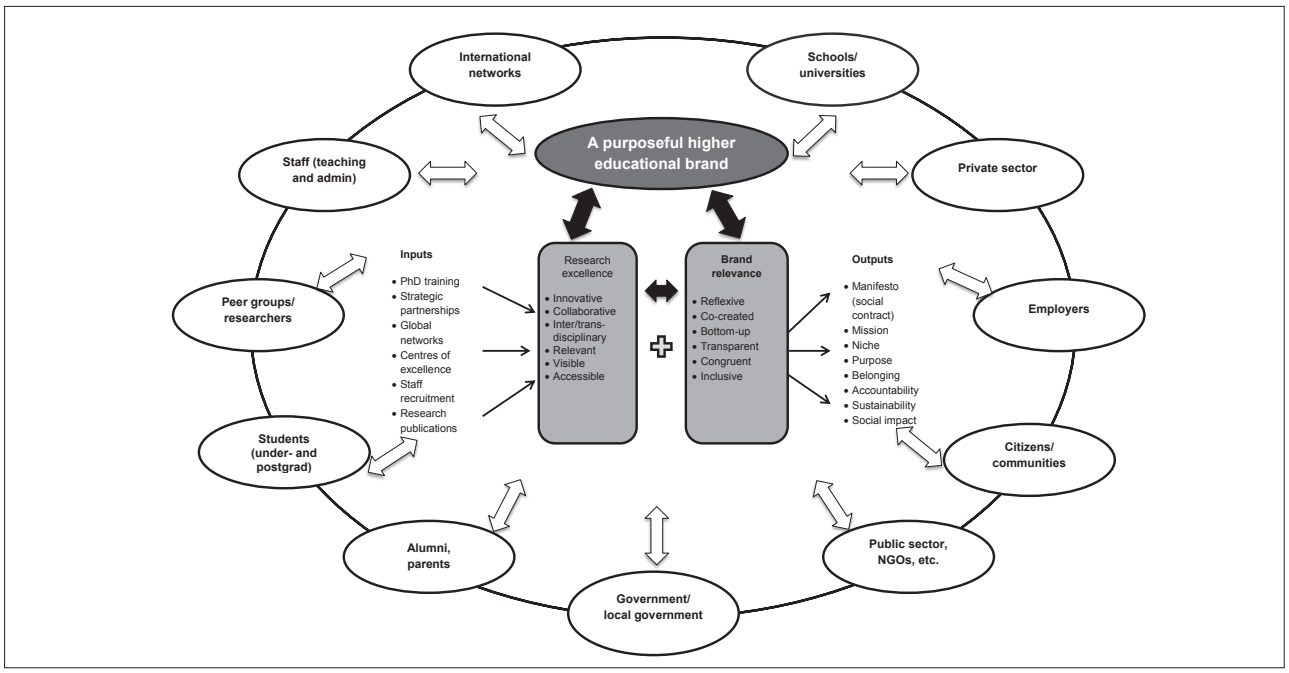


Figure 1: An alternative model for strategic higher education reputation building.

Conclusion

In this article we argue that power has shifted from institutional communicators to stakeholders. Our conclusion is based on analysis of trends and shifts in strategic science and strategic communication, in which the linear model has given way to a lateral, stakeholder-inclusive approach to strategic higher education reputation building.

The sustainability of 'research excellence' as the basis for higher education reputation building is questionable, for the following reasons:

- It is conceptualised in such a way that it can only be operationally measured by and ranked according to linear, singular indicators of research output. These outputs are no longer synchronous with new discourses in strategic science and strategic communication, and involve a limited number of stakeholders.
- It may lead to the conformity trap; excellence alone is not sufficient to differentiate higher education institutions and can lead to convergence, sameness and mediocrity.
- It can lead to the binary trap, thereby increasing reputational risk to the brand. This is because of exclusionary connotations associated with the concept of excellence, especially in transforming societies.

We therefore proposed an alternative lateral and 'stakeholder-centric' reputation model. Our model focuses on the brand instead of the product, so that the whole is bigger than the sum of the parts. Research excellence is regarded as a necessary product input, as it indicates intrinsic value. However, a higher education brand is differentiated only by shifting the focus to the perceived value of the higher education brand, and through reflection on the higher education brand's relevance and purpose. Such reflection needs to be a collaborative effort with diverse stakeholders, including internal stakeholders, end-users and communities.

The output should be the co-creation of a social contract between the higher education institution and stakeholders, or a brand manifesto. Against this contract, the institution can be held accountable and the socioeconomic impact of both the brand and the research can be tracked.

Although further work is required to refine the constructs of our model, we believe our contribution is original and heuristic, and could stimulate further much-needed research and debate. Reflections on strategic research and its relevance in society can, perhaps for the first time, combine with reflections on strategic communication and the development of a purposeful higher educational brand.

A good reputation, like charity, begins at home. Only when higher education institutions have found their niche and their purpose for being, and are able to measure their brands' role in delivering long-term socioeconomic value, can those institutions turn their attention towards building global reputation and excellence.

Authors' contributions

N.O-dK concentrated on the brand-relevance and communication aspects of the article and M.S. on the excellence aspects. The central argument and all other analytical and conceptual contributions were fully collaborative. Indeed, the whole is bigger than the sum of the parts.

References

1. Castells M. The university system: Engine of development in the new world economy. In: Ransom A, Koo S, Selvaratnam W, editors. Improving higher education in developing countries. Washington DC: The World Bank; 1993. p. 14–40.
2. Gibbon T, Ryan R, Bawa R, Moteetee A, Swart A. Strengthening the academic core for global reputation. Game changing strategies to grow reputation. Paper presented at the GIBS Leadership Development Programme; 2014 June; Johannesburg, South Africa.
3. Cloete N. The South African higher education system: Performance and policy. Stud High Educ. 2014;39(8):1355–1368. http://dx.doi.org/10.1080/03075079.2014.949533
4. Van Vught F. Mission diversity and reputation in higher education. High Educ Policy. 2008;21(2):151–174. http://dx.doi.org/10.1057/hep.2008.5
5. EC MASIS report. Challenging futures of science in society. MASIS Expert Group set up by the European Commission; 2009.
6. Irvine J, Martin RB. Foresight in science: Picking the winners. London: F. Pinter; 1984.
7. Biswas A, Kircherr J. Citations are not enough: Academic promotion panels must take into account a scholar's presence in the social media [homepage on the Internet]. c2015 [cited 2015 July 14] Available from: http://blogs.lse.ac.uk/impactofsocialsciences/2015/04/09/academic-promotion-scholars-popular-media
8. National Research Foundation. An exploratory evaluation of the socioeconomic impacts of selected NRF funding instruments. Pretoria: NRF; 2015.
9. San Francisco Declaration on Research Assessment. Putting science into the assessment of research. Declaration by a group of editors and publishers of scholarly journals during the Annual General Meeting of The American Society of Cell Biology. 2012 Dec 16; San Francisco.
10. Vale RD. Evaluating how we evaluate. Mol Biol Cell. 2012;23(17):3285–3289. http://dx.doi.org/10.1091/mbc.E12-06-0490
11. Grunwald A. Strategic knowledge for sustainable development: The need for reflexivity and learning at the interface between science and society. Int J Fore Innov Pol. 2004;1:150–167. http://dx.doi.org/10.1504/IJFIP.2004.004619
12. Freeman RE. Strategic management: A stakeholder approach. Boston, MA: Pitman; 1984.
13. Parmar BL, Freeman RE, Harrison JS, Wicks AC, Purnell L, De Colle S. Stakeholder theory: The state of the art. Acad Manag Ann. 2010;4(1):403–445. http://dx.doi.org/10.1080/19416520.2010.495581
14. Coombs WT. Crisis management: Advantages of a relational perspective. In: Ledingham JA, Bruning SD, editors. Public relations as relationship management: A relational approach to the study and practice of public relations. New Jersey: Lawrence Erlbaum; 2000. p. 73–94.
15. Cloete F. Building appropriate capacity to fast-track an Africa-rooted evaluation approach. SAMEA Conference; 2015 Oct 12–16; Hilton Hotel, Sandton, Johannesburg, South Africa.
16. The Bellagio Report. African thought leaders forum on evaluation and development: Expanding leadership in Africa. Johannesburg: CLEAR, P&DM Wits University; 2013.
17. Nowotny H, Scott P, Gibbons M. Re-thinking science: Knowledge and the public in an age of uncertainty. London: Polity Press; 2001.
18. Gibbons M, Limoges C, Nowotny H, Schwartzman S, Scott P, Trow M. The new production of knowledge: The dynamics of science and research in contemporary societies. London: SAGE; 1994.
19. Cilliers P. Complexity and postmodernism. London: Routledge; 1998.
20. Du Plessis H, Sehume J, Martin L. The concept and application of transdisciplinarity in intellectual discourse and research. Johannesburg: Mapungubwe Institute of Strategic Reflection; 2014.
21. Feyerabend P. Against method: Outline of an anarchistic theory of knowledge. London: NLB; 1975.
22. Max-Neef MA. Foundations of transdisciplinarity. Ecol Econ. 2005;53(1):5–16. http://dx.doi.org/10.1016/j.ecolecon.2005.01.014
23. Muller J. Differentiation: Africa lags (again) – or does it? Univ World News. 2014 Feb 28;(309).
24. Mouton J. Measuring differentiation in knowledge production at South African Universities. Differentiation Seminar; 2014 Oct 3; Stellenbosch, South Africa.
25. Nzimande B. Universities, knowledge production and people-centered development: Reflections from Cuba. Umsebenzi Online. 2012 Feb 23.
26. Butler-Adam J. Opening up access to research and information isn't a luxury – it's a necessity. The Conversation. 2015 Oct 20.
27. Science in the web age: Joint efforts. Nature. 2005;(438):548–549.

28. Overton-de Klerk N, Verwey S. Towards an emerging paradigm of strategic communication: Core driving forces. Communicatio. 2013;39(3):362–382. http://dx.doi.org/10.1080/02500167.2013.837626

29. Foucault M. Power/knowledge. New York: Pantheon; 1980.

30. Lyotard JF. Le différend: Phases in dispute. Theory and history of literature. Vol.46. Minneapolis, MN: University of Minnesota Press; 1988.

31. Dozier DM, Lauzen MM. Liberating the intellectual domain from the practice: Public relations, activism, and the role of the scholar. J Public Relat Res. 2000;12(1):3–22. http://dx.doi.org/10.1207/S1532754XJPRR1201_2

32. Hassard J. Essai: Organizational time: Modern, symbolic and postmodern reflections. Organ Stud. 2002;23(6):885–892. http://dx.doi.org/10.1177/0170840602236010

33. Hallahan K, Holtzhausen D, Van Ruler B, Verčič D, Sriramesh K. Defining strategic communication. Int J Strateg Commun. 2007;1(1):3–35. http://dx.doi.org/10.1080/15531180701285244

34. Hanssen L, Van Katwijk M. Paradigm shift in science communication: From transmission to transaction thinking. In: Willems J, editor. Basisboek wetenschapscommunicatie. Amsterdam: Boom; 2007. p. 130–149.

35. Ströh U. Relationships and participation: A complexity science approach to change communication. Int J Strateg Commun. 2007;1(2):123–137. http://dx.doi.org/10.1080/15531180701298916

36. Carrier M, Nordmann A. Science in the context of application: Methodological change, conceptual transformation, cultural reorientation. In: Carrier M, Nordmann A, editors. Science in the context of application. Dordrecht: Springer; 2011. p. 1–7. http://dx.doi.org/10.1007/978-90-481-9051-5_1

37. Gibbons N. Science's new social contract with society. Nature. 1999;402:11–17. http://dx.doi.org/10.1038/35011576

38. Bosman M. The staggering technology trends that are reshaping consumer behaviour and business growth. Global Summit of Consumer Goods Forum; 2010 June 24; London, England.

39. Deetz S. Transforming communication, transforming business: Stimulating value negotiation for more responsive and responsible workplaces. Int J Value-Based Manag. 1995;8(3):255–278. http://dx.doi.org/10.1007/BF00942839

40. Lichtenstein BB. Self-organized transitions: A pattern amid the chaos of transformative change. Acad Manag Exec. 2000;14(4):128–141. http://dx.doi.org/10.5465/ame.2000.3979821

41. De Beer E, Rensburg R. Towards a theoretical framework for the governing of stakeholder relationships: A perspective from South Africa. J Public Aff. 2011;11(4):208–225. http://dx.doi.org/10.1002/pa.414

42. Pratt CB. Managing sustainable development in sub-Saharan Africa: A communication ethic for the global corporation. In Sriramesh K, Verčič D, editors. The global public relations handbook: Theory, research and practice. New York: Routledge; 2009.

43. Bakardjieva M. Subactivism: Lifeworld and politics in the age of the Internet. Inf Soc. 2009;25(2):91–104. http://dx.doi.org/10.1080/01972240802701627

44. Nothhaft H, Wehmeier S. Coping with complexity: Sociocybernetics as a framework for communication management. Int J Strateg Commun. 2007;1(3):151–168. http://dx.doi.org/10.1080/15531180701434785

45. Schultz DE, Tannenbaum S, Lauterborn R. Integrated marketing communications. Chicago, IL: NTC Publishing Group; 1993.

46. Satterthwaite C. Managing reputation in a transparent world. In: Marsden A, editor. Modern brand management fundamentals and latest thinking. The marketing and management collection [homepage on the Internet]. c2010 [cited 2011 Aug 23]. Available from: http://hstalks.com/lib.php?t=HST102.2516&c=250

47. Aziyona CM. Millennials' communication of brand perceptions on Facebook: implications for reputation management [master's dissertation]. Johannesburg: University of Johannesburg; 2015.

48. Neumeier M. The brand gap. Berkeley, CA: Peachpit Press; 2005.

49. Ray J. Challengers do things differently. AdVantage Magazine. 2013, May. p. 24.

50. Kotler P, Keller K. Marketing management. 13th ed. London: Pearson Prentice-Hall; 2009.

51. Hazelkorn E. Impact of global rankings on higher education research and the production of knowledge [document on the Internet]. Unesco Forum on Higher Education, Research and Knowledge, Occasional Paper No. 18. c2009 [cited 2015 Sep 10]. Available from http://unesdoc.unesco.org/images/0018/001816/181653e.pdf

52. Teferra D. Ranking African universities is a futile endeavour. The Conversation: Africa Pilot. 2015 Sept 3.

53. Badat S. Transforming South African higher education 1990-2003: Goals, policy initiatives and critical challenges. In: Cloete N, Pillay P, Basat S, Moja T, editors. National policy and a regional response in African higher education. Cape Town: James Curry & David Phillip; 2004.

54. Leonard W. An alternative path to creating a global brand. Univ World News. 2015 May 15;(367).

55. Rensburg I. Differentiation: New elite ranking by stealth? Paper presented at Africa Higher Education summit. 2015 March 11–12; Dakar, Senegal.

56. Lascaris R. Brand leadership. Global Reputation Conversation. University of Johannesburg Executive Leadership Group; 2015 Mar 20; Sandton, South Africa.

The composition of ambient and fresh biomass burning aerosols at a savannah site, South Africa

AUTHORS:
Minna Aurela[1]
Johan P. Beukes[2]
Pieter van Zyl[2]
Ville Vakkari[1]
Kimmo Teinilä[1]
Sanna Saarikoski[1]
Lauri Laakso[1,2]

AFFILIATIONS:
[1]Finnish Meteorological Institute, Helsinki, Finland
[2]Unit for Environmental Sciences and Management, North-West University, Potchefstroom Campus, Potchefstroom, South Africa

CORRESPONDENCE TO:
Minna Aurela

EMAIL:
Minna.Aurela@fmi.fi

POSTAL ADDRESS:
Finnish Meteorological Institute, Erik Palmenin aukio 1, FIN-00560 Helsinki, Finland

DATES:
Received: 10 June 2015
Revised: 09 Oct. 2015
Accepted: 01 Dec. 2015

KEYWORDS:
air quality; chemical composition; rural environment; submicron particles; sulfate

HOW TO CITE:
Aurela M, Beukes JP, Van Zyl P, Vakkari V, Teinilä K, Saarikoski S, et al. The composition of ambient and fresh biomass burning aerosols at a savannah site, South Africa. S Afr J Sci. 2016;112(5/6), Art. #2015-0223, 8 pages. http://dx.doi.org/10.17159/sajs.2016/20150223

Atmospheric aerosols play a key role in climate change, and have adverse effects on human health. Given South Africa's status as a rapidly-developing country with increasing urbanisation and industrial growth, information on the quality of ambient air is important. In this study, the chemical composition of ambient particles and the particles in fresh biomass burning plumes were studied at a savannah environment in Botsalano, South Africa. The results showed that Botsalano was regularly affected by air masses that had passed over several large point sources. Air masses that had passed over the coal-fired Matimba power station in the Waterberg, or over the platinum group metal smelters in the western Bushveld Igneous Complex, contained high sulfate concentrations in the submicron ranges. These concentrations were 14 to 37 times higher compared with air masses that had passed only over rural areas. Because of the limited nature of this type of data in literature for the interior regions of southern Africa, our report serves as a valuable reference for future studies. In addition, our biomass burning study showed that potassium in the fresh smoke of burning savannah grass was likely to take the form of KCl. Clear differences were found in the ratios for potassium and levoglucosan in the smouldering and flaming phases. Our findings highlight the need for more comprehensive chamber experiments on various fuel types used in southern Africa, to confirm the ratio of important biomass burning tracer species that can be used in source apportionment studies in the future.

Introduction

With regard to air quality, Africa is one of the least studied continents in the world. However, extensive aerosol measurements have been carried out in southern Africa within the framework of the SAFARI 92, SAFARI 2000 and EUCAARI campaigns.[1,2] The first two campaigns focused mainly on the emissions of biomass burning and regional transport in the atmosphere. The EUCAARI campaign focused on understanding the interactions between climate and air pollution around the world. Several other air pollution measurement studies in South Africa have focused on nucleation events, trace gases and optical properties of aerosols.[3-8] The chemistry of particles has been studied to a lesser extent.[3-8]

A few studies on the chemical characterisation or source apportionment of ambient aerosols have been conducted in Tanzania during the wet and dry seasons, and in various environments in Kenya.[9-13] However, apart from the SAFARI 2000 campaign (which included multiple studies in southern Africa), few reports have been published that focus on the chemical characterisation of aerosols in South Africa.[1,14] The most recent long-term measurements of aerosol chemical composition were conducted relatively close (approximately 100 km) to the Johannesburg–Pretoria megacity conurbation. In that study, different source regions were determined and the chemical characteristics of organics were investigated.[15,16]

The main objective of this paper is to contribute knowledge on the chemical composition of particles at a regional background location in North-West, a province of South Africa for which only sparse information on air quality is available. The chosen location is downwind of the Waterberg area on the dominant anti-cyclonic circulation route of air mass movement in this part of South Africa. The construction of a large new coal-fired power station, Medupi Power Station, in addition to the currently operational coal-fired power station in the Waterberg area, Matimba Power Station, highlights the need for such measurements to serve as a reference for future studies. Data on background aerosol chemical composition are also important because South Africa is a developing country with increasing urbanisation and industrial growth.

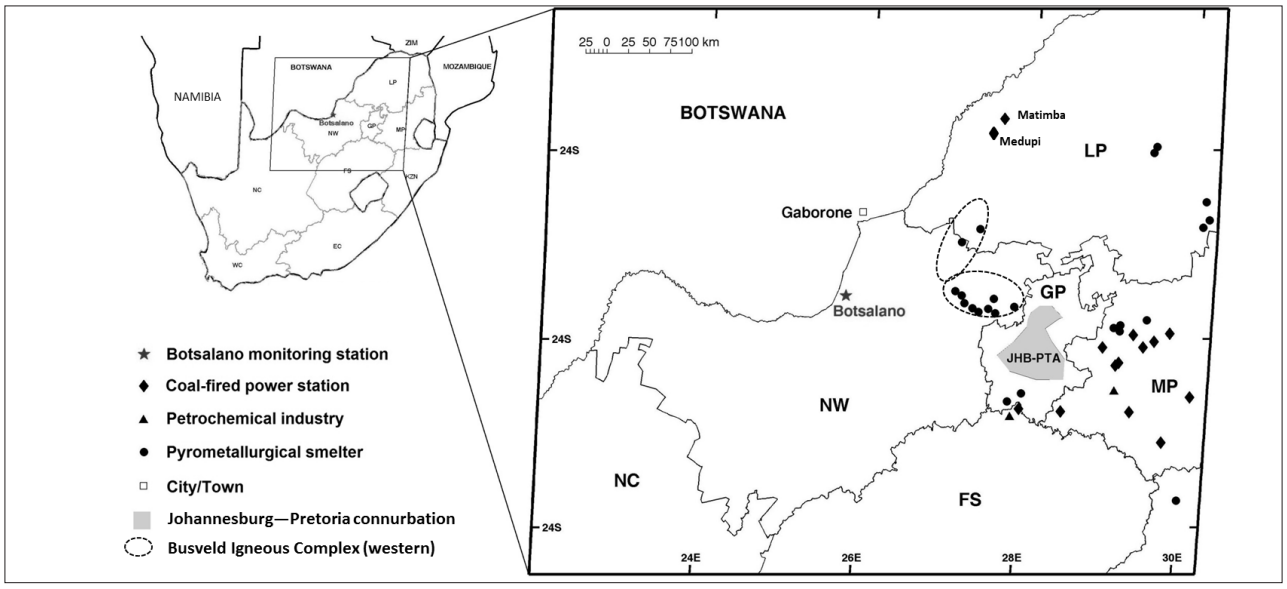
The sample site was chosen to represent a relatively clean background area with very little local pollution. The site gave us an opportunity to investigate how the chemical composition of aerosols changed depending on the origin of air masses. As far as we could assess, no other aerosol chemical composition studies have been published for this important area. Earlier studies have focused on the physical properties of aerosols at the site.[17-19] In this study we concentrated mainly on submicron particles, because particles that originate from anthropogenic and natural wildfire combustion sources are typically smaller than 1 μm in aerodynamic size.

In addition to analysing the aerosol chemical composition of regional air masses, a biomass burning experiment was conducted onsite. This enabled us to investigate the chemical composition of aerosols originating from a fresh biomass burning plume.

Methodology

Site description and measurement periods

A mobile station for atmospheric measurement was deployed in the Botsalano game reserve in North-West Province, South Africa (25.541S, 25.754E, 1424 m AMSL).[17,19] This setting can be considered a dry savannah regional background site, with no major local anthropogenic sources. Figure 1 shows the main large sources around the measurement site. Briefly, the sector from north to south, in an easterly direction, contains several large sources. Possibly, one of the largest regional pollution source areas might be the mining and pyro-metallurgical smelting activities in the western limb of the Bushveld Igneous Complex (BIC).[20-22]



Key to provinces: LP, Limpopo; GP, Gauteng; MP, Mpumalanga; FS, Free State; NC, Northern Cape; NW, North West

Figure 1: Geographical map of southern Africa, indicating the location of the Botsalano measurement site (25.541S, 25.754E) and nearby large air pollution point sources.

The ambient measurements reported in this paper were carried out in two short campaigns during 9–15 October 2007 (spring) and 30 January–5 February 2008 (summer). In total, 11 sets of impactor samples with a typical 24-hour sampling time were collected. The meteorological conditions of Botsalano during the campaigns are presented in Figure 2. The beginning of the spring campaign was slightly colder than the summer campaign; otherwise the seasons had similar temporal variation in temperatures. The relative humidity and temperature were on average 19% and 22% higher, respectively, during the summer campaign compared with the spring campaign. The summer and spring campaign represented typical seasons, according to the study by Laakso et al. (2008) at the same site in 2006–2007.[17]

In addition to the ambient samples, a small-scale biomass burning measurement experiment was performed. During the biomass burning experiment, organic materials consisting mainly of dry grass and branches collected upwind of the site were burned. The distance between the sampling equipment and the fire was approximately 50 m. Two samples were collected. The first sample was taken from the main plume, which was sampled for 44 min, and the second sample was taken over 124 min during the period when the fire intensity and fuel amount were lower.

Measurements

The samples were collected with a three-stage cascade impactor with aerodynamic cut-off diameters of 10 μm, 2.5 μm and 1 μm, followed by a backup filter (Dekati®PM10). The PM10 inlet was used for cutting off particles greater than 10 μm. The collection substrates were pre-heated quartz fibre filters (Tissuquartz, PALL). The sampling flow rate was 30 L/min and the sampling duration was approximately 24 h, except during the biomass burning experiment, when shorter sampling periods were utilised.

Chemical analysis

Organic carbon (OC) and elemental carbon (EC) were analysed from a 1 cm^2 sample punched out of the quartz fibre filters, using a thermal optical carbon analyser (TOA; Sunset Laboratory Inc. Tigard, OR, USA).

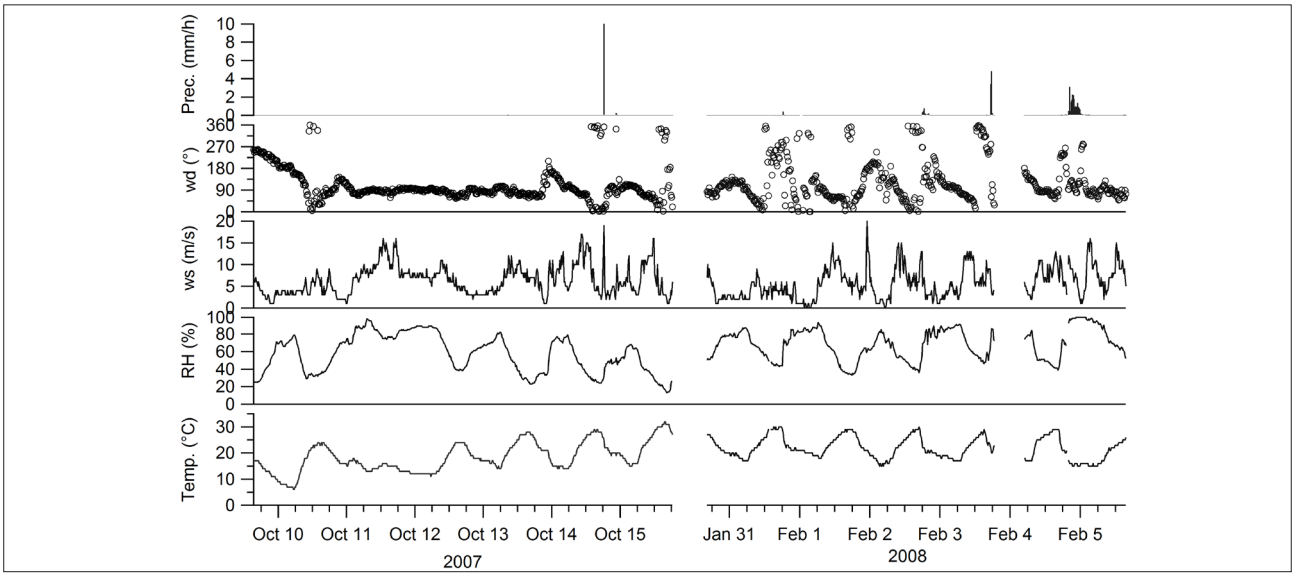


Figure 2: Meteorological parameters (temperature, relative humidity, wind speed, wind direction, precipitation) during the campaigns.

The instrument uses a two-phase thermal method to separate OC and EC (EUSAAR_1 and EUSAAR_2).[23] Optical correction was performed to separate pyrolysed organic carbon from elemental carbon.

The sampled submicron particles (PM_1) showed uniform deposition on the filters, but the larger fractions (particles >1 μm) were not uniformly spread. A previous study had shown that unevenly-spread samples yielded acceptable results when applying optical correction for impactor samples.[24] Therefore, we analysed all the samples. The amount of gaseous OC collected on the filters was not measured and hence was not subtracted from the concentrations of OC. The overall uncertainty of the TOA method was estimated to be 10% for OC and 20% for EC in concentrations above the quantification limit.

The field blank concentrations of OC for the backup filters (PM_1), calculated for a one-day sampling period, were just above the detection limit during both the spring (measured amount was 0.07 $\mu g/m^3$) and summer (measured amount was 0.10 $\mu g/m^3$) campaigns. For the impactor stages, the combined corresponding blank concentrations were 0.08 $\mu g/m^3$ and 0.09 $\mu g/m^3$ during the spring and summer campaigns respectively. The blank concentration of EC was below the detection limit for the impactor stages (0.02 $\mu g/m^3$) and for the backup filters (0.07 $\mu g/m^3$). The measured OC concentrations were multiplied by a factor of 1.6 to obtain an estimate for total particulate organic matter (POM) mass concentration.[25]

The remaining portions of the quartz filters were analysed with ion chromatography (IC) to determine selected ions, namely chloride (Cl^-), nitrate (NO_3^-), sulfate (SO_4^{2-}), oxalate, sodium (Na^+), ammonium (NH_4^+), potassium (K^+), magnesium (Mg^{2+}) and calcium (Ca^{2+}). The uncertainty of the IC analysis was estimated according to the analysis of standards as 5% to 10%, depending on the ion analysed. The field blank concentration of each ion in the backup and impactor-stage filters was 0.018 $\mu g/m^3$ and 0.008 $\mu g/m^3$ respectively, during both campaigns. The determination limit was below 0.002 $\mu g/m^3$ for all the ions.

Concentrations of monosaccharide anhydrides (levoglucosan, galactosan and mannosan) collected on backup (PM_1) filters were measured using an IC coupled to a quadrupole mass spectrometer.[26] The blank concentration of monosaccharide anhydrides was below the determination limit (1 $\mu g/L$ to 5 $\mu g/L$), which corresponded to 0.001 $\mu g/m^3$ to 0.005 $\mu g/m^3$ for a 24-hour sample.

Auxiliary data

Earlier publications[17,19] presented detailed descriptions of measurements obtained at Botsalano using a Differential Mobility Particle Sizer (DMPS), consisting of a Vianna-type Differential Mobility Analyser and TSI model 3010 Condensation Particle Counter. The measures included meteorological parameters (temperature, pressure, relative humidity, precipitation, wind speed and wind direction), trace gas concentrations (SO_2, NO/NO_x, CO and O_3) and submicron aerosol particle size distribution.

Total mass concentration of PM_1 was not measured directly but was estimated from the DMPS size distribution, for the range 10 nm to 840 nm. For particles smaller than 840 nm, the particle number concentration of each DMPS size channel was converted to volume concentration (assuming spherical particles), multiplied by the estimated particle density, and summed to obtain the total mass concentration. The particle density value was calculated from the analysed chemical composition for each sample using an approach suggested by Saarnio et al.,[27] and it varied from 1.38 to 1.75 g/cm^3.

Air mass history

Air mass backward trajectories were calculated using the NOAA HYSPLIT backward trajectory model.[28] The 96-hour backward air mass trajectories were calculated to arrive at every hour of sampling, with an arrival height of 100 m above the ground.

Results and discussion

Overall ambient air results

An overview of the meteorological pattern in Botsalano showed that the north-easterly wind direction had the highest frequency.[19] This can be expected, as the meteorological pattern over the interior of South Africa is dominated by an anticyclonic circulation pattern.[29] In addition, although Botsalano is a background site without significant local air pollution point sources, it is regularly affected by air masses that have passed over several large point sources. The Matimba coal-fired power station at Lephalale in the Waterberg, the silicon and platinum group metal (PGM) smelters near Polokwane, the PGM smelters in the Northam and Thabazimbi area, and the city of Gaborone are all situated on the dominant anticyclonic circulation path for air masses arriving at Botsalano. The mining and metallurgic activities in the western region, lying on the axis between the towns of Brits and Rustenburg (Bushveld Igneous Complex, BIC), are also likely to affect the site regularly.

The PM_1 mass concentrations calculated from the DMPS results varied between 1.4 $\mu g/m^3$ and 21.5 $\mu g/m^3$ (mean 11.3 $\mu g/m^3 \pm 5.3$ $\mu g/m^3$). The total concentrations of all analysed components varied between 2.9 $\mu g/m^3$ and 14.5 $\mu g/m^3$ (mean 9.0 $\mu g/m^3 \pm 3.2$ $\mu g/m^3$). The lowest concentrations showed the greatest difference between mass concentration (estimated from DMPS) and the sum of analysed components, as shown in Figure 3. The mean ratio between these two measures was 0.84 and they were strongly correlated (Pearson $r = 0.93$). Hence the analysed chemical components accounted for almost the entire total mass of the PM_1 fraction. The lower mass concentration derived from DMPS was likely the result of a lower cut-off for particle size, compared with the PM sampler (840 nm vs 1 μm).

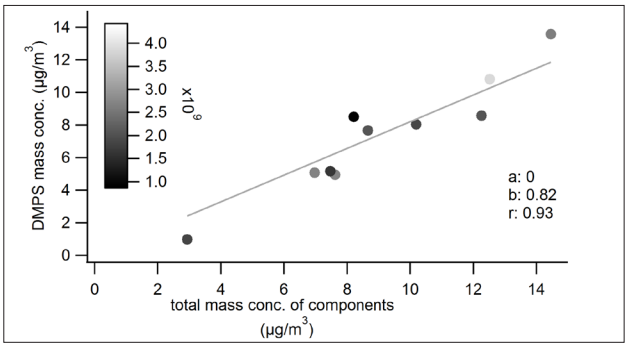


Figure 3: Mass concentrations of particles estimated from DMPS and the sum of the analysed components, greyscale-coded by total particle number concentration (#/m^3).

Figure 4 shows the PM_1 fractional concentrations of the analysed components for the combined as well as the individual spring and summer campaigns. The mean values are shown in Table 1. For the combined results (spring and summer), sulfate had the highest individual component concentration at 44% (3.9 $\mu g/m^3 \pm 2.3$ $\mu g/m^3$). POM was the second highest component at 39% (3.5 $\pm$ 0.6 $\mu g/m^3$). Two main cations, ammonium and potassium, contributed 11% and 2% respectively. The rest of the components together contributed approximately 4%.

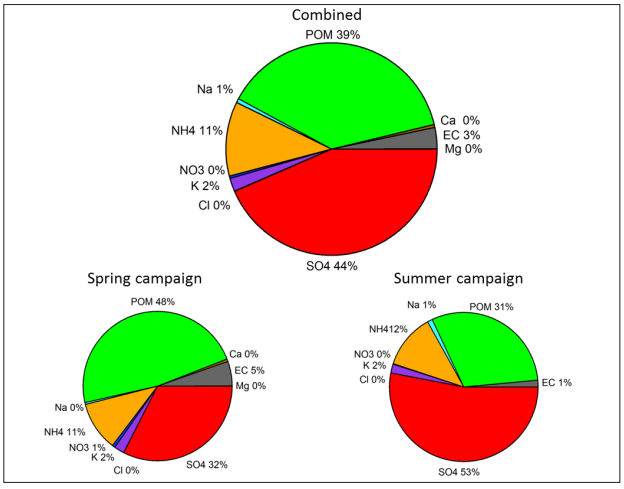


Figure 4: PM_1 fractional component composition for ambient samples: summer, spring, and combined campaigns.

Table 1: Concentrations of particles (μg/m^3) and gases (ppb) for spring, summer, all, and for the ambient case studies

	Spring	Summer	All	Case study 1	Case study 2	Case study 3
POM	3.63	3.25	3.46	2.21	3.28	3.47
SO_4^{2-}	2.48	5.65	3.92	0.20	2.82	7.48
$NH4^+$	0.81	1.27	1.02	0.04	0.86	1.76
EC	0.40	0.16	0.29	0.28	0.34	0.20
NO_3^-	0.05	0.01	0.03	0.04	0.04	0.01
K^+	0.17	0.20	0.18	0.03	0.16	0.36
Oxalate†	0.10	0.13	0.11	0.02	0.10	0.14
Ca^{2+}	0.04	N/A		0.09	0.03	N/A
Na^+	0.03	0.12	0.05	0.02	0.02	0.08
Mg^{2+}	0.003	N/A		0.011	0.001	N/A
Cl^-	0.004	BDL		0.005	0.003	BDL
Total#:	7.62	10.66	8.97	2.93	7.56	13.36
SO_2 (g)	1.16	0.64	0.93	0.05	1.36	0.52
NO_x (g)	2.70	1.94	2.37	2.11	2.66	1.62
CO (g)	144	94	123	93	146	92
O_3 (g)	36	36	36	29	35	30

Case Study 1, regional background; Case Study 2, Waterberg overpass; Case Study 3, Western BIC overpass. Biomass burning experiments were excluded.

†Oxalate is also included in POM analysis

#Total particles, excluding oxalate

N/A, not analysed; BDL, below determination limit; g, gaseous compounds

The equivalent ratio of sulfate to main cations for PM_1 indicated that sulfuric acid was not neutralised totally. The average contribution of sulfate in Botsalano was higher than the amounts measured in similar campaigns lasting approximately one year in various environments in Europe and China, and at Welgegund near Potchefstroom in South Africa.[15,30,31] However, at Welgegund a similar contribution has been noticed during the wet season, and similar findings have occurred in shorter campaigns at rural or remote sites in USA, Ireland and Japan.[15,32] According to aerosol mass spectrometer (AMS) studies, the overall average contribution of submicron sulfate and organic matter around the world are 32% (range 10% to 67%) and 45% (range 18% to 70%) respectively.[32]

A comparison of the results from the spring and summer campaigns showed that POM was the dominant component during spring, whereas sulfate dominated in summer. Sulfate concentrations were clearly higher during summer than spring, whereas the POM concentrations did not vary significantly between the two seasons (Table 1). It is well known that oxidation of SO_2 to sulfate occurs faster at higher relative humidity, as observed during summer, than in drier periods. The concentrations of SO_2 and the ratio of SO_2 to sulfate support this interpretation, with SO_2 clearly being lower during the summer campaign compared with spring. This finding indicates a faster removal of SO_2 in summer (Table 1).

Meteorological data for Botsalano, confirming the relative humidity levels, have previously been presented.[17] It has also been suggested that ozone (O_3) concentrations might be positively correlated with the sulfur conversion rate.[33] However, in our study O_3 concentration did not show a substantial difference between summer (range 17 ppb to 63 ppb) and spring (range 11 ppb to 58 ppb), although solar radiation in summer was on average 28% higher than in spring.

The EC level during the spring campaign was substantially higher than during the summer campaign (Table 1), which confirms the influence of combustion during spring. Although the concentration of levoglucosan, a biomass burning marker,[34] remained relatively low throughout our study, it was substantially higher during spring than summer; the mean values were 0.010 μg/m^3 in spring and 0.006 μg/m^3 in summer. In addition, CO, which can be used as a gaseous tracer of combustion,[35] showed similar trends, with means of 145 ppb and 92 ppb during the spring and summer campaigns respectively.

As previously mentioned, we found that the total concentration of the components analysed was a good approximation for the total PM_1 mass, as calculated from DMPS data. Additionally, PM_1 particle depositions on the filters were uniform, whereas the larger-size fractions were not. Because mass closure could not be performed for the $PM_{1-2.5}$ and $PM_{2.5-10}$ size fractions, we do not discuss these fractions individually in detail but only in comparison with the PM_1 fraction. Figure 5 shows the analysed components for the various size fractions for the combined spring and summer ambient air sampling campaigns. From these data it is obvious that SO_4^{2-}, POM, NH_4^+, EC and K^+ were predominantly found in the PM_1 size fraction, whereas NO_3^- occurred mostly in the coarser fractions. The size fraction analysis clearly indicates that the fine fraction (PM_1) is dominated by water-soluble chemical compounds. Soil components, which were not analysed in this study, are likely to comprise the majority of the coarser fractions.

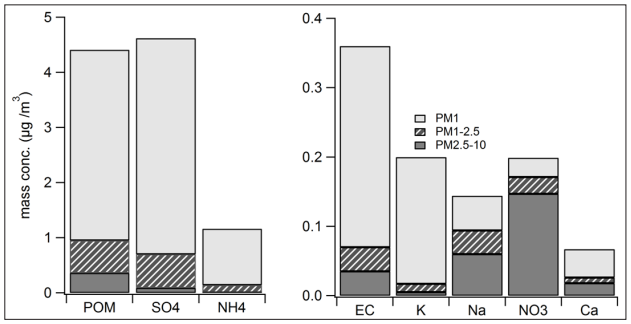


Figure 5: Comparison of PM_1, $PM_{1-2.5}$ and $PM_{2.5-10}$ compositional analysis for the combined ambient campaigns.

Ambient air case studies

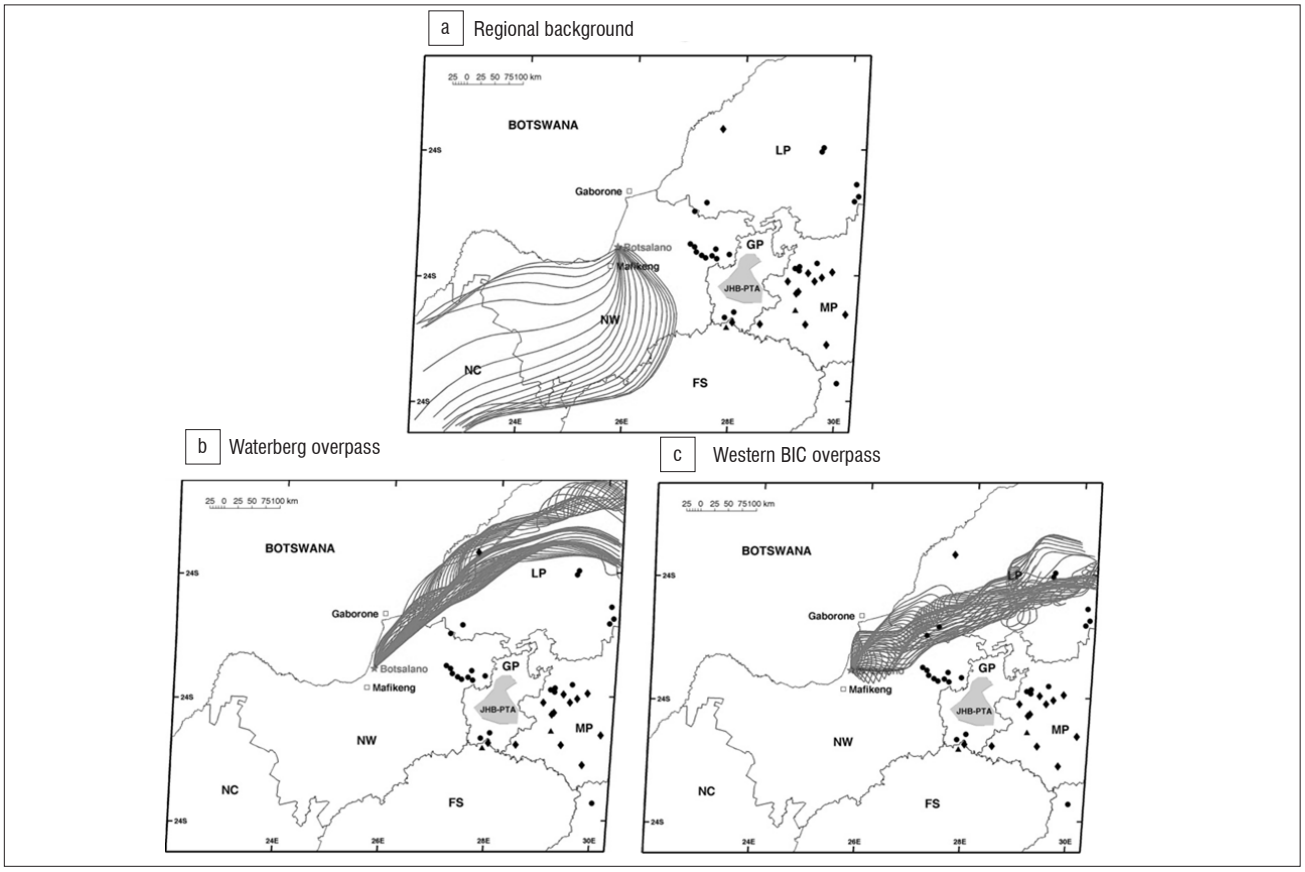
In the previous paragraphs we established that the analysed chemical compounds were mainly found in the fine fraction (PM_1). In addition, the mass balance could be calculated only for the PM_1 fraction. Therefore, the following case studies focus mainly on the PM_1 results.

Case study 1: Air masses that passed over the regional background

According to backward air mass trajectory analysis, there was one sampling day (9–10 October 2007) when the air masses arrived at Botsalano from the south-west. In this direction there are no large air pollution point sources, apart from the town of Mafikeng (population approximately 260 000), as indicated in Figure 6a.

Case study 2: Air masses that passed over the Waterberg

On three days during the sampling campaigns (11–12, 12–13 and 14–15 October 2007), air masses were likely to have been influenced by the coal-fired power station at Lephalale in the Waterberg, i.e. Matimba Power Station, without any influence from other large point sources (Figure 6b). The salience of this case study data is that at the time of writing, a second coal-fired power station – Medupi Power Station – had been commissioned for the Waterberg area. Matimba Power Station does not have equipment for removing sulfur and nitrogen gases from the emissions (de-SO_x and de-NO_x). By contrast, the newly-constructed Medupi Power Station will include such technology, although it is not yet clear if it will be implemented in the initial operations. The data in this case study could thus serve as a valuable reference for future studies in this region, to compare the concentrations of particles with and without NO_x and SO_x removal.



Key to provinces: LP, Limpopo; GP, Gauteng; MP, Mpumalanga; FS, Free State; NC, Northern Cape; NW, North West

Figure 6: Air mass history of identified ambient case studies: (a) regional background, (b) Waterberg overpass, (c) Western BIC overpass.

Case study 3: Air masses that passed over the western BIC

On two days during the sampling campaign (30 January–1 February 2008), the air masses arriving at Botsalano had passed directly over the two closest PGM smelters in the western BIC, situated east of Rustenburg (Figure 6c). These air masses were also likely to have travelled over the PGM and silicon smelters near Polokwane, but not over any other large air pollution point sources. The large air pollution point sources in the western BIC are mainly ferrochrome and PGM smelters as well as base metal refineries.[20] The sampled air masses therefore do not represent all types of point sources that occur in the western BIC, but they do represent the PGM smelters occurring in this region.

Comparison of case studies

In Table 1 the comparative results for the three case studies identified are listed. From these results it is clear that the background air masses we sampled were substantially cleaner than the anthropogenically-influenced air masses. This was despite a possible influence of pollution from Mafikeng on the background air masses sampled. The total concentrations of the components measured in the PM_1 fraction – which was also found to be a good approximation of the total PM_1 mass, as indicated earlier – were as follows: 2.9 $\mu g/m^3$ for the regional background, 7.6 $\mu g/m^3$ for the Waterberg overpass, and 13.4 $\mu g/m^3$ for the western BIC overpass.

The anthropogenic influence of the large point sources on the chemical composition of the PM_1 fraction is especially apparent when sulfate compositions are considered. The regional background had a sulfate concentration of a mere 0.2 $\mu g/m^3$, whereas air masses in the Waterberg overpass during spring and the Western BIC overpass during summer had levels of 2.8 $\mu g/m^3$ and 7.5 $\mu g/m^3$ respectively. During the Western BIC event, the sulfate-to-SO_2 ratio was higher than on average during the summer campaign, indicating faster oxidation of SO_2 or a longer transport time. The temperature and relative humidity during the event were similar to the averages, but the wind speed was lower than average. During the Waterberg event, the sulfate-to-SO_2 ratio was similar to the spring average.

Samples from biomass burning experiments

During the ambient air sampling campaigns, no clear biomass burning plume could be identified, hence no ambient biomass burning plume case study is presented. The spring sampling season in mid-October was well past the peak activity of regional biomass burning, so we did not expect to find biomass-burning influences in the air masses. The highest levoglucosan concentration was 0.030 $\mu g/m^3$, which was one order of magnitude lower than the background concentration (0.3 $\mu g/m^3$) measured during the SAFARI 2000 campaign. The SAFARI 2000 campaign was conducted during a period of large-scale regional biomass combustion.[35]

Biomass burning is an important air pollution source in southern Africa, with implications for both air quality and climate change. To augment the ambient data already presented in this paper, data from the biomass burning experiment we conducted are presented in this section.

A fractional distribution graph of the time-weighted average for the various components analysed in the two biomass burning plume samples gives an overall indication of the relative importance of the components in fresh biomass burning plumes. These data are presented in Figure 7.

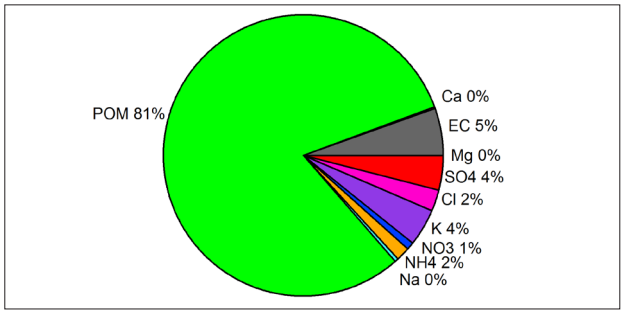


Figure 7: Time-weighted combined PM_1 fractional component composition of sampled biomass burning plumes.

The table shows that POM was much higher than the other components in the PM_1 fraction. We realise that some of the detected OC might have been gaseous OC, as the amount of gaseous OC was not measured. Previous studies have shown the amount of gaseous OC in ambient measurement to be between 6% and 19%.[9,27,37] In our study, approximately a quarter of the analysed OC was vaporised during the first temperature step of the TOA analysis (T=75–200 °C), implying that these compounds were the most volatile. Some of these components might therefore have been gaseous OC absorbed or adsorbed into the filter substrate, or onto the particles collected on it. Despite this possibility, the dominance of POM in the PM_1 fraction is evident.

The above finding is important, as it indicates that organic components released by biomass combustion during the dry season in southern Africa could play an important role during new particle formation and subsequent growth. This is in contrast to the wet season in southern Africa, when biogenic activities are mainly responsible for the release of organic components. The situation in southern Africa is therefore completely different from European and North American conditions, where biogenic activity is the main source of atmospheric organic compounds – at least in rural areas. The chemical and physical characteristics in the atmosphere are therefore likely to differ substantially across these continents.

The dominant contribution of POM to PM_1 aerosol from biomass burning smoke supports the findings presented in a recent paper.[38] The relevant study had focused on the effect of biomass burning plumes on the formation and growth of ambient atmospheric aerosols in southern Africa, based on long-term data.[38] Apart from the high (81%) POM contribution to PM_1 (Figure 7), further differences between biomass burning plume samples and other samples were the concentrations or the contributions of monosaccharide anhydrides, oxalate and potassium, which can be attributed as biomass burning tracers.[36] Their concentrations increased substantially during the biomass burning event, compared with the ambient samples. The concentration of submicron oxalate during the biomass burning event increased almost four-fold compared with the average of the ambient samples. However, the relative contribution of oxalate to POM (0.1%) was smaller than the oxalate contribution to all the ambient samples combined (2.7%). The contribution of K^+ in PM_1 during biomass burning (4% to 5%) was double that of the ambient samples (2%).

The ratios of monosaccharide anhydrides give information on the biomass burning material. The ratios of levoglucosan to mannosan and levoglucosan to galactosan in the first plume were 17.1 and 14.0 respectively; in the second plume, the ratios were 16.4 and 17.4 respectively. These ratios were quite similar to values measured in laboratory experiments in which savannah grass with acacia wood was burned (levoglucosan to mannosan ratio of 21.7, and levoglucosan to galactosan ratio of 15.2).[39] It has been reported that fresh smoke of savannah grass contains potassium chloride (KCl) particles, whereas aged smoke contains potassium sulfate (K_2SO_4) and potassium nitrate (KNO_3).[36,40] In our study, Cl^- was the most abundant anion (3%) in the first plume, followed by sulfate (2%) and nitrate (1%). These results indicate that most of the K^+ was probably present as KCl. The calculation of equivalent ratios showed an excess of K^+ compared with Cl^-, SO_4^{2-} and NO_3^-. However, if SO_4^{2-} was assumed to react first with NH_4^+, there was no excess sulfate left for K^+. The additional K^+ might therefore be present as carbonaceous material such as K_2CO_3, which was detected in an earlier biomass burning study in southern Africa.[41]

Because components analysed in the coarse fractions ($PM_{1\text{-}2.5}$ and $PM_{2.5\text{-}10}$) are unlikely to approach mass closure, we present our results only in a comparative manner to the PM_1 results (Figure 8). The dominance of the PM_1 fraction in the overall PM_{10} chemical composition in fresh biomass burning plumes is evident from these results. More than 90% of the POM, EC, K^+, Cl^-, SO_4^{2-} and NH_4^+ ammonium were in a PM_1 fraction. The NO_3^-, Na^+ and oxalate were mainly in the PM_1 fraction (62% to 83%), whereas at least 50% of Ca^{2+} was in the PM_1 fraction.

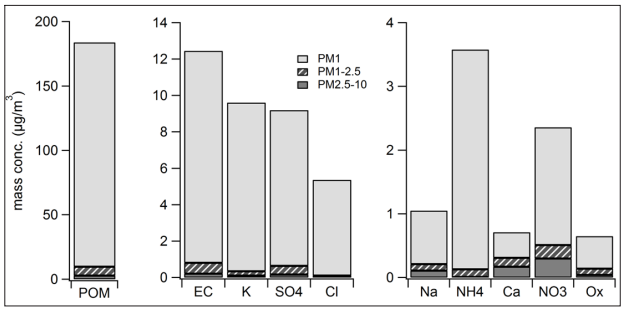


Figure 8: Comparison of PM_1, $PM_{1\text{-}2.5}$ and $PM_{2.5\text{-}10}$ compositional analysis for the combined time-weighted average biomass burning plumes sampled.

Conclusion

Size-segregated aerosol samples were collected using a three-stage impactor with backup filter at a relatively clean savannah environment at Botsalano Game Reserve in South Africa. The backup filter collected particles smaller than 1 μm. Other size ranges were >10 μm, 2.5–10 μm and 1–2.5 μm. Two campaigns were performed, one in spring (October 2007) and the other in summer (January–February 2008). In total throughout the entire study period, 11 sets of impactor samples were collected. In addition to these ambient measurements, a biomass burning experiment was performed to determine the chemical composition of fresh biomass burning fire plumes. Organic and elemental carbon, selected ions and monosaccharide anhydrides (e.g. levoglucosan) were analysed. These components essentially comprised the total mass of the PM_1 fraction.

Results of the ambient air mass analyses indicated that sulfate, organic carbon, ammonium, elemental carbon and potassium were mostly associated with fine particles (PM_1). Other components (sodium, chloride, nitrate and oxalate) were divided into fine and coarse fractions, but their concentrations were very small. POM was the dominant component during spring, whereas sulfate dominated in summer. Some indications of differing oxidation rates for SO_2 were found between the seasons, explaining the higher sulfate concentrations during the summer campaign. Substantially higher elemental carbon concentrations in spring indicated the occurrence of more combusting processes compared with the summer campaign.

Although Botsalano is a background site, it is regularly affected by air masses that have passed over several large point sources. The anthropogenic influence of these large point sources on the chemical composition of the PM_1 fraction was apparent, especially when the sulfate composition was considered. Air masses that passed over the coal-fired power station in Waterberg or the PMG smelter increased the sulfate concentration by 14 or 37 times compared with background air. Characterisation of air masses that had passed over the Waterberg area was especially important, because an additional large coal-fired power station was being commissioned at the time of writing (2015). The data presented here can be compared with the findings of similar studies, after the new coal-fired power station comes online. Comparisons can also be made between our data and those obtained after the de-SO_x and de-NO_x technologies of the new power station become operational. Such comparisons would help to assess the impact of large industrial developments in this region of South Africa.

The biomass burning study confirmed that potassium in fresh biomass plumes probably takes the form of KCl, for the most part, rather than K_2SO_4 and KNO_3. Clear differences were also found in the ratios of potassium and levoglucosan in the smouldering and flaming phases. Our data highlight the need for more comprehensive chamber-type experiments on the major fuel types, to confirm the ratio of important biomass burning tracer species. Such information can, in future, be used to better quantify the contribution of biomass burning in source apportionment studies.

Acknowledgements

The authors acknowledge the financial support received from the Academy of Finland under the project Air Pollution in Southern Africa (APSA) (project number 117505). We also thank the Head of the Botsalano Game Reserve, M. Khukhela, and the game reserve employees, for their kind and invaluable help during our measurements. In addition we are grateful to M. Jokinen and E. Sjöberg for their help in organising the project.

Authors' contributions

M.A. was the main investigator and wrote the primary manuscript; the study formed part of her PhD study, with S.S. as the degree supervisor. J.P.B revised the manuscript, V.V. and L.L. maintained the research station and helped to generate the data, P.v.Z. made conceptual contributions, and K.T. assisted with chemical analysis. All authors reviewed and approved the final manuscript.

References

1. Swap RJ, Annegarn HJ, Suttles JT, King MD, Platnick S, Privette JL, et al. Africa burning: A thematic analysis of the Southern African Regional Science Initiative (SAFARI 2000). J Geophys Res. 2003;108(D13):8465. http://dx.doi.org/10.1029/2003JD003747
2. Kulmala M, Asmi A, Lappalainen HK, Baltensperger U, Brenguier J-L, Facchini MC, et al. General overview: European Integrated project on Aerosol Cloud Climate and Air Quality interactions (EUCAARI) – integrating aerosol research from nano to global scales. Atmos Chem Phys. 2011;11:13061–13143. http://dx.doi.org/10.5194/acp-11-13061-2011
3. Laakso L, Merikanto J, Vakkari V, Laakso H, Kulmala M, Molefe M, et al. Boundary layer nucleation as a source of new CCN in savannah environment. Atmos Chem Phys. 2013;13:1957–1972. http://dx.doi.org/10.5194/acp-13-1957-2013
4. Laakso L, Vakkari V, Virkkula A, Laakso H, Backman J, Kulmala M, et al. South African EUCAARI measurements: seasonal variation of trace gases and aerosol optical properties. Atmos Chem Phys. 2012;12:1847–1864. http://dx.doi.org/10.5194/acp-12-1847-2012
5. Lourens AS, Beukes JP, Van Zyl PG, Fourie GD, Burger JW, Pienaar JJ, et al. Spatial and temporal assessment of gaseous pollutants in the Highveld of South Africa. S Afr J Sci. 2011;107(1/2):269–276. http://dx.doi.org/10.4102/sajs.v107i1/2.269
6. Martins J, Dhammapala R, Lauchmann G, Galy-Lacaux C, Pienaar J. Long-term measurements of sulphur dioxide, nitrogen dioxide, ammonia, nitric acid and ozone in southern Africa using passive samplers. S Afr J Sci. 2007;103:336–342.
7. Hirsikko A, Vakkari V, Tiitta P, Manninen HE, Gagné S, Laakso H, et al. Characterisation of sub-micron particle number concentrations and formation events in the western Bushveld Igneous Complex, South Africa. Atmos Chem Phys. 2012;12:3951–3967. http://dx.doi.org/10.5194/acp-12-3951-2012
8. Hirsikko A. Vakkari V, Tiitta P, Hatakka J, Kerminen VM, Sundström AM, et al. Multiple daytime nucleation events in semi-clean savannah and industrial environments in South Africa: analysis based on observations. Atmos Chem Phys. 2013;13:5523–5532. http://dx.doi.org/10.5194/acp-13-5523-2013
9. Mkoma S, Chi X, Maenhaut W. Characterization of carbonaceous materials in PM2.5 and PM10 size fractions in Morogoro, Tanzania, during 2006 wet season campaign. Nucl Instrum Methods Phys Res B. 2010;B268:1665–1670. http://dx.doi.org/10.1016/j.nimb.2010.03.001
10. Mkoma S, Wang W, Maenhaut W. Seasonal variation of water-soluble inorganic species in the coarse and fine atmospheric aerosols at Dar es Salaam, Tanzania. Nucl Instrum Methods Phys Res B. 2009;B267:2897–2902. http://dx.doi.org/10.1016/j.nimb.2009.06.099
11. Mkoma S, Maenhaut W, Chi X, Wang W, Raes N. Characterisation of PM10 atmospheric aerosols for the wet season 2005 at two sites in East Africa. Atmos Environ. 2009;43:631–639. http://dx.doi.org/10.1016/j.atmosenv.2008.10.008
12. Mkoma S, Maenhaut W, Chi X, Wang W, Raes N. Chemical composition and mass closure for PM10 aerosols during the 2005 dry season at a rural site in Morogoro, Tanzania. Xray Spectrom. 2009;38:293–300. http://dx.doi.org/10.1002/xrs.1179
13. Gaita SM, Boman J, Gatari MJ, Pettersson JBC, Janhäll S. Source apportionment and seasonal variation of PM2.5 in a Sub-Saharan African city: Nairobi, Kenya. Atmos Chem Phys. 2014;14:9977–9991. http://dx.doi.org/10.5194/acp-14-9977-2014
14. Maenhaut W, Schwarz J, Cafmeyer J, Annegarn H. Study of elemental mass size distributions at Skukuza, South Africa, during the SAFARI 2000 dry season campaign. Nucl Instrum Methods Phys Res B. 2002;189:254–258. http://dx.doi.org/10.1016/S0168-583X(01)01053-9
15. Tiitta P, Vakkari V, Croteau P, Beukes JP, Van Zyl PK, Josipovic M, et al. Chemical composition, main sources and temporal variability of PM_1 aerosols in southern African grassland. Atmos Chem Phys. 2014;14:1909–1927. http://dx.doi.org/10.5194/acp-14-1909-2014
16. Booyens W, Van Zyl PG, Beukes JP, Ruiz-Jimenez J, Kopperi M, Riekkola M-L, et al. Size-resolved characterisation of organic compounds in atmospheric aerosols collected at Welgegund, South Africa, J Atmos Chem. 2015;72:43–64. http://dx.doi.org/10.1007/s10874-015-9304-6
17. Laakso L, Laakso H, Aalto PP, Keronen P, Petäjä T, Nieminen T, et al. Basic characteristics of atmospheric particles, trace gases and meteorology in a relatively clean southern African savannah environment. Atmos Chem Phys. 2008;8:4823–4839. http://dx.doi.org/10.5194/acp-8-4823-2008
18. Vakkari V, Beukes JP, Laakso H, Mabaso D, Pienaar JJ, Kulmala M, et al. Long-term observations of aerosol size distributions in semi-clean and polluted savannah in South Africa. Atmos Chem Phys. 2013;13:1751–1770. http://dx.doi.org/10.5194/acp-13-1751-2013
19. Vakkari V, Laakso H, Kulmala M, Laaksonen A, Mabaso D, Molefe M, et al. New particle formation events in semi-clean South African savannah. Atmos Chem Phys. 2011;11:3333–3346. http://dx.doi.org/10.5194/acp-11-3333-2011
20. Hirsikko A, Vakkari V, Tiitta P, Manninen HE, Gagné S, Laakso H, et al. Characterisation of sub-micron particle number concentrations and formation events in the western Bushveld Igneous Complex, South Africa. Atmos Chem Phys. 2012;12:3951–3967. http://dx.doi.org/10.5194/acp-12-3951-2012
21. Venter AD, Vakkari V, Beukes JP, Van Zyl PG, Laakso H, Mabaso D, et al. An air quality assessment in the industrialised western Bushveld Igneous Complex, South Africa. S Afr J Sci. 2012;108(9/10):1059–1068. http://dx.doi.org/10.4102/sajs.v108i9/10.1059
22. Lourens ASM, Butler TM, Beukes JP, Van Zyl PG, Beirle S, Wagner TK, et al. Re-evaluating the NO hotspot over the South African Highveld. S Afr J Sci. 2012;108(11/12):1146–1152. http://dx.doi.org/10.4102/sajs.v108i11/12.1146
23. Cavalli F, Viana M, Yttri KE, Genberg J, Putaud J-P. Toward a standardised thermal-optical protocol for measuring atmospheric organic and elemental carbon: the EUSAAR protocol. Atmos Meas Tech. 2013;3:79–89.
24. Jaffrezo J-L, Aymoz G, Cozic J. Size distribution of EC and OC in the aerosol of Alpine valleys during summer and winter. Atmos Chem Phys. 2005;5:2915–2925. http://dx.doi.org/10.5194/acp-5-2915-2005
25. Turpin B, Lim H-J. Species Contributions to PM2.5 mass concentrations: revisiting common assumptions for estimating organic mass. Aerosol Sci Technol. 2001;35:602–610. http://dx.doi.org/10.1080/02786820119445
26. Saarnio K, Teinilä K, Aurela M, Timonen H, Hillamo R. High-performance anion-exchange chromatography–mass spectrometry method for determination of levoglucosan, mannosan, and galactosan in atmospheric fine particulate matter. Anal Bioanal Chem. 2010;398:2253–2264. http://dx.doi.org/10.1007/s00216-010-4151-4
27. Saarnio K, Aurela M, Timonen H, Saarikoski S, Teinilä K, Mäkelä T, et al. Chemical composition of fine particles in fresh smoke plumes from boreal wild-land fires in Europe. Sci Total Environ. 2010;408:2527–2542. http://dx.doi.org/10.1016/j.scitotenv.2010.03.010
28. Draxler RR, Rolph GD. 2003 HYSPLIT (HYbrid Single-Particle Lagrangian Integrated Trajectory) Model. Available from: http://www.arl.noaa.gov/ready/hysplit4.html.
29. Tyson PD, Garstang M, Swap R. Large-scale recirculation of air over Southern Africa. J Appl Meteorol. 1996;35:2218–223. http://dx.doi.org/10.1175/1520-0450(1996)035<2218:LSROAO>2.0.CO;2

30. Putaud JP, Raes F, Van Dingenen R, Brüggemann E, Facchini MC, Decesari S, et al. A European aerosol phenomenology – 2: Chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. Atmos Environ. 2004;38:2579–2595. http://dx.doi.org/10.1016/j.atmosenv.2004.01.041

31. He K, Zhao Q, Ma Y, Duan F, Yang F, Shi Z, Chen G. Spatial and seasonal variability of $PM_{2.5}$ acidity at two Chinese megacities: Insights into the formation of secondary inorganic aerosols. Atmos Chem Phys. 2012;12:1377–1395. http://dx.doi.org/10.5194/acp-12-1377-2012

32. Zhang Q, Jimenez JL, Canagaratna MR, Allan JD, Coe H, Ulbrich I, et al. Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced northern hemisphere midlatitudes. Geophys Res Lett. 2007;34:L13801. http://dx.doi.org/10.1029/2007GL029979

33. Khoder MI. Atmospheric conversion of sulfur dioxide to particulate sulfate and nitrogen dioxide to particulate nitrate and gaseous nitric acid in an urban area. Chemosphere. 2002;49:675–684. http://dx.doi.org/10.1016/S0045-6535(02)00391-0

34. Simoneit BRT, Schauer JJ, Nolte CG, Oros DR, Elias VO, Fraser MP, et al. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. Atmos Environ. 1999;33:173–82. http://dx.doi.org/10.1016/S1352-2310(98)00145-9

35. Cabada J, Pandis S, Subramanian R, Robinson A, Polidori A, Turpin B. Estimating the Secondary Organic Aerosol Contribution to PM2.5 Using the EC Tracer Method. Aerosol Sci Tech. 2004;38(S1):140–155. http://dx.doi.org/10.1080/02786820390229084

36. Gao S, Hegg DA, Hobbs PV, Kirchstetter TW, Magi BI, Sadilek M. Water-soluble organic components in aerosols associated with savanna fires in southern Africa: Identification, evolution, and distribution. J Geophys Res. 2003;108(D13): 8491. http://dx.doi.org/10.1029/2002JD002324

37. Sillanpää M, Frey A, Hillamo R., Pennanen AS, Salonen RO. Organic, elemental and inorganic carbon in particulate matter of six urban environments in Europe. Atmos Chem Phys. 2005;5:2869–2879. http://dx.doi.org/10.5194/acp-5-2869-2005

38. Vakkari V, Kerminen V-M, Beukes JP, Tiitta P, Van Zyl P, Josipovic M, et al. Rapid changes in biomass burning aerosols by atmospheric oxidation. Geophys Res Lett. 2014;41:2644–2651. http://dx.doi.org/10.1002/2014GL059396

39. Iinuma Y, Brüggemann E, Gnauk T, Müller K, Andreae MO, Helas G, et al. Source characterization of biomass burning particles: The combustion of selected European conifers, African hardwood, savanna grass, and German and Indonesian peat. J Geophys Res. 2007;112:D08209. http://dx.doi.org/10.1029/2006JD007120

40. Li J, Pósfai M, Hobbs P, Buseck P. Individual aerosol particles from biomass burning in southern Africa: 2, Compositions and aging of inorganic particles. J Geophys Res. 2003;108(D13):8484. http://dx.doi.org/10.1029/2002JD002310

41. Pósfai M, Simonics R, Li J, Hobbs PV, Buseck PR. Individual aerosol particles from biomass burning in southern Africa: 1. Compositions and size distributions of carbonaceous particles. J Geophys Res. 2003;108(D13):8483. http://dx.doi.org/10.1029/2002JD002291

The early history of research funding in South Africa: From the Research Grant Board to the FRD

AUTHORS:
Ndivhuwo M. Luruli[1]
Johann Mouton[1,2]

AFFILIATIONS:
[1]Centre for Research on Evaluation, Science and Technology (CREST), Stellenbosch University, Stellenbosch, South Africa
[2]DST–NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy, Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Johann Mouton

EMAIL:
jm6@sun.ac.za

POSTAL ADDRESS:
Centre for Research on Evaluation, Science and Technology (CREST), Private Bag X1, Matieland 7600, South Africa

DATES:
Received: 10 Mar. 2015
Revised: 16 Nov. 2015
Accepted: 16 Dec. 2015

KEYWORDS:
Centre for Science Development; differentiated funding; Foundation for Research Development; National Research Council and Board; National Research Foundation

HOW TO CITE:
Luruli NM, Mouton J. The early history of research funding in South Africa: From the Research Grant Board to the FRD. S Afr J Sci. 2016;112(5/6), Art. #2015-0097, 6 pages. http://dx.doi.org/10.17159/sajs.2016/20150097

The South African government has a long tradition of supporting research at public higher education institutions. Such support commenced in the early 20th century, although the exact nature of the support at that time is poorly documented. The oldest research funding model in the country was agency funding, which started as early as 1911 through the Royal Society of South Africa. A few years later, in 1918, a more coordinated funding body called the Research Grant Board (RGB) was established in the Union of South Africa. The RGB offered competitive funding to individual academics in the natural and physical sciences. The human sciences were only supported much later with the establishment of the Council for Educational and Social Research in 1929. Here we review the history of research funding in South Africa, with a special focus on the work of the RGB between 1918 and 1938.

Introduction

Many people assume that a limited number of models exist for the public funding of modern research systems. At the highest level, a distinction is made between 'core' and 'project' funding of research conducted at public research organisations and universities. Core funding (also referred to as 'block funding') for universities is usually channelled through a Ministry of Education or Higher Education. The term 'core funding' refers to state support for the core business of universities and other public research bodies, which is usually understood to comprise teaching and learning, research, and community engagement. There are basically two ways in which core funding to universities is calculated, namely *formula-based* and *performance-based* (or some combination of the two).

Formula-based core funding consists of calculating the core funds to be allocated to a specific university on the basis of an agreed-upon formula. The formula usually takes into account student numbers, growth in student numbers, staff numbers, infrastructure, and so on. By contrast, performance-based core funding is based on the past performance of a university. In the field of research, performance is usually linked to the research output of the university. In the field of teaching and learning, performance could involve any number of 'measures' such as student completion rates, student throughput rates, and absolute numbers of graduates and post-graduate students. It is not uncommon to have a system of core funding that consists of both formula funding and performance funding.

South African universities receive an annual core funding amount that is calculated in terms of students, staff, and infrastructure, as well as performance-based funding. Performance-based funding was introduced in 1985 and revised in 2003 and 2015, and it rewards the most research-productive universities.

Project funding, which involves supporting research projects at public research organisations, can be channelled either directly or through an agent. Direct channelling is not the norm in most countries. Agents that channel research funding could be a research funding council, such as the Economic and Social Research Council in the UK, or a foundation, such as the National Science Foundation in the USA. Foundations are usually accountable to a Ministry of Science and Technology but also sometimes to a Ministry of Higher Education. Project funding is often referred to as *competitive* funding, as such funds are usually disbursed on the basis of open competition, even where certain priority areas are designated or 'ring-fenced'. The process involves calls for proposals, subsequent peer-review, and monitoring of project deliverables and outcomes.

Funding and scientific systems have evolved and transformed over time. Historical trajectories and changes in the political and social climate create shifting spaces in which funding councils and scientific systems need to function. Context is vital to the functioning of these funding bodies and funding bodies need to adapt to survive. According to Rip, 'Funding agencies, with their aggregation machines, function in a particular historical context and translate contextual changes…' into their functioning.[1]

A review of the literature shows that there exists a clear consensus regarding the definition and main functions of science granting councils. Science granting councils or agencies serve as intermediary, quasi-public institutions that are positioned between the state and individuals or institutions that perform research.[1] The primary purpose of research councils, traditionally, has been to 'organise part of the funding relationship between government and universities as a peer-review based competition for project funding'.[2] The councils are 'expected to mediate the political and policy interests in scientific research into the world of science and technology and promote the interests of science and technology in the policy world'.[2] For example, Lepori et al.[3] consider a funding agency to be the body that disburses grants, irrespective of the origins of the funds. These agencies operate in an intermediary position between the knowledge production system and state policy, and between state and academy.

Research funding councils can be seen as a link in a chain of principal–agent relationships. The government acts as principal to the research council, and the research council as principal to the scientists. A research council would be both agent (in relation to the government) and principal (in relation to the scientists) at once. In simple terms, research councils are positioned both as agents of state funders and societal interests, with their task being to deliver the goods, as well as being the principal with respect to individual research providers and scientists.

Our focus in this paper is on the early history, specifically the institutionalisation, of the agency funding model in South Africa.

Early history of science funding in South Africa

Scientific activities have been taking place in South Africa from as early as the 18th century. In the early years, they were somewhat unregulated despite the existence of prominent institutions such as the Royal Society of South Africa and the South African Museum.[4] The Royal Society of South Africa was founded in 1877[5] as the South African Philosophical Society,[4] and received formal status through a Royal Charter in 1908 signed by King Edward VII.[5] It was also in 1908 that the Onderstepoort Veterinary Research Laboratory was built through generous government funding of GBP80 000.[6] A few years earlier, in 1903, the South African Association for the Advancement of Science, known as the S2A3, had been established as the regulatory body for all scientific activities in the country.[4]

In addition to research in veterinary sciences, other significant research activities that took place in the Union of South Africa during these early years included geological research by J.P. Johnson and herpetology research by G.A. Boulenger.[7] During this period, public funding of research in South Africa was not institutionalised, although some funding for research was available through donations made by prominent individuals, or in some cases by institutions such as the South African Literary and Scientific Institutions.[4]

Because of a perceived lack of co-ordinated research funding, the then President of the Royal Society of South Africa, Mr H.H. Hough, wrote to the Prime Minister of the Union of South Africa on 1 July 1910, requesting that the Society be recognised as a research agency.[8] In his letter, Mr Hough stated that:

> *The Royal Society of South Africa desires to draw the attention of the Union Government to the importance of considering at the present time the best means of promoting methodological scientific research, this being an agency on which, as is well known, so much of the material and moral welfare of a country depends. In the past, unfortunately, there has been no continuity in any such efforts made in our country, with the result that no really adequate return has been obtained for the money thus spasmodically spent.*

This plea was followed by a grant award of GBP500 from the Ministry of Education, through a budget vote, to the Royal Society of South Africa. The award was aimed at research support for the year 1911.[9] In what can be considered the first case of government funding for research in South Africa, the Royal Society awarded five grants totalling GBP250. In the following year (1912), the Society funded another six projects totalling GBP275.[10]

The Royal Society was faced with budget cuts between 1914 and 1916, when the Department of Education reduced its annual budget, first from GBP500 to GBP300 and then to GBP50.[5,11] This prompted a delegation from the Society to pay a visit to the Minister of Mines on 23 May 1917, as the mandate of providing funding for research had by then been transferred from the Ministry of Education to the Ministry of Mines. The delegation lobbied for the reinstatement of the original grant of GBP500 and was led by Dr L. Peringuey, secretary of the Royal Society of South Africa. Following their request, government agreed to have the grant to the Society increased in 1917 to GBP300.[12]

While the Royal Society of South Africa battled with a decreased budget and continued to negotiate for an increase over the following years,[13] discussions were taking place within government for the establishment of a national research funding body. This body would later be called the Research Grant Board. However, the fact that the Royal Society disbursed research grants to the universities on behalf of the government for the period 1911–1917 means it can be seen as the first 'research agency' in South Africa. Nonetheless, this function would always be seen as additional to the Society's main mission as an academy of science. It is therefore not surprising that by 1916–17, government was ready to establish a new body that would assume the role of a research agency.

The Research Grant Board

The history of the Research Grant Board (RGB) dates back to 1916, when the Industries Advisory Board was established on 13 October 1916.[14] At the first meeting of the Industries Advisory Board, held in Pretoria on 18 October 1916, the functions of the Board were explained to the eleven members appointed to the Board. The Board was required to deal with statistics of production, scientific and industrial research, factory legislation, encouragement of industries, development and utilisation of natural resources, and paper manufacture.[14,15]

The founding members of the Industries Advisory Board were all industrialists, and the Board reported to the Ministry of Mines and Industries. Membership was however extended in 1917 to include individuals with scientific and technical skills from the Scientific and Technical Committee. In 1918, the Minister of Mines and Industries approved a proposal to merge the Industries Advisory Board and the Scientific and Technical Committee.[16,17] The two bodies had argued that their consolidation would lead to better coordination of activities. The new institution that resulted from the merger was called the Advisory Board of Industry and Science. The Advisory Board of Industry and Science, within its first year of existence, recommended to the government that it should form a Research Grant Board, which would be based within the Department of Education.[18]

The RGB was subsequently established in October 1918 as a subcommittee of the Advisory Board of Industry and Science, reporting to both the Minister of Education and the Minister of Mines and Industries. In addition to advising the government on issues of research at universities and museums, the RGB was given the mandate to manage all research grants allocated to universities from government funds.[13] This step effectively inaugurated the institutionalisation of research agency funding in South Africa. On instruction by the Minister of Education, the RGB also, during the 1920–21 financial year, took over the research funding components of both the Royal Society of South Africa and the South African Association for the Advancement of Science.[19,20]

Funding through the Research Grant Board

The RGB provided government research grants to university-based researchers, mainly those researchers who were permanent residents of the Union.[21] Prominent scientists such as Dr Basil Schonland, Dr Meiring Naudé, and Dr J.L.B. Smith were among the individuals who benefited from support by the RGB.[17] Over the years the RGB supported research in a variety of topics and disciplines. Examples of projects funded in 1919 include:

- Bushman and other native studies (A.M. Duggan-Cronin)
- Relative values of locomotive smoke box-char and various wood-charcoals as fuel for suction gas engines (W.S.H. Cleghorne)
- Flat worm parasites in South African wild and domestic animals and a survey of the trematodes in all classes, vertebrates and invertebrates, of South African animals (C.S. Grobbelaar).

Perhaps not surprisingly, the majority of projects supported through the RGB were in the natural sciences. The social sciences did not have a dedicated source of funding until 1929, when the National Bureau of Educational Research (NBER) was established under the Department of Education.[22] The broad social sciences field was only represented on the RGB through the inclusion of members with Arts and Humanities background in 1920.[23] Smit reported that because the NBER was established during an economic crisis in South Africa, some of its functions were compromised.[24]

In 1934, the mandate of the NBER was broadened to include the social sciences, and in line with this addition the name of the institution was changed to Council for Educational and Social Research.[22] Later the name was changed again, to the National Bureau for Educational and Social Research.[24] The initial funding administered by the Council for Educational and Social Research was obtained from the Carnegie Corporation of New York.[18]

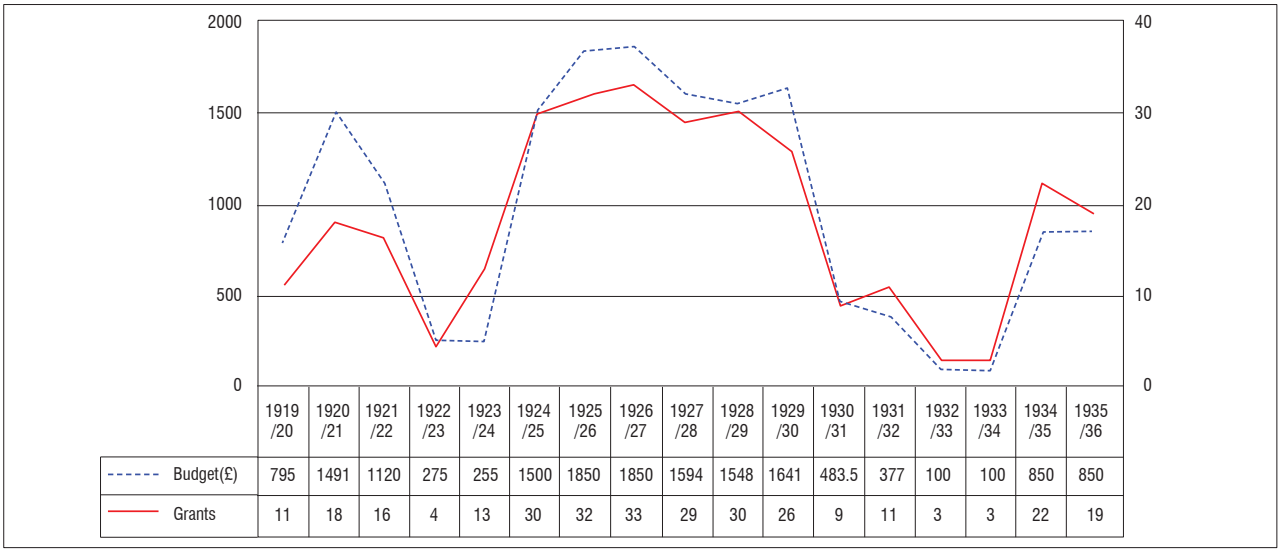


Figure 1: Number of RGB grants (1919–1935) and annual budget.

Key: Number of government research grants, 1919–1935 (solid line, right axis); budget in GBP allocated each year (dashed line, left axis)

Modern research councils or foundations typically utilise a wide array of funding instruments ('funding categories') to execute their mission. The introduction and administration of such instruments usually respond to the specific needs and policy imperatives of the research system. This is also true of the current National Research Foundation (NRF), which offers a range of fairly standard instruments such as postgraduate scholarships, grants for emerging scholars and established scholars, research chairs, and centres of excellence. The NRF also has instruments aimed at steering the research system to achieve national goals, including grants for 'women in science', Thutuka, and formerly the Technology and Human Resources for Industry Programme (THRIP).

Even in its early years, the RGB differentiated between a number of research funding instruments:

- Carnegie Research Grants (of New York), and Carnegie Travelling Fellowships, started in 1928
- University Research Grants and University Research Scholarships, started in 1934
- Mineral Research Scholarships, started in 1935 (and managed by the Director of the Mineral Research Laboratory at the University of Witwatersrand).[21]

Thus, in addition to funding received from government, the RGB administered research funding entrusted to it by the Carnegie Corporation of New York, which made available to it an allocation of GBP10 000 for the period 1928–1932, with a further USD30 000 for 1933–1937.[21] The fact that the RGB acted as an 'agent' of the Carnegie Corporation of New York is noteworthy, as it signifies the trust already accorded to the RGB, but is not unusual in itself. It has become a common feature of national research funding councils to become the conduit for various funding sources, including international funds.

Although the RGB operated under the auspices of the Advisory Board of Industry and Science, it enjoyed a substantial degree of independence. When the Advisory Board of Industry and Science was dissolved in 1923,[25] the RGB became a separate body aligned only to the Department of Mines and Industries. The RGB was ultimately transferred to the Department of Commerce and Industries in 1933.[21]

Between 1919 and 1936, the RGB supported 309 projects totalling an investment of over GBP16 000. The highest number of projects funded within a single financial year was 33 projects, during the 1926–27 financial year (see Figure 1). There was also variation in the average grant amounts, as shown in Table 1. For the most part, there was great variation in the number of funded projects and the amount of funding awarded from one year to the next. Fluctuation in funding between years was solely the result of the varying parliamentary vote or budget that was allocated annually to the relevant department towards the functions of the RGB. This unpredictable method of funding, with a different amount each year, was found to be ineffective by the RGB and resulted in uncertainty about the future of the Board. The RGB wanted the Government of the Union to implement a funding mechanism similar to that of Carnegie research grants, where funding was awarded for a five-year period.[21]

Table 1: Number of grants, total awarded, and average amount per grant (1919–1935)

Year of funding	Number of grants awarded	Total value of grants (in GBP)	Average grant value (in GBP)
1919–20	11	795	72
1920–21	18	1491	83
1921–22	16	1120	70
1922–23	4	275	69
1923–24	13	255	20
1924–25	30	1500	50
1925–26	32	1850	58
1926–27	33	1850	56
1927–28	29	1594	55
1928–29	30	1548	52
1929–30	26	1641.1	63
1930–31	9	483.5	54
1931–32	11	377	34
1932–33	3	100	33
1933–34	3	100	33
1934–35	22	850	39
1935–36	19	850	45
Total	**309**	**16679.6**	**54**

Source: NASA[21]

The Carnegie funding for the RGB was over and above the funding received from the South African government. During its initial funding period, the Carnegie Corporation awarded a substantial amount of GBP10 000 for the period 1928–1932, followed by an award of USD30 000 for 1933–1937.[21] Sue Krige[26] highlights that funding by the Carnegie Corporation of New York played a significant role in developing and extending research in South Africa between the 1920s and the 1950s.

From RGB to the National Research Council and Board

During the mid-1930s, proposals were submitted to advocate for the establishment of a new institution, a National Research Council that would replace the RGB. One of the proposals was addressed to Jan Hofmeyr, then Minister of Education, by Professor M.M. Rindl, then president of the South African Association for the Advancement of Science. The proposal suggested that

> *… the new Council should incorporate the functions of the Research Grant Board, and that the moneys administered at present by the Research Grant Board be transferred to the general income of the National Research Council.*[27]

Two years later, the Department of Mines issued a memorandum supporting the proposal to establish a National Research Board and a National Research Council to replace the RGB.[28] The memorandum suggested that the proposed institution should be placed within the Department of Education, and would thus be removed from the Department of Commerce and Industries, where the RGB was placed. This proposal was motivated by the fact that the scope of the RGB had grown over the years, such that it was no longer appropriately placed within the Department of Commerce and Industries.

The growth in the RGB's scope resulted from the extension of funding responsibilities to include support not only for universities and museums, but also for other institutions that conducted research, and in general all areas of knowledge production. Furthermore, in 1923 when the Union of South Africa joined the International Research Council, later known as the International Council of Scientific Union (ICSU), the RGB took on the responsibility of managing this affiliation.[21] Other reasons for the reorganisation of the RGB were that the constitution needed to be changed, and there was a need for better coordination of research activities by different government departments. Furthermore, the departments concerned had expressed the view that in future, 'more stable financial provision should be made'.[28]

A committee was convened to lead the restructuring process. When the process was complete, recommendations were made and submitted to the Minister of Education. Some of the main recommendations were as follows:

> a. *The present Research Grant Board shall cease to function at 31st March, 1938; and in its place there shall be set up a National Research Council [and a National Research Board]. These bodies shall function under the Minister of Education, and*
>
> b. *The functions of these bodies shall correspond to those at present exercised by the Research Grant Board.*[29]

The RGB was reorganised in 1938 to form a 'larger and more representative body', and was subsequently replaced by two institutions, namely the National Research Board and the National Research Council.[30] The National Research Board took over the administrative duties of the RGB, and the National Research Council became an advisory body to the Minister of Education, offering advice on ways to improve research in South Africa.[31] These two institutions were collectively referred to as the National Research Council and Board (NRC&B), and were officially inaugurated on 25 July 1938.[29] In his inaugural speech, Jan Hofmeyr referred to the NRC&B as the

> *South African Parliament of Research – its primary function being to consider measures for the improvement of the research position in the Union, and to suggest directions along which research is desirable.*[29]

Despite the achievements of the RGB and its successors over the years, a high level of dissatisfaction with the status of research in the Union of South Africa remained, mostly among individuals who were in charge of research development, i.e. those who were part of the NRC&B. For the most part, their dissatisfaction centred on the lack of coordination of research activities and the lack of collaboration among researchers. The NRC&B was only in existence for few years before talks began calling for further change to the shape of the research institution. Among the suggestions for a new format was that the Union should possess an institution similar to the National Research Council of Canada. Early discussions also focused greatly on the calibre of the individual who would be in charge of managing the institution. It was highlighted that

> *… in this connection, the Council recognizes that the success or failure of the whole scheme, when established, will depend in great measure on the Executive Officer and that consequently every effort should be made to secure a man with the qualities indicated.*[32]

The right person for this job was also described as

> *… a man of high scientific attainments who is at the same time energetic, tactful and experienced in negotiations … and his mental horizon should be wide enough for him to take a statesman's view of researches in such diverse fields as, let us say, social anthropology and geophysics.*[32]

The post-war beginnings of the current research funding system

The NRC&B was reorganised at the end of the Second World War, in 1945, to form the Council for Scientific and Industrial Research (CSIR). Dr Basil Schonland was its first Chief Executive Officer. The CSIR took over part of the functions of the NRC&B (supporting research in the fields of industry and natural science). Those functions that fell under the scope of the social sciences were transferred, in 1946, to a new institution named the National Council for Social Research (NCSR).[22] The NCSR absorbed the responsibilities of the National Bureau of Educational and Social Research[18] in addition to those that were transferred from the NRC&B.

The CSIR was established under the Scientific Research Council Act 33 of 1945 (published in *Government Gazette* 3514 on 22 June 1945). It had a twofold mandate[33,34]:

- First, to conduct scientific and industrial research in its own laboratories, to complement research done at universities.
- Second, to support, through the provision of funding, research conducted at universities throughout the country.

To fulfil its dual mandate, the CSIR received a grant allocation from the Department of National Education (through Parliament). Funding for university research took the form of the CSIR awarding grants to academic staff and bursaries to students. With regard to its own onsite research, the CSIR started out with three laboratories: the National Physical Laboratory, the National Chemical Research Laboratory, and the National Building Research Institute.[35] The first head of the National Physical Laboratory was Dr Meiring Naudé, who later succeeded J.P. Du Toit to become the third president of the CSIR in 1952 (until 1971).[36]

The support and development of research at universities started during the first year of the CSIR's existence. In this regard, Dr Schonland developed university research grants to provide funding for academics and students alike. Research grants were managed under the University Research Division (URD), which supported research of the

scientist's own free choice or self-initiated research.[37] During its first year of funding, there was little demand for this kind of support, with only GBP16 526 being requested from a total budget of GBP27 800. However, the demand for funding increased over the years. In 1962, for example, the CSIR received requests of up to ZAR537 338 from a budget of ZAR299 754.[37] In the mid-1970s, the URD became the Research Grants Division (RGD) and started supporting researchers at museums and technikons as well as at universities.[38]

The CSIR also established several discipline-based research units. The first was the Medical Research Unit, established in the 1950s. By the mid-1960s, nine research units had been established. The research units were headed by eminent researchers and were based at various universities and research institutes.

Alongside the RGD, the CSIR introduced the Co-operative Scientific Programmes (CSP) in 1975, initially referred to as the National Scientific Programmes. The aim of the CSPs was 'to identify problems peculiar to South Africa which, because of their magnitude and complexity, required the co-ordinated effort of a number of different organizations in planned research programmes'.[37] The CSPs therefore supported projects aimed at addressing problems of national importance through multi-disciplinary research.

In 1984, the Research Grants Division and the CSPs were combined to form the Foundation for Research Development (FRD).[37,39] The mandate of this new body was 'the provision of appropriate human resources in science and technology to meet the requirements of the national economy'.[40] The FRD officially became a funding agency of the CSIR on 1 April 1984, and later became the main research support programme within the CSIR.

In 1990, the FRD was awarded autonomous status through the Research Development Act (Act No. 75 of 1990). The Act identified the mandate of the FRD as research development, which included not only providing financial support to higher education institutions and museums, but also managing some expensive national facilities. The latter group was made up of the National Accelerator Centre (NAC), now iThemba Labs; the South African Astronomical Observatory (SAAO); the Hartebeeshoek Radio Astronomical Observatory (HartRAO); and the Hermanus Magnetic Observatory (HMO), now SANSA Space Science. The FRD thus became the largest research support agency in the country during the 1990s, although it supported only the natural sciences and engineering.

In 1999, FRD was merged with the Centre for Science Development (CSD), its counterpart from the Human Sciences Research Council (HSRC). This merger resulted in the formation of the National Research Foundation (NRF).

Concluding remarks

South Africa has come a long way in research development, with government showing its commitment through financial support. Furthermore, the institutionalisation of research funding through an agency, consolidated with the establishment of the RGB in 1918, laid a solid foundation and ensured structured support for research in the country. The placement of the funding agency was an important consideration during the early years, as the institution needed to be strategically placed to meet the needs of the entire research sector. The RGB and its predecessors faced several challenges over the years, including constant budget cuts that left the agency ineffective in supporting research. The high demand for research funding therefore started much earlier than is generally assumed, although it has intensified in recent years. A grant from an external source, the Carnegie Corporation of New York, allowed the RGB to support more research and provided greater stability in the system – which had been lacking with government funding alone.

Another milestone during the early years was the introduction in 1929 of a dedicated funding institution for the social sciences, through the National Bureau of Educational Research. This move marked the beginning of differentiation among funding instruments for various scientific fields, which continued under the CSIR and its successors (the FRD, CSD, and NRF).

Many of the funding principles established under the RGB were carried through over the next decades. These principles included, for example, the allocation of funding based on the outcome of a peer-review process, and utilisation of a diversity of funding instruments. Such principles continue to form an integral part of resource allocation by funding agencies in South Africa, almost a century later.

Authors' contributions

This publication is a result of research done by N.M.L. towards her PhD; J.M. was her supervisor.

References

1. Rip A. Higher forms of nonsense. Eur Rev. 2000;8(4):467–486. http://dx.doi.org/10.1017/S1062798700005032
2. Van der Meulen B. New roles and strategies of a research council: Intermediation of the principal-agent relationship. Sci Public Policy. 2003;30(5):323–336. http://dx.doi.org/10.3152/147154303781780344
3. Lepori B, Van den Besselaar P, Dinges M, Potì B, Reale E, Slipersæter S, et al. Comparing the evolution of national research policies: What patterns of change? Sci Public Policy. 2007;34(6):372–388. http://dx.doi.org/10.3152/030234207X234578
4. Dubow S. A commonwealth of knowledge: Science, sensibility and white South Africa 1820–2000. Oxford: Oxford University Press; 2006.
5. Royal Society of South Africa. Brief history [homepage on the Internet]. No date [cited 2009 Oct 12]. Available from: http://www.royalsocietysa.org.za/?page_id=122
6. Plug C. Southern African science in the year 1908. S Afr J Sci. 2008;104(1/2):25–26.
7. Plug C. Southern African science in the year 1910-100^{n}. S Afr J Sci. 2010;106(1/2):11–12.
8. National Archives of South Africa (NASA). Letter of 1st July 1910 addressed to the Honourable Prime Minister from the Royal Society of South Africa. [Reference: E18/1]. Pretoria: NASA; 1910.
9. National Archives of South Africa (NASA). Letter of 6th November 1911 addressed to Dr Peringuey from the Minister of Education. [Reference: E18/1]. Pretoria: NASA; 1911.
10. National Archives of South Africa (NASA). Royal Society of South Africa. Applications for grants-in-aid, for the year 1912. [Reference: E18/1]. Pretoria: NASA; 1912.
11. National Archives of South Africa (NASA). Notes of interview Deputation of the Royal Society and Minister of Mines, 23rd May 1917. [Reference: E18/1]. Pretoria: NASA; 1917.
12. National Archives of South Africa (NASA). Letter of 14th June 1917 addressed to M. Hofmeyr, Cape Town. [Reference: E18/1]. Pretoria: NASA; 1917.
13. National Archives of South Africa (NASA). Letter of April 1917 addressed to M. Hofmeyr, Cape Town. [Reference: E18/1]. Pretoria: NASA; 1918.
14. National Archives of South Africa (NASA). Industries Advisory Board. Report, 1916–1917. [Reference: B61/1]. Pretoria: NASA; 1917.
15. Luruli NM. Research funding and modes of knowledge production: A comparison between NRF-funded and industry-funded researchers in South Africa [unpublished PhD thesis]. Stellenbosch: Stellenbosch University; 2014.
16. National Archives of South Africa (NASA). Letter of 5th October 1918 addressed to the Secretary, Scientific and Technical Committee, Johannesburg. [Reference: MM3063/18]. Pretoria: NASA; 1918.
17. National Archives of South Africa (NASA). Letter of 10th October 1918 addressed to the Chief Conservator of Forests, Pretoria. [Reference: MM3063/18]. Pretoria: NASA; 1918.
18. National Archives of South Africa (NASA). Report of the Research Grant Board, 1926. [Reference: MM611/26]. Pretoria: NASA; 1927.
19. National Archives of South Africa (NASA). Letter of 11th February 1919 addressed to the secretary of the Research Grant Board. [Reference: A668]. Pretoria: NASA; 1919.
20. National Archives of South Africa (NASA). Letter of 28th February 1919 addressed to the secretary for Mines and Industries. [Reference: S11/1/1]. Pretoria: NASA; 1919.

21. National Archives of South Africa (NASA). Report of the Research Grant Board, 1918–1935. [Reference: LA213]. Pretoria: NASA; 1936.

22. Human Sciences Research Council (HSRC). General Information, Report No IN–9. Pretoria: HSRC; 1971.

23. National Archives of South Africa (NASA). First annual report of the Research Grant Board. October 1918–December 1919. [Reference: MISC 13]. Pretoria: NASA; 1920.

24. Smit AP. Staatsondersteunde natuur- en geesteswetenskaplike navorsing in Suid-Afrika; 1918–1969. 'N voorgeskiedenis van die WNNR en die RGN. Pretoria: Raad vir Geesteswetenskaplike Navorsing; 1984.

25. National Archives of South Africa (NASA). Industrial advisers – dissolution of the Board. [Reference: A668] Pretoria: NASA; 1923.

26. Krige S. Manuscript on the history of the evaluation/rating system in the granting of funds for research, 1918 to 1996. 2008 (Unpublished).

27. National Archives of South Africa (NASA). Report of the National Research Council and Board. [Reference: F8/209]. Pretoria: NASA; 1941.

28. National Archives of South Africa (NASA). A plea for the establishment of a National Research Council and for the initiation of a National Research Policy in South Africa. [Reference: R3276/2]. Pretoria: NASA; 1935.

29. National Archives of South Africa (NASA). A memorandum to the Department of Mines, on the proposed establishment of a National Research Board and a National Research Council. [Reference: F8/209]. Pretoria: NASA; 1937.

30. National Archives of South Africa (NASA). Address delivered at the inaugural meeting of the National Research Council by Robert B. Young. [Reference: R3276/3]. Pretoria: NASA; 1938.

31. National Archives of South Africa (NASA). Address delivered at the inaugural meeting of the National Research Council by the Honorable J.H. Hofmeyr. 1938 [Reference: R3276/3]. Pretoria: NASA; 1938.

32. National Archives of South Africa (NASA). Report on the consideration by the National Research Council of proposals for the re-organisation of research activities in the Union of South Africa. [Reference: R3276/2]. Pretoria: NASA; 1942.

33. Boshoff SC, Grebe E, Mouton J, Ravat E, Ravjee N, Waast R. Science in South Africa: History, institutions and statistics. Stellenbosch: Centre for Interdisciplinary Studies, Stellenbosch University; 2000.

34. Marais HC. Perspectives on science policy in South Africa. Pretoria: Network Publishers; 2000.

35. Austin B. Schonland: Scientist and soldier. Johannesburg: Witwatersrand University Press; 2001. http://dx.doi.org/10.1201/9781420033571

36. Wagener P. The formation of the SAIP [http://www.saip.org.za]. Pretoria: South African Institute of Physics; c2005 [cited 2009 April 24]. Available from http://www.saip.org.za/documents/History_of_the_SAIP.pdf.

37. Kingwill DG. The CSIR – the first 40 years. Pretoria: CSIR; 1990.

38. Garbers CF. Amptelike opening SNO-Gebou – 4 April 1989. 1989.

39. De Wet JS. Investigation into the award system for research support in the natural sciences and engineering at universities, museums and technikons. Research undertaken by Prof Jack S de Wet, research grants adviser to CSIR executive. Pretoria: FRD; 1987.

40. Foundation for Research Development (FRD). Science policy in South Africa – government funding of academic and related research. Pretoria: FRD; 1991.

AUTHORS:
Gaokgatlhe Tawane[1]
Daniel García-Martínez[1,2,3]
Jennifer Eyre[4,5]
Markus Bastir[1,2]
Lee Berger[1]
Peter Schmid[1,7]
Shahed Nalla[1,6]
Scott A. Williams[1,4,5]

AFFILIATIONS:
[1]Evolutionary Studies Institute and Centre for Excellence in Palaeosciences, University of the Witwatersrand, Johannesburg, South Africa
[2]Palaeoanthropology Group, National Museum of Natural Sciences (MNCN-CSIC), Madrid, Spain
[3]Biology Department, Faculty of Sciences, Autonomous University of Madrid, Madrid, Spain
[4]Center for the Study of Human Origins, Department of Anthropology, New York University, New York, New York, USA
[5]New York Consortium in Evolutionary Primatology, New York, New York, USA
[6]Department of Human Anatomy and Physiology, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa
[7]Anthropological Institute and Museum, University of Zurich, Zurich, Switzerland

CORRESPONDENCE TO:
Daniel García-Martínez

EMAIL:
dan.garcia@mncn.csic.es

POSTAL ADDRESS:
Palaeoanthropology Group, Museo Nacional de Ciencias Naturales (MNCN-CSIC) – Paleobiology, José Gutierrez Abascal 6, 28006 Madrid, Spain

DATES:
Received: 29 July 2015
Revised: 11 Dec. 2015
Accepted: 08 Jan. 2016

KEYWORDS:
upper thorax; rib cage; *Australopithecus africanus*; geometric morphometrics; Gauteng Province

HOW TO CITE:
Tawane G, García-Martínez D, Eyre J, Bastir M, Berger L, Nalla S, et al. A hominin first rib discovered at the Sterkfontein Caves, South Africa. S Afr J Sci. 2016;112(5/6), Art. #2015-0278, 7 pages. http://dx.doi.org/10.17159/sajs.2016/20150278

A hominin first rib discovered at the Sterkfontein Caves, South Africa

First ribs – the first or most superior ribs in the thorax – are rare in the hominin fossil record, and when found, have the potential to provide information regarding the upper thorax shape of extinct hominins. Here, we describe a partial first rib from Member 4 of the Sterkfontein Caves, South Africa. The rib shaft is broken away, so only the head and neck are preserved. The rib is small, falling closest to small-bodied *Australopithecus* first ribs (AL 288-1 and MH1). Given that it was recovered near the StW 318 femur excavation, which also represents a small individual, we suggest that the two may be associated. Three-dimensional geometric morphometric analyses were used to quantify the rib fragment morphology and compare it to extant hominoid and other fossil hominin ribs. While only the proximal end is preserved, our analyses show that South African *Australopithecus* share derived features of the proximal first rib more closely resembling *A. afarensis* and later hominins than great apes.

Introduction

Ribs are rare in the human fossil record because of their delicate structure; this scarcity is particularly pronounced in the early hominin fossil record. Moreover, those that have been discovered are usually fragmentary and distorted, and their interpretation is difficult because of the complicated morphological quantification of their 3D curvature.[1-5] In spite of the scarcity of these remains, it is important to note that, among ribs, the first rib is generally the best preserved in the fossil record as a result of its unique morphology relative to the rest of the ribs in the thorax.[6,7] The first rib is also important because it bears diagnostic features informative of upper thorax morphology.[7-9] For example, the modern human first rib usually has a univertebral articulation (with the T1 vertebra), as do those of early hominins, whereas the other great apes have a bivertebral articulation (with the C7 and T1 vertebrae).[10]

In early hominins, first ribs are known from *Australopithecus afarensis* (AL 288-1ax)[11], *Homo erectus* (KNM-WT 15000 AG and AY&AZ)[12], *A. sediba* (MH1, UW88-148; MH2, UW88-198 and UW88-187)[8] and *H. naledi* (UW101-83)[13]. *A. afarensis* AL 288-1 and *A. sediba* (MH1 and MH2) were hypothesised as presenting a narrow upper thorax, as extant apes do, and this thoracic shape is probably related to suspensory locomotor behaviour.[8,11,14,15] This narrow upper thorax has also been observed in the recently published small-bodied species *H. naledi.*[13] In contrast, other early hominin specimens, such as *H. erectus* (KNM-WT 15000) and the large-bodied *A. afarensis* (KSD-VP-1/1), have been proposed to have modern human-like, expanded upper thoraces.[12,16]

Sterkfontein Caves (located 40 km from Johannesburg) – one of the most important South African sites regarding the quantity and quality of the fossils discovered[17-20] – has yielded several costal remains, but a first rib has not been reported from this site to date. The *A. africanus* partial skeleton Sts 14, discovered at Sterkfontein in 1947, is associated with a number of ribs.[21] Robinson's analysis positioned the preserved ribs toward the lower thorax. According to Robinson[21], these ribs are smaller than those of modern *H. sapiens* but are characterised by a similar degree of curvature. The features observed on the ribs suggested a modern human-like lower thorax shape, but there is still some degree of uncertainty about the upper thorax shape of *A. africanus*.

In 1987, another partial skeleton, StW 431, this time of a larger-bodied, presumed male *A. africanus* was discovered at Sterkfontein in Bed B of Member 4.[22] StW 431 preserves a right rib with the head, neck and tubercle, which was assessed as a probable third rib.[22] However, the preservation of this fragment was not sufficient to discuss upper thorax morphology in *A. africanus* from an evolutionary point of view.

Here we report, present and discuss, in a comparative anatomical context, a well-preserved proximal part of a first rib previously recovered at Sterkfontein but neither identified nor published, as well as its provenance and its relation to the previously discovered remains from the Sterkfontein Caves.

Materials and methods

The new fossil, a proximal fragment of a first rib from the right side (StW 670; Figure 1) was discovered at Sterkfontein site (Cradle of Humankind, Gauteng, South Africa) in grid square V 46, at 21 feet 1 inches to 22 feet 4 inches (27'2–28'2), Member 4[23] (Figure 2). The date of the discovery of the rib is unknown. The collection from Sterkfontein Caves housed at the University of the Witwatersrand dates from 1966 to present excavations. Material collected from Sterkfontein prior to 1966 is housed at Ditsong Museum (formerly known as the Transvaal Museum).

Preservation, morphology and ontogenetic assessment of StW 670

A detailed description of the preservation status as well as the morphology was carried out on StW 670 based on the principal anatomical features of the first rib. Frequencies of single/double articular facet/s of rib head in the comparative sample were also studied (Supplementary table 1 of the supplementary material).

The maturation state of the first rib was evaluated based on the epiphyseal fusion of the articular tubercle of the rib and the articular facet of the rib head.[24] However, it should be noted that the maturation rate of *Australopithecus* epiphyseal fusion could differ from that of *H. sapiens*.

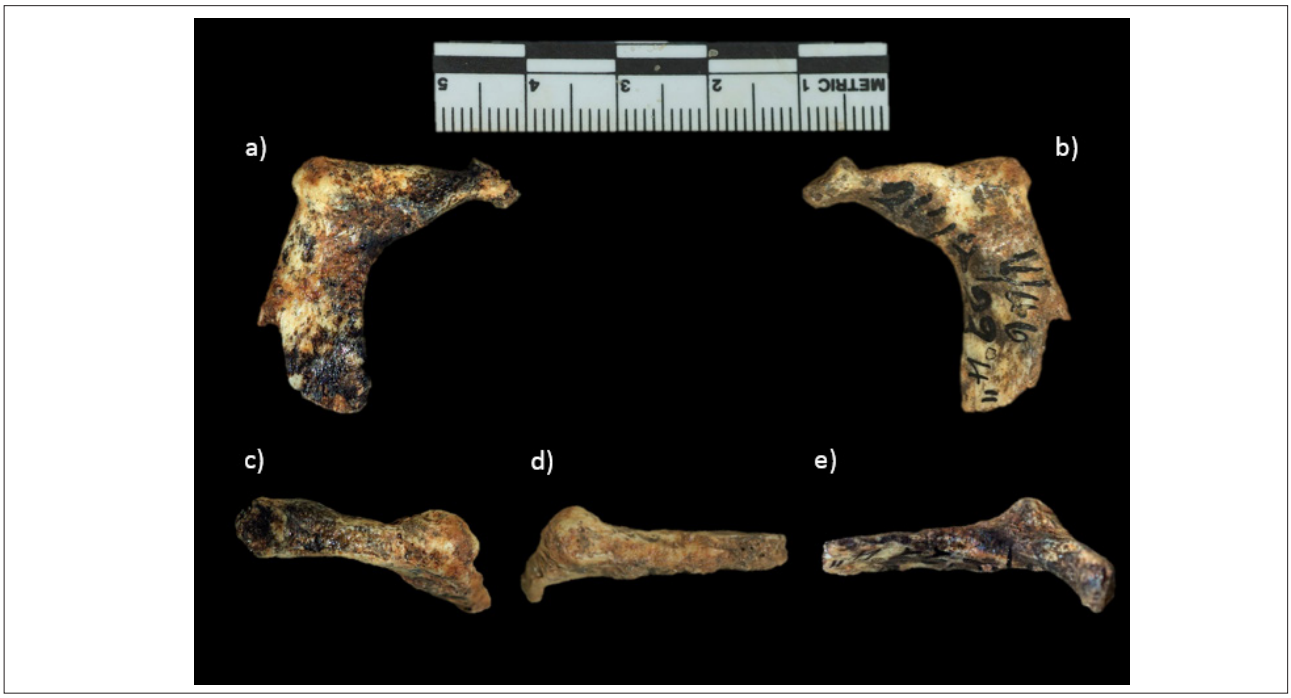


Figure 1: Rib StW 670 in different views: (a) superior, (b) inferior, (c) posterior, (d) exterior and (e) interior.

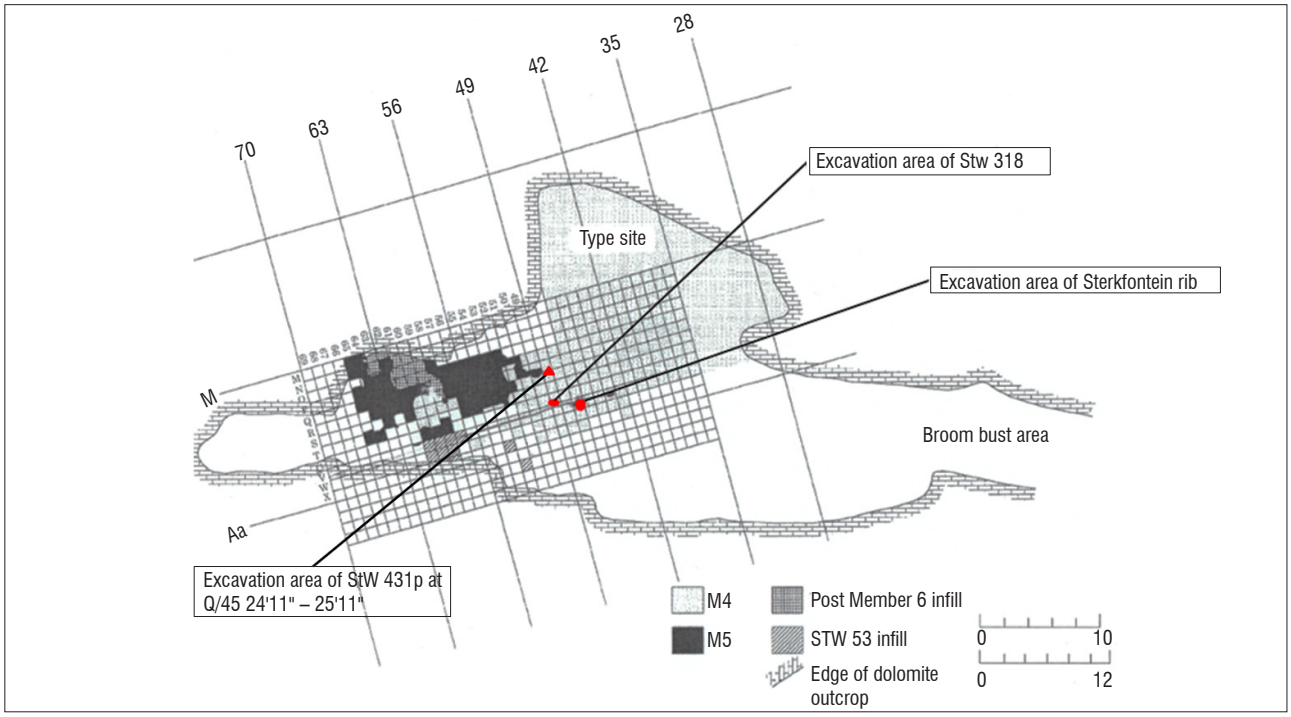


Figure 2: Stratigraphy of Sterkfontein Members 4 and 5, indicating the locality of excavation of StW 670 and StW 318.

Linear measurements

Linear measurements taken on StW 670 were carried out on the original remains. The linear measurements taken were: head height (superoinferior diameter of the rib head), head width (anteroposterior diameter of the rib head), neck length (distance from the medial-most point on the rib head to the medial-most point on the articular facet of the tubercle) and neck width (taken midway between the rib head and articular tubercle in anteroposterior dimension).[2,9,25] Each measurement was calculated from the average of three measurements in order to reduce intra-observer error.[1,2,25] The measurements are given in Table 1.

Table 1: Linear measurements taken on StW 670

	Head height (mm)	Head width (mm)	Neck length (mm)	Neck width (mm)
StW 670	6.4	6.7	18.8	5.5

Measurements on StW 670 were made with standard anthropometric instruments following the definitions used in previous studies.[2,9,25]

Geometric morphometric analyses

For 3D geometric morphometrics, 3D high-resolution laser scans of original fossils – StW 670, *A. sediba* MH1 (UW88-148) and MH2 (UW88-198) – and of a high-quality cast of AL 288-1ax *(A. afarensis)* were used. These scans were obtained using a NextEngine 3D laser scanner in 'wide mode' (with a resolution of 0.38 mm and an accuracy of 6 points per mm). Additionally, computerised tomography (CT) scans of the original remains of *H. erectus* KNM-WT 15000 (AG and AY&AZ) were employed for the analyses (Supplementary table 1).

For the 3D geometric morphometrics analyses, remains of 33 *H. sapiens* (20 Europeans, 10 sub-Saharan Africans, 1 small-bodied member of the San population and 2 small-bodied individuals from the Andaman Islands) were analysed (Supplementary table 1). These scans were also obtained through a NextEngine 3D laser scanner in 'wide mode' (with a resolution of 0.38 mm and an accuracy of 6 points per mm). Additionally, 8 ribs of *Pan troglodytes* (chimpanzee), 3 ribs of *Gorilla gorilla* (gorilla), 3 ribs of *Pongo* sp. (orangutan) and 3 ribs of *Hylobates* sp. (gibbon) from Kyoto University Primate Research Institute (www.pri.kyoto-u.ac.jp) were used to represent non-hominin morphological pattern (Supplementary table 1). Technical specifications of CT scanning can be obtained from www.pri.kyoto-u.ac.jp.

Six 3D type 1 landmarks were collected on each rib, quantifying morphological information of the rib head and the articular tubercle as well as the thickness of the rib neck (Figure 3). Landmarks were collected at the rib head on the most lateral, most medial, most caudal and most cranial points of the articular facet/s of the rib head and at the most lateral point of the articular tubercle. Additionally, one landmark was located at the point defined by the shortest distance from the most lateral point of the articular tubercle to the internal margin of the rib curvature ('d' distance in Figure 3). Because StW 670 is missing a large part of the body of the rib and the sternal end, no landmarks were taken on this area of the comparative sample.

Size was measured as centroid size.[26,27] The size relations between the Sterkfontein first rib and the rest of the individuals in our sample were explored using a box plot. These analyses were carried out in PAST software version 3.[28] In order to study the shape relations along the variability of the sample, we superimposed the landmark configurations using generalised Procrustes analysis.[26,27] The superimposed coordinates were then submitted to principal component analysis (PCA) in shape space in order to reduce the dimensionality of data and to visualise the main axes of variation.

Ordinations were computed into MorphoJ 1.05f software[29] and the shape differences of the surface associated with PC1 and PC2 axes were warped and visualised using the EVAN Toolkit software version 1.62 (www.evan-society.org/). A box plot of Procrustes distances was used to compare and explore shape relations from the individual under study (StW 670) to the different groups of the sample. These analyses were carried out in the Virtual Morphology Lab of the National Museum of Natural Sciences (Madrid, Spain).

Results

Preservation, morphology and ontogenetic assessment of StW 670

There is sufficient preservation of features in the first rib fragment (StW 670) to make description, comparison and analysis possible, even though the fossil is incomplete. The fragment has a well-preserved head, neck, tubercle and most of the proximal part of the body, although the rest of the body is lacking distally.

The rib fragment StW 670 was determined as a first rib from the right side. Siding was determined because the head is positioned at an angle caudal to the tubercle and shaft which produces a mediolateral slope of the neck. Moreover, as the caudal surface of the shaft is smoother than the cranial surface, this feature also confirms the side of the rib as a first rib from the right side.

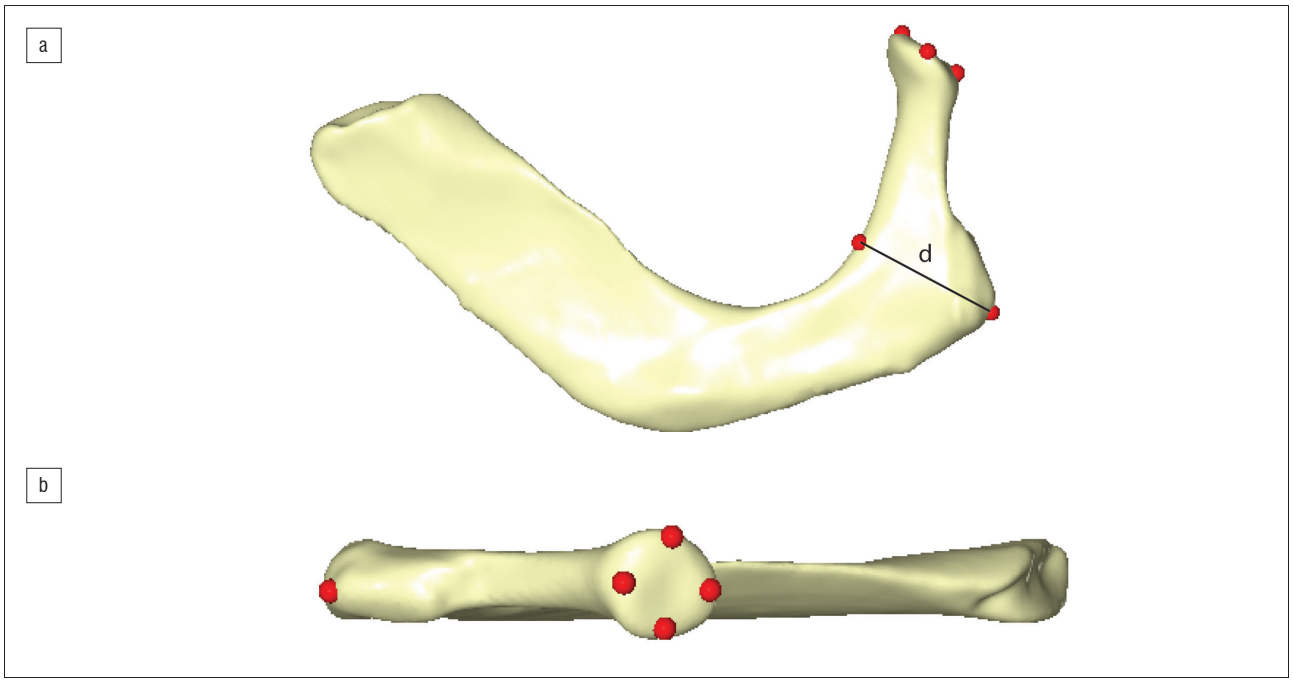


Figure 3: Anatomical location of the six landmarks type 1 taken on the sample. In (a), landmarks at the rib head on the most lateral, most medial and most cranial points of the articular facet/s of the rib head as well as at the most lateral point of the articular tubercle and the landmark defined by the distance 'd' are observed. In (b), the four landmarks at the rib head as well as the one at the tubercle are observed. Measurement 'd' is defined by the shortest distance from the most lateral point of the articular tubercle to the internal margin, which is important to calculate the landmark defined by this measurement.

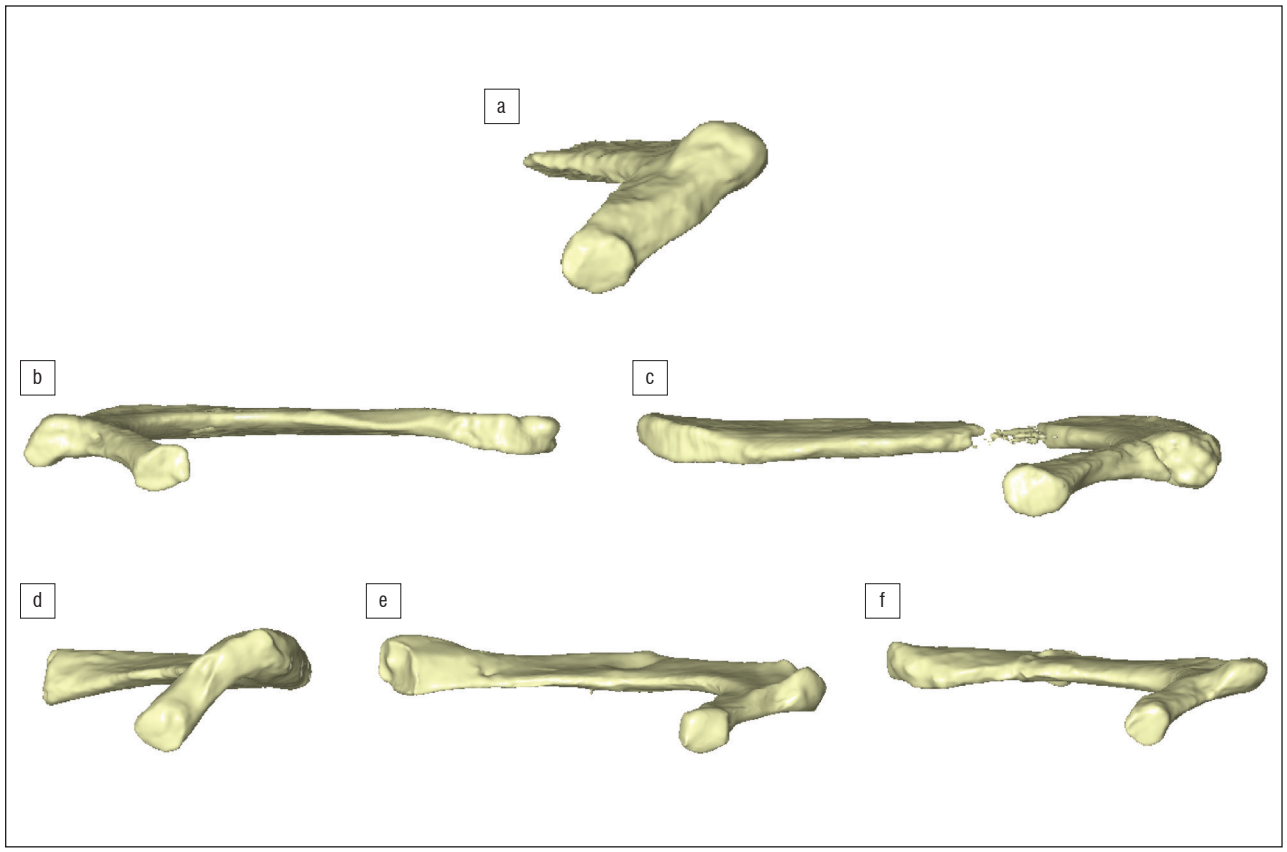


Figure 4: Detail of the single articular facet at the rib head of every fossil specimen studied in the sample: (a) StW 670, (b) KNM-WT 15000 left, (c) KNM-WT 15000 right, (d) AL 288-1, (e) MH1 and (f) MH2. As the goal of this image is to display the single facets, every specimen has been scaled to the same size.

The head has a rim that is thickened for the attachment of the costovertebral ligament. The angle and the tubercle coincide – a feature that is presented in extant and fossilised primate first ribs. The neck is narrow proximally and widens distally as the tubercle is approached. The neck is also rounded proximally and flattened distally. The tubercle is well developed and has a well-defined, smooth and rounded articular facet. The non-articular part of the tubercle for the attachment of the lateral costotransverse ligament is present.

The head has a single articular facet, which is rounded – a morphology that most closely resembles modern humans. The rest of the fossil specimens studied also presented a single articular facet (Figure 4). Regarding the comparative sample, 95% of the European modern humans presented a single articular facet at the rib head for articulation with T1, while 100% of the African modern humans presented this feature. In contrast, all of the non-hominin hominoids (*Pan*, *Gorilla*, *Pongo* and *Hylobates*) presented a double facet for articulation with C7 and T1.

The preservation of the fragment does not allow for the finer detail of anterior scalene and anterior serrate muscle insertions elevations, but it does allow us to study the scalene medium muscle insertion, which is only slightly marked in StW 670. The epiphysis at the articular tubercle and the rib head are well preserved and totally fused with the metaphysis, suggesting that the individual was adult at the time of death.[24]

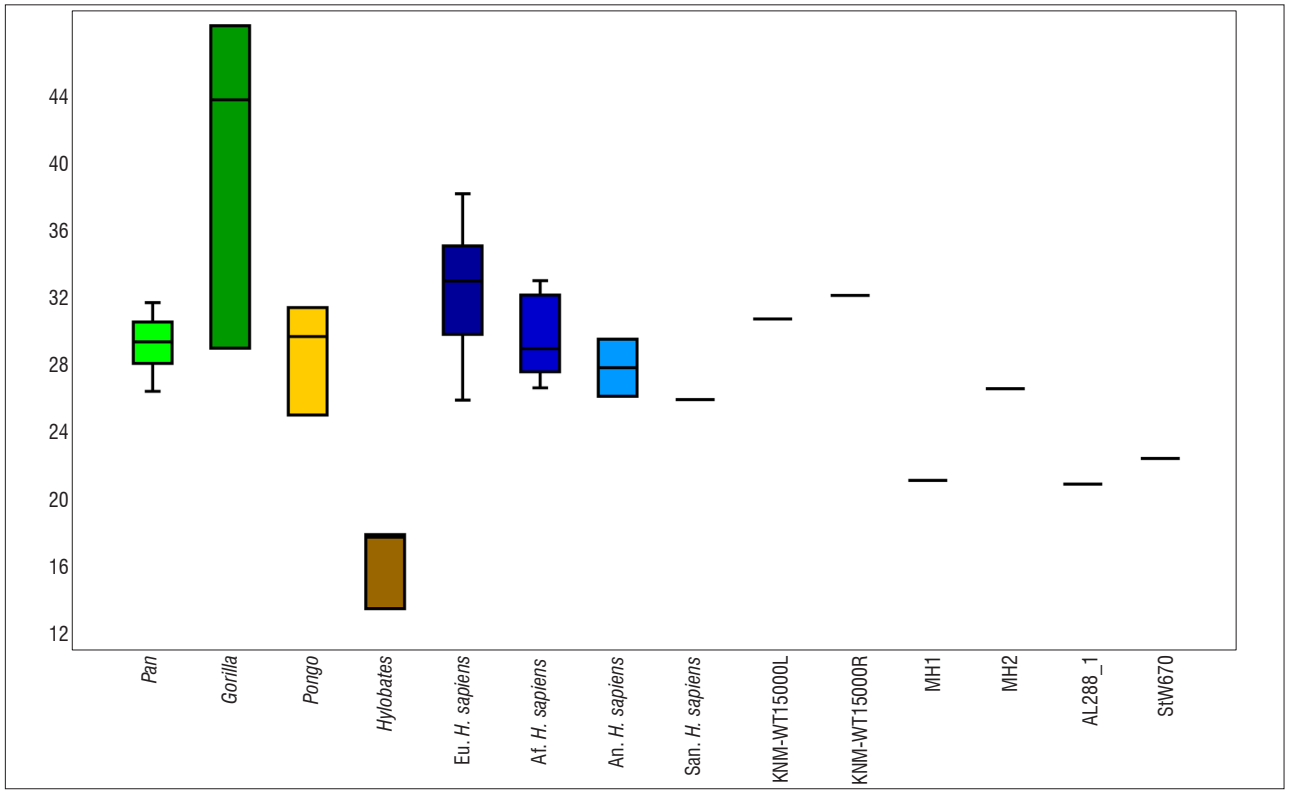
Geometric morphometric analyses

Figure 5 shows a box plot of centroid size distribution of the different groups included in the sample. The centroid size of the StW 670 fossil is smaller than that of *H. erectus* (KNM-WT 15000, both antimeres), *A. sediba* (MH2), the great apes (*Pan, Gorilla* and *Pongo*) and *H. sapiens* (including the small-bodied individuals), and larger than that of *Hylobates*, and is located closest to the centroid size of *A. afarensis* AL 288-1 and *A. sediba* MH1.

Regarding shape, in Figure 6, we can observe that PC1 (44.3% of the total variation) polarises most of the non-hominin hominoids on the positive side of the axis, whereas the scatter of *H. sapiens* of different populations is located mostly towards the negative values of the PC1 axis. As it can be observed on the associated warps (Figure 6), the positive values of PC1 (associated with great apes) correspond to first ribs with a shorter relative distance between the head and the articular tubercle and a larger neck width at the articular tubercle. The negative values of PC1 (mainly associated with hominins) correspond to first ribs with a larger relative distance between the head and the articular tubercle and a shorter neck width at the articular tubercle. Moreover, PC1 polarisation differentiates between different rib head orientations. That is, in the positive values of PC1 (associated with non-hominin hominoids), the plane of the rib head is more oblique to the plane of the neck, while in the negative values of PC1 (mainly associated with modern humans and fossil specimens – so hominins), the plane of the rib head is more parallel to the plane of the neck.

As PC2 (16.7% of the total variation), which explains intraspecific variation, does not polarise between groups observed in PC1, as is observed in Figure 6, we will not discuss PC2. The rest of the principal components explain little of the variation so they also will not be discussed here.

Finally, looking at shape similarities according to the Procrustes distances, it is observed in the box plot (Figure 7) that the closest groups to StW 670 are *A. sediba* (mean distance=0.09), small-bodied Andaman *H. sapiens* (mean distance=0.09), European *H. sapiens* (mean distance=0.11), *H. erectus* (mean distance=0.11), small-bodied San *H. sapiens* (distance=0.12), *A. afarensis* (distance=0.12) and the sub-Saharan African *H. sapiens* (mean distance=0.13). The farthest groups are *P. troglodytes* (mean distance=0.16), *Hylobates* sp. (mean distance=0.21), *G. gorilla* (mean distance=0.22) and *Pongo* sp. (mean distance=0.28).



Eu, European; Af, African; An, Andaman

Figure 5: Centroid size distribution of StW 670 compared to first ribs in the comparative sample. StW 670 centroid size is smaller than that of *Homo erectus* (KNM-WT 15000, both antimeres), the great apes (*Pan, Pongo* and *Gorilla)* and *H. sapiens* (including small-bodied ones); bigger than that of *Hylobates;* and is located between the centroid sizes of the two *Australopithecus* species (*A. afarensis* AL 288-1 and *A. sediba* MH1). Comparative extant sample sizes are listed in Supplementary table 1.

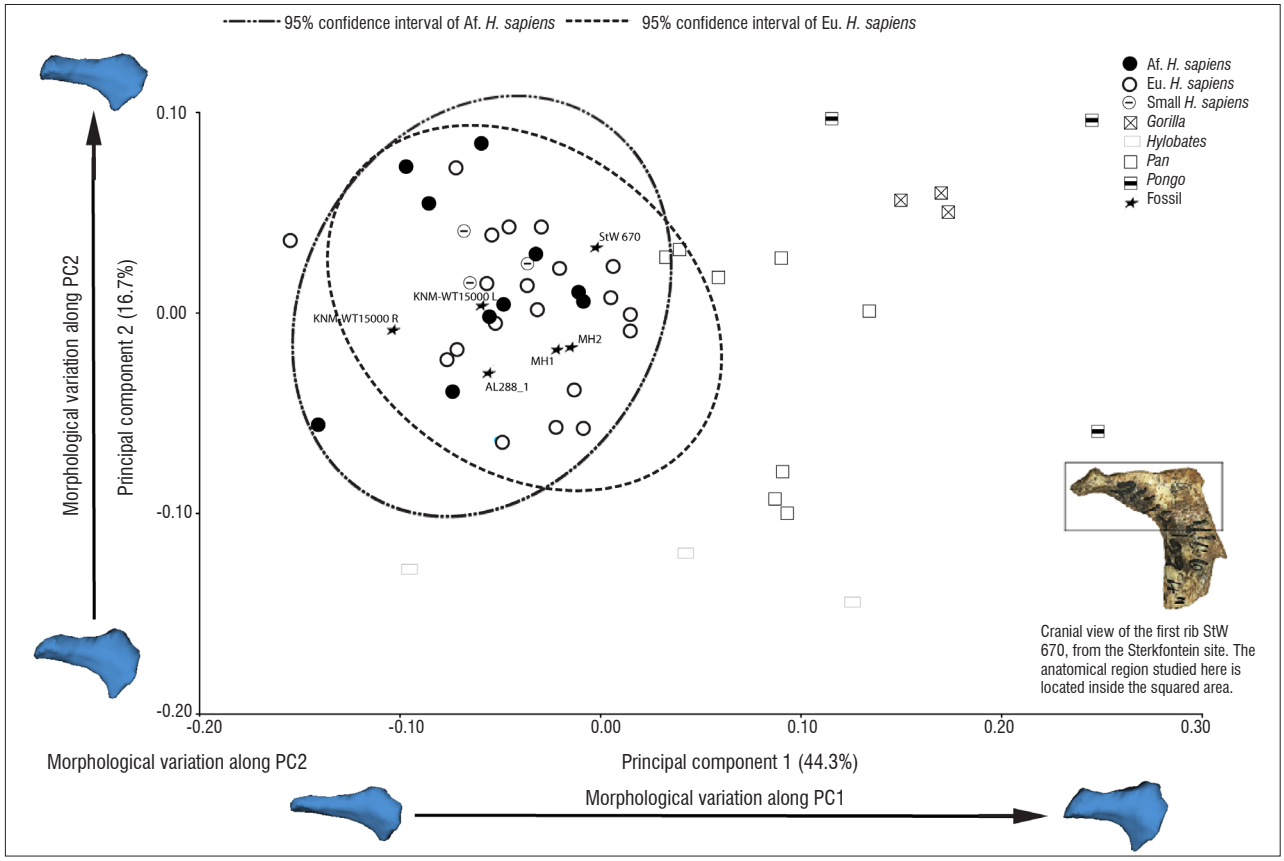


Figure 6: Principal component (PC)-shape analysis of first ribs of all specimens. Scatterplot of PC1 (44.3% of total variance) and PC2 (16.7% of total variance), and warped 3D models of proximal parts of first ribs in axial view. Note that fossil specimens (stars) are within the 95% confidence ellipse of European (Eu) and African (Af) *Homo sapiens*.

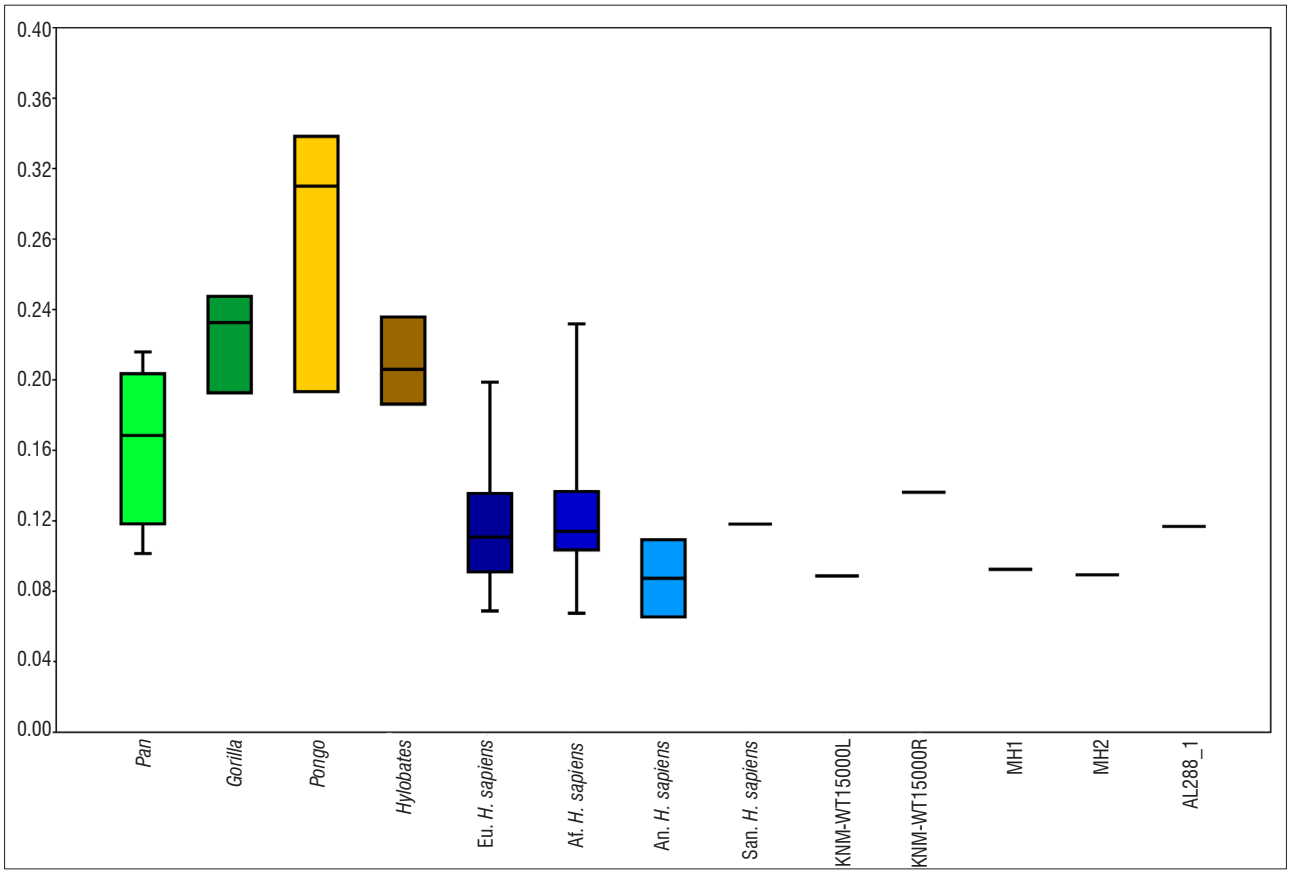


Figure 7: Box plot comparing the Procrustes distances between StW 670 (distance: 0.00) and the comparative sample. The closest groups to StW 670 are *Australopithecus sediba* (mean distance=0.09), small-bodied Andaman (An) *Homo sapiens* (mean distance=0.09), European (Eu) *H. sapiens* (mean distance=0.11), *H. erectus* (mean distance=0.11), small-bodied San *H. sapiens* (distance=0.12), *A. afarensis* (0.12) and sub-Saharan African (Af) *H. sapiens* (mean distance=0.13). The farthest groups are *Pan* (mean distance=0.16), *Hylobates* (mean distance=0.21), *Gorilla* (mean distance=0.22) and *Pongo* (mean distance=0.28).

Discussion and conclusion

Sterkfontein Member 4 has yielded quite a number of hominin remains, with Sts 14 and StW 431 (both *A. africanus*) among the best-known and well-studied hominins from this locality.[21,22] Four beds are recognised within Member 4 – labelled A, B, C and D.[23] StW 431 was excavated in Member 4 Bed B. The excavation beds of Sts 14 and StW 670 were apparently never recorded, so we do not know for certain from which bed StW 670 was excavated. However, StW 670 was excavated close to the excavation area of StW 318, a femur found at U/49 24'0"–25'0". Berger and McHenry[30] surmised the femur to be from an individual of small stature, weighing about 45–50 kg. The small size of the StW 670 rib may mean it is associated with this femur.

Three-dimensional geometrics morphometric analyses, which take into account the 3D spatial relationship between rib head and tubercle, show that the morphology of the preserved portion of StW 670 first rib is quite similar to *A. sediba* as it is close to this taxon in the box plot of the Procrustes distances. Regarding the morphological similarities with the rest of the groups, we can see that, besides *A. sediba*, the morphology of StW 670 is more similar to the *H. sapiens* groups, *H. erectus* and *A. afarensis* than it is to non-hominin hominoids (*Pan*, *Pongo, Gorilla* and *Hylobates).* In fact, in the PCA analyses, every fossil specimen (StW 670, MH1, MH2, AL 288-1 and KNM-WT 15000) falls inside the 95% confidence interval ellipse of *H. sapiens* (Figure 6).

Morphologically, the similarity of the fossils to modern humans observed in this PCA is because these groups are characterised by a larger relative distance between the head and the articular tubercle, and a shorter neck width at the articular tubercle, and because the plane of the rib head is more parallel to the plane of the neck in modern humans and the fossils than in the non-human hominoids. This latter feature observed at the rib head should probably be reflected in the orientation of the transverse processes of the T1 vertebrae, thus future studies should investigate this possibility.

Additionally, the presence of a single facet observed in every early hominin studied (*Australopithecus* and *H. erectus*) also supports the modern-like pattern of this fragment. This observation contrasts strongly with the pattern observed in the great apes (*Pan*, *Pongo*, *Gorilla* and *Hylobates*) as all of them present two articular facets at the rib head. Although we are conscious of the limitations of our sample size, we think that the observations made here are relevant to discuss the modernity of this character in StW 670. So, in the light of these observations, we conclude that the single facet of the rib head of the first rib is a feature unique to the hominin lineage, which emerged, at least, with the genus *Australopithecus.*[10] Moreover, aspects observed in the rib head should be reflected in the adjacent vertebrae as well, but because an associated vertebra does not exist and an assessment of vertebral morphology is not the goal of this work, it should be addressed in future studies.

The similarity of the proximal rib morphology observed in *Australopithecus* and *H. sapiens* first ribs is interesting because it has traditionally been thought that the upper thorax morphology of *Australopithecus* is characterised by a great ape-like form.[6,7] So, another hypothesis to be tested in future studies is that if we accept the derived morphology of the proximal part of the first rib in *Australopithecus* (*A. afarensis*, *A. sediba* and *A. africanus*), we can infer that the archaic pattern of the upper *Australopithecus* thorax should be found at the distal part of the ribs. This would suggest that the proximal part of the rib became derived in the direction of the *H. sapiens* morphology earlier than did the distal part of the first rib. This suggestion is consistent with the fact that the proximal and distal parts of the first rib belong to different morphogenetic modules in ontogeny,[31] which allows both units to evolve at different rates according to different requirements.

Great apes are often described as having a narrow upper thorax, while humans have a more expanded one.[32] It would be useful to be able to use fragmentary ribs in the hominin fossil record to determine when the transition between the two morphologies took place. The 2D analyses performed in previous studies do not make this goal look promising because of the complex morphology of ribs and the ribcage more broadly; however, the 3D analyses performed here show that the spatial configuration of the rib head and articular tubercle could be informative about the archaic or derived morphology, at least of the proximal part of the first rib. This fact, together with features such as the single facet of the rib head[10], could be informative in finding evolutionary traits in the upper thorax of early hominins. More research employing 3D techniques is needed on complete first ribs as well as the rest of the ribs in the costal sequence (from first to twelfth) in order to better quantify and understand *Australopithecus* rib morphology and the role of the thorax in hominin evolution more broadly.

Acknowledgements

We thank Bernhard Zipfel and the Access Committee of the Evolutionary Studies Institute, University of the Witwatersrand, as well as the South African National Centre of Excellence in Palaeosciences for hosting us whilst the materials were studied. D.G-M. and M.B. were funded through CGL2012-37279-MINECO (Spain) and a Leakey Research Grant to D.G-M. S.A.W. and J.E. were funded through the New York University Research Challenge Fund. We thank Emma Mbua, Fred Spoor and the National Museums of Kenya Earth Science Department for providing CT scans of the KNM WT-15000 axial skeleton. We also thank the reviewers for their comments which improved the previous version of the manuscript.

Authors' contributions

G.T. identified the rib fragment as hominin and described it. D.G-M. and J.E. analysed the 3D and linear data, respectively. G.T., D.G-M., J.E., S.N. and S.A.W. wrote the paper. M.B., P.S. and L.B. provided data and supervisory support, respectively.

References

1. Franciscus RG, Churchill SE. The costal skeleton of Shanidar 3 and a reappraisal of Neandertal thoracic morphology. J Hum Evol. 2002;42:303–356. http://dx.doi.org/10.1006/jhev.2001.0528
2. Gómez-Olivencia A, Eaves-Johnson KL, Franciscus RG, Carretero JM, Arsuaga JL. Kebara 2: New insights regarding the most complete Neandertal thorax. J Hum Evol. 2009;57:75–90. http://dx.doi.org/10.1016/j.jhevol.2009.02.009
3. Bastir M, García-Martínez D, Recheis W, Barash A, Coquerelle M, Rios L, et al. 3D analysis of human ribcage ontogeny. Am J Phys Anthropol. 2013;S56:75.
4. Garcia-Martinez D, Barash A, Recheis W, Utrilla C, Sanches T, Garcia Rio F, et al. On the chest size of Kebara 2. J Hum Evol. 2014;70:69–72. http://dx.doi.org/10.1016/j.jhevol.2014.02.003
5. García-Martínez D, Recheis W, Bastir M. Ontogeny of 3D rib curvature and its importance for the understanding of human thorax development. Am J Phys Anthropol. 2016;159(3):423–431. http://dx.doi.org/10.1002/ajpa.22893
6. Aiello L, Dean C. An Introduction to human evolutionary anatomy. London: Academic Press; 1990.
7. Nalla S. The morphology of the upper thorax of *Australopithecus sediba* within the context of the selected hominoids [PhD thesis]. Johannesburg: University of the Witwatersrand; 2013.
8. Schmid P, Churchill SE, Nalla S, Weissen E, Carlson KJ, De Ruiter DJ, et al. Mosaic morphology in the thorax of *Australopithecus sediba*. Science. 2013;340:163–165. http://dx.doi.org/10.1126/science.1234598
9. Bastir M, Garcia-Martinez D, Estalrrich A, García-Tabernero A, Huguet R, Ríos L, et al. The relevance of the first ribs of the El Sidron site (Asturias, Spain) for the understanding of the Neandertal thorax. J Hum Evol. 2015;80:64–73. http://dx.doi.org/10.1016/j.jhevol.2014.10.008
10. Ohman JC. The first rib of hominoids. Am J Phys Anthropol. 1986;70:209–229. http://dx.doi.org/10.1002/ajpa.1330700208
11. Johanson DC, Lovejoy CO, Kimbel WH, White TM, Ward SC, Bush ME, et al. Morphology of the Pliocene partial hominid skeleton (A.L. 288-1) from the Hadar Formation, Ethiopia. Am J Phys Anthropol. 1982;57:403–451. http://dx.doi.org/10.1002/ajpa.1330570403
12. Jellema L, Latimer B, Walker A. The rib cage. In: Leakey RE, Walker A, editors. The Nariokotome *Homo erectus* skeleton. Cambridge, MA: Harvard University Press; 1993. p. 294–325. http://dx.doi.org/10.1007/978-3-662-10382-1_13
13. Berger LR, Hawks J, De Ruiter DJ, Churchill SE, Schmid P, Delezene LK, et al. *Homo naledi*, a new species of the genus *Homo* from the Dinaledi Chamber, South Africa. eLife. 2015;4, e09560, 35 pages. http://dx.doi.org/10.7554/eLife.09560
14. Schmid P. The trunk of the australopithecines. Paris: Origine(s) de la Bipédie chez les Hominidés Presses du CNRS; 1991. p. 225–234.
15. Williams SA. Modern or distinct axial bauplan in early hominins? Comments on Haeusler et al. (2011). J Hum Evol. 2012;63:552–556. http://dx.doi.org/10.1016/j.jhevol.2012.01.007
16. Haile-Selassie Y, Latimer BM, Alene M, Deino AL, Gibert L, Melillo SM, et al. An early *Australopithecus afarensis* postcranium from Woranso-Mille, Ethiopia. Proc Natl Acad Sci USA. 2010;107(27):12121–12126. http://dx.doi.org/10.1073/pnas.1004527107
17. Broom RR, Robinson JT. Man contemporaneous with the Swartkrans ape-man. Am J Phys Anthropol. 1950;8:151–156. http://dx.doi.org/10.1002/ajpa.1330080211
18. Hughes AR, Tobias PV. A fossil skull probably of the genus *Homo* from Sterkfontein, Transvaal. Nature. 1977;265:310–312. http://dx.doi.org/10.1038/265310a0
19. Kuman K, Clarke RJ. Stratigraphy, artefact industries and hominid associations for Sterkfontein, Member 5. J Hum Evol. 2000;38:827–847. http://dx.doi.org/10.1006/jhev.1999.0392
20. Partridge TC, Granger DE, Caffee MW, Clarke RJ. Lower Pliocene hominid remains from Sterkfontein. Science. 2003;300:607–612. http://dx.doi.org/10.1126/science.1081651
21. Robinson JT. Early hominid posture and locomotion. Chicago: University of Chicago Press; 1972.
22. Toussaint M, Macho GA, Tobias PV, Partridge TC, Hughes AR. The third partial skeleton of a late Pliocene hominin (StW 431) from Sterkfontein, South Africa. S Afr J Sci. 2003;99:215–223.
23. Partridge TC, Watt IB. The stratigraphy of the Sterkfontein hominid deposit and its relationship to the underground cave system. Palaeontol Afr. 1991;28:35–40.
24. Rios L, Cardoso HFV. Age estimation from stages of union of the vertebral epiphyses of the ribs. Am J Phys Anthropol. 2009;140:265–274. http://dx.doi.org/10.1002/ajpa.21065
25. Gómez-Olivencia A, Carretero JM, Lorenzo C, Arsuaga JL, Bermúdez de Castro J-M, Carbonell E. The costal skeleton of Homo antecessor: preliminary results. J Hum Evol. 2010;59:620–640. http://dx.doi.org/10.1016/j.jhevol.2010.07.023
26. Zelditch ML, Swiderski DL, Sheets HD, Fink WL. Geometric morphometrics for biologists: A primer. London: Academic Press; 2004.
27. O'Higgins P. The study of morphological variation in the hominid fossil record: Biology, landmarks and geometry. J Anat. 2000;197:103–120. http://dx.doi.org/10.1046/j.1469-7580.2000.19710103.x
28. Hammer Ø, Harper D, Ryan P. PAST: Paleontological statistics software package for education and data analysis. Palaeontol Electron. 2001;4(1), Art. #4, 9 pages. http://palaeo-electronica.org/2001_1/past/issue1_01.htm
29. Klingenberg CP. MorphoJ: An integrated software package for geometric morphometrics. Mol Ecol Resour. 2011;11:353–357. http://dx.doi.org/10.1111/j.1755-0998.2010.02924.x
30. Berger LR, McHenry HM. Body proportions in *Australopithecus afarensis* and *A. africanus* and the origin of the genus *Homo*. J Hum Evol. 1998;35:1–22. http://dx.doi.org/10.1006/jhev.1997.0197
31. Aoyama H, Mizutani-Koseki S, Koseki H. Three developmental compartments involved in rib formation. Int J Dev Biol. 2005;49:325–333. http://dx.doi.org/10.1387/ijdb.041932ha
32. Schultz AH. Vertebral column and thorax. Primatologia. 1961;4:1–66.

Note: This article includes supplementary material

Creationist and evolutionist views of South African teachers with different religious affiliations

AUTHORS:
Michèle Stears[1]
Pierre Clément[2]
Angela James[1]
Edith Dempster[1]

AFFILIATIONS:
[1]School of Education, University of KwaZulu-Natal, Durban, South Africa
[2]Learning, Teaching, Assessment and Training (ADEF) Laboratory, Aix-Marseille University, Marseille, France

CORRESPONDENCE TO:
Michèle Stears

EMAIL:
stearsm@ukzn.ac.za

POSTAL ADDRESS:
School of Education, University of KwaZulu-Natal, Private Bag X03, Durban 3605, South Africa

DATES:
Received: 17 June 2015
Revised: 18 Oct. 2015
Accepted: 12 Jan. 2016

KEYWORDS:
biology teaching; comparative study; religious orientation; teachers

HOW TO CITE:
Stears M, Clément P, James A, Dempster E. Creationist and evolutionist views of South African teachers with different religious affiliations. S Afr J Sci. 2016;112(5/6), Art. #2015-0226, 10 pages. http://dx.doi.org/10.17159/sajs.2016/20150226

Concerns have been raised in the scientific community that many teachers do not accept evolution as a scientific, testable phenomenon, and this is evident in their teaching. The non-acceptance of evolution theory is often heavily influenced by religious groups that endeavour to eliminate evolution from the curriculum. In South Africa, the inclusion of evolution in the curriculum is a recent event. This study focused on teachers' views of evolution in relation to their religious affiliations. A questionnaire was developed and was validated by the Biohead–Citizen Project, and was then administered to more than 300 South African teachers and student teachers. Equal numbers of pre-service and in-service teachers were sampled. The groups included equal numbers of biology, English, and generalist teachers at primary school level. The results showed differences between teachers from different religions with regard to their views of evolution. Among teachers who identified as agnostic or atheist, 17% held creationist views. Among teachers who identified as Protestant, other Christian, or Muslim, 70% held creationist views. This study also examined, for the first time, the views of teachers belonging to religions not included in previous research. Of these, only 25% of Hindus held creationist views. Fewer adherents of African Independent Churches held creationist views compared with teachers from traditional Protestant denominations; for example, only 30% of Zionist followers and 40% of Shembe followers held creationist views. This study adds important knowledge by including the views of teachers from religions not previously researched.

Introduction

On 21 June 2006, the Inter-Academy Panel (IAP[1]),a global network of 68 science academies, published a joint statement on the teaching of evolution. The statement read as follows:

> *We, the undersigned Academies of Sciences, have learned that in various parts of the world, within science courses taught in certain public systems of education, scientific evidence, data, and testable theories about the origins and evolution of life on Earth are being concealed, denied, or confused with theories not testable by science. (p. 1)*

This statement acknowledged that several student teachers or qualified teachers did not accept evolution as a scientific, testable phenomenon, and this bias was evident in their teaching. This state of affairs was initially reported on by a number of authors[2-4] in the USA, where the teaching of evolution is still a contentious issue in many communities.

Research has shown that in several countries, differences exist with regard to the inclusion of evolution in the curriculum.[5-8] Countries such as Italy and Germany have experienced controversy over the teaching of evolution.[9] Some controversy also exists in the United Kingdom regarding the teaching of evolution.[10] Although the level of acceptance of evolution theory is generally higher in Western Europe than in the USA, special creationist ideas are widespread.[11] More Protestants who belong to non-mainstream denominations, as well as conservative Muslims, accept the theory of special creation compared with any other religious groups. This means they believe God created all living things and that no changes have occurred since creation[12,13] (see Table 2 for percentages of people who hold this belief).

A comprehensive body of research points to the influence of religion on the acceptance of evolutionary theory. A study reported by Martin[14], whilst not conducted with teachers, provides valuable insight into the views of the public based on the views of governing bodies of various Christian denominations. Martin's research showed that more people in the USA accept evolution theory than those who reject it.[14] However, many members of special creationist groups – such as Pentecostal Protestants – find no compatibility between their faith and evolution.[14]

Several scholars, namely BouJaoude, Asghar, Wiles, Jaber, Sarieddine, and Alters[15] as well as Clément[16] report that in Lebanon, Christian teachers and Muslim Druze teachers are more inclined to accept the theory of evolution than other Muslim groups. The same authors report that some individuals from Muslim groups who do accept the theory have reinterpreted it to exclude human evolution. Clément[16] showed that Muslim teachers' views of evolution differed significantly from one country to another. In Burkina Faso, Muslim (Sunni) teachers accept evolution more readily than their Protestant colleagues. In Lebanon, there is no significant difference between Christian, Druze, and Shiite teachers' views of evolution, with only their Sunni colleagues' views being a little more creationist.

Asghar[17] conducted a study to assess the evolutionary views of Muslim science teachers from diverse contexts such as Canada and Pakistan. Her study showed that most teachers were prepared to accept the evolution of living organisms, apart from human beings. They felt that human evolution contradicted their Islamic beliefs. BouJaoude, Wiles, Asghar, and Alters[18] conducted a similar study with learners from three different Muslim groups in Lebanon and Egypt. The study showed that the religious beliefs of Sunni and Shiite Muslim learners from both countries influenced their views of evolution.

In the South African context, the teaching of evolution in schools emerged as an issue with the implementation of a new Life Sciences curriculum, spelled out in the National Curriculum Statement.[19] According to Lever[20], the new curriculum endeavoured to include content that had been omitted because it was seen as being alien to the ethos of Christian National Education, which had underpinned the previous curriculum. The theory of biological evolution was one such topic.[20] Most teachers of biology had obtained their qualifications at colleges of education where evolution was not taught as an integral part of biology. The exception was a small number of colleges that fell under the administration of the various provincial administrations rather than the Department of Bantu Education. According to Sanders and Ngxola, even teachers who had obtained bachelor degrees in the biological sciences agreed with their less-qualified counterparts that they had inadequate knowledge of evolution or how to teach the subject.[21] This state of affairs has sparked a number of research projects[22,23] to assess how teachers view evolution theory, and their attitudes towards the teaching of the subject.

In the current South African curriculum,[24] evolution constitutes 22% of the Grade 12 Life Sciences curriculum in terms of marks allocated and teaching time. This places a great responsibility on teachers to teach this section competently, to enable their learners to pass the exit-level matriculation examinations at the end of Grade 12. However, many South African teachers do not accept the theory of evolution[22,23] and these teachers initially felt a sense of inner conflict. They were expected to teach a topic they had not studied in their initial training, had no experience of teaching, and about which they held negative views.[25-27]

The South African education system is centralised and prescribes what teachers are required to teach. It is dominated by a national exit-level examination, the National Senior Certificate, which determines access to tertiary education. The highly prescriptive nature of the curriculum and external assessment make it impossible for a teacher to omit the subject of evolution at Grade 12 level. At professional development workshops organised to help teachers improve their knowledge of evolution, some teachers voiced their concerns about teaching evolution.[20] As recently as 2013, Keke[28] found that teachers still listed evolution among the topics they found most difficult to teach.

The relatively recent introduction of evolution in the Life Science curriculum in South Africa, and the identified teacher concerns mentioned above, have spawned a number of research endeavours. Kyriacou, De Beer and Ramnarain[29] identified several problems related to currently-employed teachers' knowledge of evolution. One of these was objections raised by some fundamentalist religious groups. Similarly, a study by Mpeta[27] showed that religion played an important role in both learners' and teachers' views on evolution. Furthermore, Abrie[25] found that student teachers who were more religiously observant were more opposed to teaching evolution than their less observant peers.

Pillay's[30] research included both Christian and Muslim teachers. Although lack of content knowledge played an important role in their opposition to evolutionary theory, all the Muslim teachers and most of the Christian teachers were of the view that their religious beliefs contradicted the theory of evolution. The work of Yalvac[31] showed that Muslim teachers who were interviewed saw themselves as creationists. In that study, many Christian teachers surveyed were also opposed to evolutionary theory. Naidoo[32] argues that Hindu teachers and learners have no problem with evolutionary ideas because Hinduism has no creation story depicting creation as a once-off event.

The Biohead–Citizen research project[33] is an international study that has obtained data on teachers' conceptions across numerous countries. Initially, eighteen countries participated in the project and subsequently the Biohead–Citizen questionnaire has been administered in a further twelve countries. Data obtained from this project show that countries sometimes show different trends in teachers' acceptance of evolutionary theory. For instance, Clément and Quessada[34] found that the percentage of teachers holding creationist views differed from country to country, even for teachers of the same religious group. Fundamentalist creationist conceptions ranged from 0% to 62% among Roman Catholic teachers, and from 2% to 76% among Protestant teachers in various countries. Within each of 26 countries studied, no significant difference existed among the various religions with regard to the percentage of teachers who held fundamentalist creationist views.[12,34,35] However, there are exceptions. In Brazil, Protestant teachers hold more creationist beliefs than those of their colleagues in other religions.[36] In Lebanon, the ideas of Sunni Muslim teachers are a little more creationist than their colleagues'. In Burkina Faso, the views of Muslims are less creationist than those of Protestants. In South Korea, the views of Protestant teachers are more creationist than their colleagues who are mainly agnostic, atheist or Buddhist.[37]

These findings prompted us to administer the Biohead–Citizen questionnaire in South Africa. South Africa provides an interesting case study because the country has diverse religions, including Christianity, Islam, Hinduism, Buddhism, and African traditional churches or beliefs.[38] Census 2001, the last time data on religious affiliation was collected in South Africa, identified four categories of Christian churches: mainstream (Roman Catholic, Anglican, Methodist, Presbyterian, Lutheran and Congregational), African Independent Churches (Zionist, Shembe and Ethiopian-type), 'other Christians' (not defined), and 'Pentecostal / charismatic' (not defined).[38]

Roman Catholic and mainstream Protestant churches have a long history in most countries of the world. The African Independent Churches (AICs) have developed in Africa as offshoots of Protestant-type Christian churches.[38] African traditional beliefs comprise a variety of belief systems that are inherently African, but most adherents also hold some Christian beliefs. Because AICs are unique to Africa, they are separate from mainstream Protestant churches. In Census 2001, almost 32% of South Africans belonged to one of the AICs, with the Zion Christian Church (ZCC) being the largest.[39] The ZCC has its origins in the Catholic Apostolic Church and is regarded as a form of African Pentecostalism.[40] The oldest AIC in Africa is the Nazareth Baptist Church, also known as Ibandla namaNazaretha or Shembe, after its founder. Estimates of its following vary from about 250 000 people[38] to 4.5 million people.[41] Although the church has its roots in Christianity, it is a mixture of Zulu traditional beliefs and Christianity. Members believe in the Holy Trinity but observe the Jewish Sabbath, and hold the belief in an African Messiah.[41] Most followers of Shembe live in the province of KwaZulu-Natal.

Mainstream Protestant churches in many countries have accepted the theory of evolution. Charismatic Protestant churches, including Pentecostal and the Full Gospel Church, are much more literal in their interpretation of biblical texts.[15]

Table 1 shows our reasoning for the categories we defined in the present study. In future studies, the question (P13, shown in the Methodology section of this paper) eliciting information about religion should be reworded. Our results showed that some members of Pentecostal or charismatic churches may have identified themselves as Protestant, but others wrote the name of the church to which they belonged. Thus eight members of the Full Gospel Church were identified separately from the Protestant group. Table 1 gives a brief description of the religious groupings used in this study.

One of the aims of the Biohead–Citizen Project was to compare the views about evolution among teachers of different religions.[48,49] Such information could assist in structuring teacher education programmes to address issues such as teachers who are strongly opposed to accepting evolution as an important topic in the teaching of biology. Similarly, the purpose of our study was to investigate the relationship between beliefs about evolution and religious affiliation among South African teachers. The question that drove our research was: *What views do South African teachers from different religious affiliations hold with regard to evolution?*

Conceptual framework

Various theoretical and conceptual frameworks have been used to frame studies relating to issues surrounding the teaching and learning of evolution, as well as the acceptance of evolutionary theory. The 'stages of concern' theory developed by Fuller[50] and Hall and Loucks[51] has been helpful in designing appropriate professional development courses and has been applied particularly for courses in evolution.[21] Scott's[52] creation–evolution continuum illustrates the positions people may hold in the evolution–creation controversy. Table 2 provides a summary of Scott's model.

Table 1: A brief description of belief systems of religious affiliations

Code	Belief system	Description
AGN	Agnostic Atheist	Agnostics believe it is not possible to say with certainty whether or not God exists.[42] Atheists deny the existence of a deity.[42]
CAT	Roman Catholic	The main teachings of the Roman Catholic church are God's objective existence, and acknowledgement of Him as creator of heaven and earth. The church does not have an official position on whether various life forms developed instantaneously (creationism), or over the course of time (evolution). However, if they did develop slowly, they did so under the impetus and guidance of God and their ultimate creation must be ascribed to God.[43]
PRO	Protestant	Protestants believe in God as the creator of the universe and all in it. Beliefs about how this creation occurred vary. Some mainstream Protestants have theistic evolutionist beliefs, whereas others interpret the Bible more literally. Officially these churches accept evolution as a possible explanation for the diversity of life.[44]
ELS CHR	Other Christian churches (mostly Pentecostal)	Pentecostalism is a renewal movement within Protestant Christianity that emphasises personal and direct experience of God through baptism in the Holy Spirit. This group believes in the inerrancy of the Bible, and therefore interprets biblical texts literally.[44] This means all things were created in six days and no changes have occurred since then (creationist view).
FULL	Full Gospel Church of God	The Full Gospel Movement is associated with Pentecostalism and Charismatic Christianity. While their religious doctrines are different, they hold similar views to that of the Pentecostal churches with regard to the inerrancy of the Bible.[44]
MUS	Muslim (Sunni)	Sunni Islam is a denomination of Islam that is sometimes referred to as 'orthodox' Islam. Sunni Muslims regard Allah as the only God and creator of all.[45]
HIND	Hindu	Hinduism is not predominantly earth-centred, and puts much emphasis on other planes of existence. There is no one simple account of creation, and there are many detailed and inter-related stories.[46]
ZIO	Zion Christian Church	The ZCC fuses African traditions and values with Christian faith. The bishop and ministers of ZCC preach the Gospel of Jesus Christ as laid out in the Bible.[41] They have no specific doctrine regarding creation of the universe.
SHEMBE	Shembe and other African Religions	The church of the ama-Nazarites, also generally known as the Nazareth Baptist Church, was founded by Isaiah Shembe. His theology was based on his conviction that he was the mouthpiece and instrument of Jehovah. This religion has no pertinent doctrine regarding creation.[47]

Table 2: An adaptation of Scott's creation–evolution continuum

YOUNG EARTH	Flat Earth Geocentrism Young Earth Creationism	Special Creation Creation–evolution continuum Evolution
OLD EARTH	Gap Day-age Progressive Theistic evolution Materialist evolution Intelligent design	

Source: Modified from a diagram by Scott[52]

According to Scott[52], Young Earth creationists are biblical literalists who believe, based on calculations from characters and events named in the Bible, that the Earth came into existence only a few thousand years ago. These include the Flat-Earthers, who believe the Earth is a round disc, and the geocentrists, who believe the Earth is a sphere but do not accept the findings of modern physics and geology pertaining to the age of the Earth. Because this group interprets the creation story in their religious books literally, they deny biological descent with modification (i.e. evolution).

Scott[52] refers to several groups collectively as Old Earth creationists, as they accept that the earth is ancient. The Gap Theory and Day-Age creationists have various explanations of biblical events to attempt to reconcile science and religion. Most Old Earth creationists hold some view of progressive creationism. This view blends special creationism with aspects of science. It accepts speciation but rejects macro-evolution. Intelligent design is a branch of progressive creationism that believes the study of living organisms produces evidence of God's creation. They believe complex organisms could not have evolved by chance, and must have been designed by an 'intelligent being' – that is, God. Adherents of 'intelligent design' are found across the continuum of creationist beliefs.

There is a sharp division between Young Earth creationists and Old Earth creationists. However, the division between the Old Earth creationist subgroups is less clear, as there is a gradual increase in the extent to which science influences the beliefs of these groups. At one end of the continuum are theistic evolutionists, who accept macro-evolution but believe that it is managed by a divine being. Most mainstream churches accept this view. Materialist evolutionists, by contrast, hold a non-religious view and accept only scientific explanations for life and its diversity.[52]

This continuum is useful in classifying people from the Christian faith, and to a lesser extent other monotheistic religions that originated in the Middle East. But it does not accommodate the views of people from other religions, many of which have no creation story as a central tenet of their religion. Because we wanted to include other religions too, we needed a different conceptual framework. The Biohead–Citizen Project uses three concepts: *evolutionist*, *creationist*, and simultaneously *creationist and evolutionist.* As the famous evolutionist Dobzhansky[53(p.127)] claimed in 1973*:* 'I am a creationist and an evolutionist. Evolution is God's, or Nature's, method of Creation.' This category is not far from the theistic evolution category defined by Scott[52], and does not conflict with teaching biological evolution – whereas the others forms of creationist conceptions do.

Methodology

The Biohead–Citizen Project was funded by the European Commission, and adheres to the ethical requirements of the Commission (number CICT-CT-2004-506015). However, the South African part of the project was funded by the South African authors of this paper. A questionnaire

developed and validated by the Biohead–Citizen research project[33] was used for this study. The questionnaire contained 144 questions covering a number of topics; 15 questions were dedicated to evolution and 17 questions related to personal information (gender, age, level of education, and religion or religious opinions).

A few questions followed a multiple-choice format. Others used a Likert-type scale, with four responses (ranging from 'I agree' to 'I disagree') to a statement. Each response was scored on a scale of 1 to 4. Questions on evolution were designed to elicit answers that indicated whether a teacher's conceptions were towards the creationist end of the continuum or towards the opposite evolutionist pole. Teachers could also choose answers that could be classified as simultaneously evolutionist and creationist.

For example, items B42 to B48 asked respondents to indicate the importance of a number of factors in species evolution. Factor B42 is 'Chance', and teachers were asked to rank its importance from 'great importance', 'some importance', 'little importance' to 'no importance at all'. An answer of 'great importance' in this case would be classified as strongly evolutionist, while 'no importance at all' would be classified as strongly creationist. In the same section, factor B48 is 'God'. Here, an answer of 'great importance' would be strongly creationist, and 'no importance at all' strongly evolutionist. In both cases, an answer of 'some importance' shows a degree of ambivalence that classifies the answer as simultaneously evolutionist and creationist. An answer of 'little importance' tends towards one end of the continuum, and would be classified as evolutionist or creationist depending on the factor.

Some questions were adapted to suit the South African context. For instance, respondents were asked to identify their religious affiliation, where the religions listed were those commonly found in South Africa. These are illustrated in the following question:

P13. **Are you?** (tick only **ONE** box):

☐ Agnostic/Atheist

Christian: ☐ *Catholic* ☐ *Protestant* ☐ *Zionist* ☐ *Shembe*

☐ *Other (specify):* _____

Moslem: ☐ *Sunnite* ☐ *Shiite* ☐ *Other (specify):* _____

☐ Jewish

☐ Hindu

☐ Buddhist

☐ Other religion/belief (specify): _____

☐ I don't want to answer

Our sample was limited to the province of KwaZulu-Natal, South Africa, as the first step of a possible larger project. Potential respondents were assured that their participation was voluntary, and all who agreed to participate were assured of anonymity. The sampling method was the same as that used in previous studies,[33] to allow comparisons of the effect of religious affiliation, teaching specialisation and experience on teachers' positions on the continuum between evolutionism and creationism.

A total of 336 teachers filled out the questionnaire, with their expertise or training being as follows:

- 53 in-service teachers of biology (secondary schools)
- 60 pre-service teachers of biology (final year of their training)
- 65 in-service generalist teachers in primary schools
- 58 pre-service generalist primary school student teachers (final year of their training)
- 49 in-service teachers of English (secondary schools)
- 51 pre-service student teachers of English (final year of their training).

The above list represents our expected 'hierarchy' of knowledge of evolution, from the greatest to the least.

Respondents were contacted in various ways, all of which were a form of convenience sampling. These forms of contact were:

- Distribution of questionnaires at a provincial workshop, where the workshop facilitators agreed to allow additional time after the workshop to complete the questionnaire.
- Distribution to selected biology teachers, where the researchers oversaw the completion of the questionnaires after school hours.
- Distribution to students at two tertiary institutions: a private college of education, and the School of Education at a university.

The data were collected in 2013 in KwaZulu-Natal, where the South African researchers are based, under the supervision of the South African participants in the Biohead–Citizen Project study. Answers were coded exactly as described for the BioHead-Citizen Project.[33] There were no incomplete questionnaires.

Statistical analysis was conducted by the French participant in conjunction with a qualified statistician. The software R (R Development Core Team[54]) for multivariate analyses was used for this purpose. We conducted between-class analyses (Dolédec and Chessel[55]) to discriminate between teachers' groups as defined by the main parameters (gender, age, level of education, religion, and teaching experience).

A Monte Carlo permutation test (Romesburg[56]) implemented in the ade4 library of R was then used to see if the difference between the groups was or was not statistically significant. Multivariate tests were used to test for differences in scores on any of the questions, according to the main parameters listed in the previous paragraph. Scores for the 15 questions related to evolution were analysed. We also ran a principal component analysis of orthogonal instrumental variables (PCAOIV), as described by Sabatier et al.[57], to analyse if the effect of one parameter remained significant after suppressing another significant effect. Here we tested the independence of the effect of 'religious affiliation' from the characteristics of the six groups of teaching expertise, identified as follows:

- in-service biology
- pre-service biology
- in-service English
- pre-service English
- in-service primary generalists
- pre-service primary generalists.

As shown in previous studies,[58,59] these kinds of multivariate analyses are appropriate for the data collected with the Biohead–Citizen questionnaire.

Results

Table 3 shows the religions of respondents and the number of respondents in each religious group, as well as the codes used in the statistical analysis. Many respondents wrote the name of their church rather than ticking a box provided. The churches were then grouped using the categories established for Census 2001.[38]

Full Gospel was not included as a choice in Question P13, but was specifically written in by eight respondents. A separate code was therefore created to accommodate this denomination.

Two between-class analyses were performed, to test whether significant differences emerged that could distinguish between the groups, based on their responses to all 15 questions. The outcomes we examined were knowledge and beliefs about evolution among the tested groups:

1. The between-class analysis differentiated between the six sample groups (in-service or pre-service; biology, English or primary generalist). As expected, biology teachers (both pre and in-service) had significantly better knowledge about some processes of evolution, and held slightly more evolutionist views, than any of their colleagues.
2. The most marked differences identified by between-class analysis were related to the teachers' religious beliefs, as shown in Figure 1.

Table 3: Numbers of teachers in each religious group identified in the sample

Code	Religious affiliation or belief system	Number of participants	Percentage of sample
AGN	Agnostic and atheist	17	5.1
CAT	Roman Catholic	65	19.3
PRO	Protestant	97	28.9
FULL	Full Gospel Church of God	8	2.4
ELS CHR	Other Christian religions, mostly Pentecostal	15	4.5
MUS	Muslim (Sunni)	19	5.7
HIN	Hindu	36	10.7
ZIO	Zionist	30	8.9
SHEMBE	Shembe and other African religions	20	6.0
OTHER	Other	2 (1 Buddhist)	0.6
NR	No answer	27	8.0
	TOTAL	336	

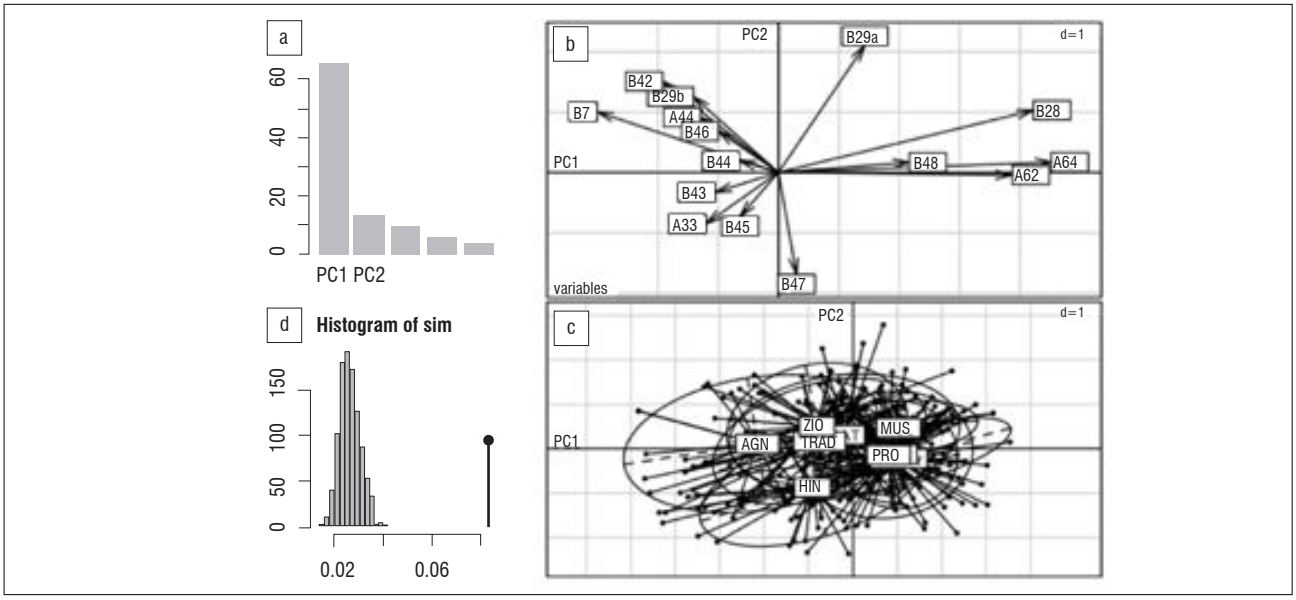


Figure 1: Between-class analysis differentiating the groups of teachers (according to religious affiliation) based on the 15 questions related to evolution. (a) Contribution of each principal component to the total variance: the first principal component (PC1) accounts for more than 65% of the total variance; the second (PC2) accounts for 12%. Subsequent components are considered to be background noise. (b) Loading of answers to the 15 questions (evolution variables) on PC1 (horizontal axis) and PC2 (vertical axis). The length of each arrow indicates the loading of each question on PC1 and PC2. (c) Each point summarises a teacher's answers to the 15 questions, and is related to the group centroid for each religious affiliation. Each ellipse encompasses two-thirds of the teachers in a religious group. (Figure 2 presents these results in simplified and enlarged format.). (d) Histogram of simulations: Results of a randomisation test (Monte Carlo type) generated from 1000 iterations, randomly assigning a religious affiliation to each teacher. The observed variance (to the right) is distinct from random, showing that the results in Figure 1c) are not random ($p<0.01$).

Figure 1a shows that the first principal component (PC1) explains over 65% of the total variance. The second component (PC2) accounted for 12%, meaning these two components together accounted for most of the variance (77%). It was therefore not necessary to investigate further components.

Figure 1b shows the contribution of each of the 15 questions on evolution to PC1 and PC2. Five questions – those with the longest arrows – have the highest loading on PC1 (B64, A62, B28, and to a lesser extent B48 and B7). All these questions are related to the continuum between evolutionism and creationism. Thus PC1 differentiates among the respondents' views in terms of their position on that continuum, with the most strongly creationist answers to the right of the axis and the most strongly evolutionist to the left. Figure 1c shows the diversity of teachers' views within each religious affiliation. The ellipses encompassing two-thirds of teachers from each religious affiliation overlap. Nevertheless, their centroids are distributed along the horizontal axis, which is easier to see in Figure 2 (an enlargement of Figure 1c).

Figure 1d shows, from a Monte Carlo randomisation test, that answers to questions relating to evolution differed significantly among the religious groups identified in this study ($p<0.01$). Nevertheless, it is *a priori* possible that the observed significant differences among religious affiliations could be a result of the differences in teachers'

views depending on their specialisation (biology, English or generalist primary school). We mentioned earlier the significant differences linked to specialisation.

To test the possibility that the differences shown in Figures 1 and 2 were influenced by specialisation, we conducted a principal component analysis of orthogonal instrumental variables (PCAOIV; Sabatier et al.[57]). The effect of the six groups (in-service or pre-service, plus subject speciality) was suppressed so that we could investigate other influences. A subsequent between-class analysis, followed by a randomisation test, showed a persistent significant difference among the religious affiliations ($p<0.01$), with the same order along the horizontal axis (PC1) as shown in Figures 1 and 2. Consequently, the differences shown in Figures 1 and 2 are not merely artefacts from different levels of training in biology among the teachers.

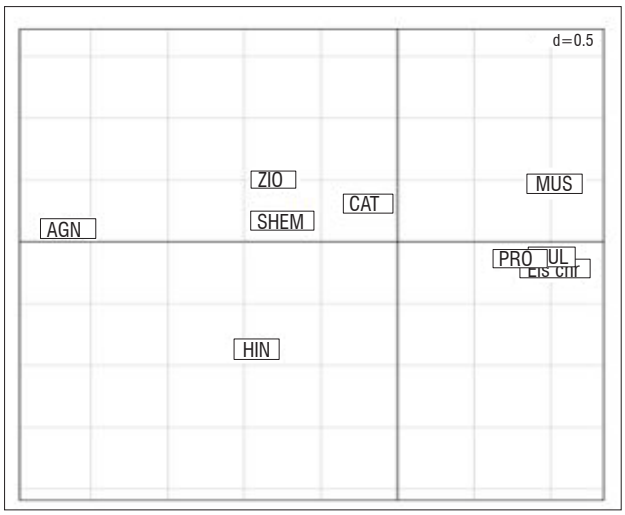


Figure 2: Group centroids characterising the positions of religious affiliations along PC1 (horizontal axis) and PC2 (vertical axis).

Figure 2 shows the group centroids for the religious affiliations in relation to PC1 and PC2. Protestant (PRO), Full Gospel (FUL) and Other Christian (ELS CHR) categories overlap entirely. The most strongly evolutionist answers were given by agnostic or atheist (AGN) teachers, and the most strongly creationist answers by Muslim (MUS), Protestant (PRO), Full Gospel (FUL) and Other Christian (Els Chr) groups. The Zionist (ZIO), Shembe (SHEM), Hindu (HIN), and to a lesser extent Roman Catholic (CAT) teachers held views that were more evolutionist than creationist.

The main finding of the PCA was that the sampled teachers could be positioned according to their religious affiliation along a continuum, from extreme evolutionism to extreme creationism. At one end were those holding evolutionist views (agnostic or atheist), in the middle was a group that held mixed evolutionist and creationist views (Zionist, Shembe, Hindu and Roman Catholic), and at the other end was a group that held strong creationist views (Muslim, Protestant, Full Gospel and Other Christian). The analysis then illustrated the responses of the groups by graphing their responses to the three questions that achieved the highest loadings on PC1 (namely questions A64, B28 and A62).

Because only eight respondents identified 'Full Gospel' as their religious affiliation, they were incorporated into 'Other Christian' (Els Chr) for further analyses. However, this merging of Full Gospel with Other Christian was only done after the PCA, which showed a complete overlap between Full Gospel with Other Christian. Hence, 'Other Christian' gives insight into the views of 23 adherents of mainly Pentecostal or charismatic churches. Figures 3, 4 and 5 show the frequency distribution for answers given by teachers who indicated their religious affiliation ($n=309$).

Question A64 is about the origin of life. Four statements are presented; the first and second statements are strongly evolutionist (the first being most dogmatic), and the fourth statement is strongly creationist. Statement three is both evolutionist and creationist, acknowledging the control of evolution by a creator (theistic evolution). The question is:

> **A64**. Which of the following four statements do you agree with the most? (tick only ONE answer).
>
> ☐ It is certain that the origin of life resulted from natural phenomena.
>
> ☐ The origin of life may be explained by natural phenomena without considering the hypothesis that God created life.
>
> ☐ The origin of life may be explained by natural phenomena that are governed by God.
>
> ☐ It is certain that God created life.

Figure 3 lists the groups, based on religious affiliation, ranked from most evolutionist at the top (agnostic or atheist) to most creationist (Muslim and Other Christian) at the bottom. Approximately 70% to 75% of Protestants, Muslims and Other Christians were certain that God created life. This proportion was about 50% among Roman Catholics, and dropped to less than 20% for the atheist and agnostic group. A large proportion of every group except atheists and agnostics selected the third answer, which allows for evolution under the control of God.

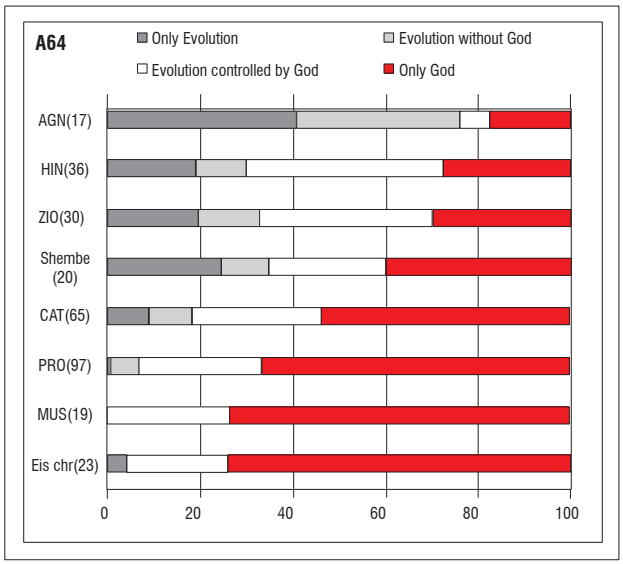


Figure 3: Percentage of teachers (grouped by religious affiliation) who selected creationist (red), evolutionist (black and grey), or mixed creationist and evolutionist (white) answers to question A64 on the origin of life.

Only the atheist and agnostic (AGN) group had a majority of responses that dispensed with the need for a creator entirely, with almost 80% of that group selecting the first or second option. However, three teachers (17%) in the AGN group selected the fourth option, that God created life. Because atheists reject the idea of God, we may assume these three teachers were agnostic. About 30% of Hindu and Zionist respondents, and 40% of Shembe adherents, chose the strongly creationist answer (the fourth statement). The Muslim group presented the most strongly creationist response, with no respondent selecting answers that excluded God.

The third option, compatible with a theistic evolution perspective, was least attractive to the AGN group, with only one AGN respondent selecting that option. It was the most popular option for Hindu and Zionist respondents (36.7% and 41.7%, respectively), and was chosen by more than 25% of all remaining groups except Other Christian, of whom 21.7% chose this option.

Question B28 was about the origins of humankind. As in question A64, four statements are presented, the first two being most strongly evolutionist and the last most strongly creationist:

B28. Which of the following four statements do you agree with most? Select ONLY one sentence:

- ☐ It is certain that the origin of the humankind results from evolutionary processes.
- ☐ Human origin can be explained by evolutionary processes without considering the hypothesis that God created humankind.
- ☐ Human origin can be explained by evolutionary processes that are governed by God.
- ☐ It is certain that God created humankind

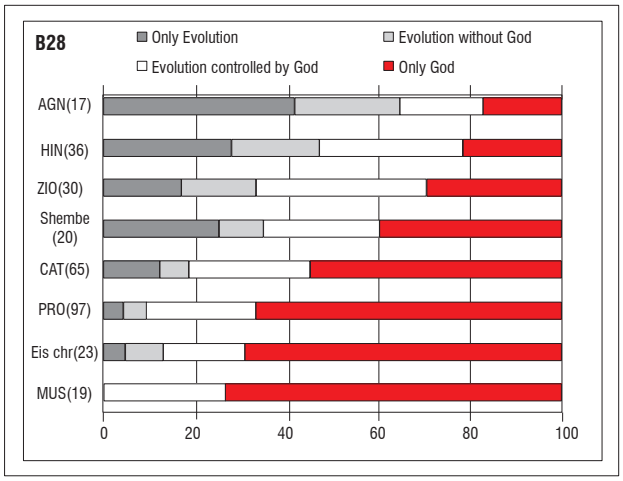


Figure 4: Percentage of teachers (grouped by religious affiliation) who selected creationist (red), evolutionist (black and grey), or mixed creationist and evolutionist (white) answers to question B28 on the origin of humankind.

Figure 4 lists the groups ranked from most evolutionist at the top (agnostic or atheist) to most creationist at the bottom (Muslim). Approximately 70% to 75% of Protestants, Muslims and Other Christians were certain that God created humankind. This proportion was about 55% among Roman Catholics, and dropped to between 40% and 20% for Shembe, Zionist and Hindu, and to less than 20% of the AGN group. Once again, between 17.4% and 36.7% of all respondents showed theistic evolution beliefs by selecting the third answer, which allows for the controlling hand of God in the natural process of human evolution.

For Question B28, agnostic or atheist teachers were the least likely to select the third answer (17.6%), with the Other Christian group being second least likely (36.7%). About 25% of each of the remaining groups chose the third option. No Muslims, and only 10% to 15% of Protestants and Other Christians, selected purely evolutionist answers – namely that humans arose by natural evolution without the intervention of God. Nearly 50% of Hindus selected evolutionist answers (first and second options), in contrast to the 30% of Hindus who selected evolutionist answers to Question A64. The pattern of answers for Zionist and Shembe adherents was similar to that for question A64, with over 30% of each group adopting an evolutionist position.

Question A62 makes another kind of statement concerning the origin of humankind. Respondents are asked to select three terms that they believe are most strongly related to the origins of humankind. Coding of the answers (0 to 3) was based on how many terms chosen were associated with creation. The first, third and fifth terms are creationist, whereas the second, fourth and sixth terms are evolutionist:

A62. In the list below, tick the THREE expressions that you think are the most strongly associated with the origins of humankind.

- ☐ Adam and Eve
- ☐ Australopithecus
- ☐ Creation
- ☐ Evolution
- ☐ God
- ☐ Natural selection

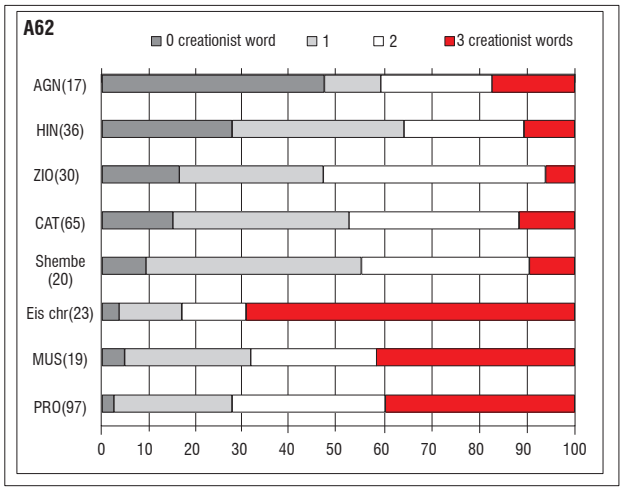


Figure 5: Percentage of teachers (grouped by religious affiliation) who selected 0, 1, 2 or 3 creationist words in Question A62 on the origin of humankind.

As shown in Figure 5, Question A62 highlighted the differences among the groups of teachers. Among Protestant, Muslim and Other Christian teachers, 40% or more selected three creationist terms associated with the origin of humankind. Fewer than 20% of teachers in all other groups chose three creationist terms. At the opposite extreme, more than 40% of atheist or agnostic teachers chose no creationist terms. Over 50% of Hindu, Roman Catholic and Shembe teachers, and about 47% of Zionist adherents, chose either none or one creationist term. Roughly 30% of Protestant and Muslim teachers, and fewer than 20% of Other Christian teachers, chose either none or one creationist term. Unexpected results were that more than 40% of the atheist–agnostic group chose two or three creationist words.

The results for the three questions – as shown in Figure 3 to Figure 5 – illustrate the strong effect that answers to Questions A64, B28 and A62 had on PC1. As mentioned earlier, PC1 was identified by PCA as the strongest component to distinguish between groups of religious affiliation along a continuum of evolutionist to creationist. The three questions consistently and clearly identify atheist and agnostic teachers as holding the most strongly evolutionist views, whereas Muslim, Protestant and Other Christians held the most creationist views. The same questions also placed Hindus, Zionists, and Shembe adherents closer to the evolutionist pole compared with Roman Catholics.

The answers given to the questions also illustrated that about 12% of agnostic and atheist teachers, 20% of Other Christians, 25% of Shembe, Roman Catholic, Protestant and Muslim teachers, and 36% of Hindu and Zionist teachers selected answers that were compatible with a theistic evolution position.

Discussion

Among our sample of South African teachers, a greater number held creationist views than their counterparts in European countries or in South Korea. Fewer South African teachers held creationist views than their counterparts in North Africa or Lebanon. The results obtained in our study differed in several respects from those of other countries sampled in the Biohead–Citizen international study. The main results from 30 countries[60] show important differences among those countries. However, within any country, with few exceptions there were no significant differences among the answers from teachers with different religious affiliations. More Protestant teachers held creationist views than their Roman Catholic colleagues in Brazil, Burkina Faso[60] and South Korea.[37] In Lebanon, slightly more Sunni teachers (but not Shiite or Druze teachers, who are also Muslim) held creationist views than their Christian colleagues.[60] The clear differences linked to religion in our study illustrate a greater diversity of views with regard to evolution than that of other countries.

In South Africa, the views of Protestant teachers appear to be more creationist than their Roman Catholic colleagues'. However, Protestants in South Africa may include adherents of Pentecostal churches in addition to mainstream Protestant churches, which could have skewed our results. Pentecostal churches believe in the inerrancy of the Bible, and one could expect adherents of these denominations to hold stronger creationist views than mainstream Protestants. The group we called 'Other Christian' included respondents who follow mostly Pentecostal and charismatic churches (as classified by Census 2001.[38]) Notwithstanding the difficulties in separating adherents of mainstream Protestant churches from those of Pentecostal or charismatic churches, we noted an almost complete overlap between Protestant, Other Christian and Full Gospel churches. This cluster comprised the broad Christian group with the strongest creationist views of all Christian groups in our study.

The result for Protestant teachers in South Africa is similar to results from some non-European countries, such as Brazil[36,61] and South Korea.[36] It is substantially different from that of European Protestant teachers, who are either Calvinist (as in France), Lutheran (as in Germany) or Anglican (as in UK).

A further finding worth noting is the proportion of teachers (between 20% and 40%) who selected the third option for questions A64 and B28, on the origin of life and humanity respectively. These responses imply a belief system that is simultaneously creationist and evolutionist – the group that Scott[52] refers to as 'theistic evolutionists'. The fact that these teachers constituted a substantial proportion of each religious affiliation in our study supports the view that among teachers with the same religious affiliation, different views are possible across the creationist–evolutionist continuum.

Teachers who hold views that are both creationist and evolutionist should find teaching the subject of evolution less problematic than teachers who hold fundamentalist creationist views. (The latter group was represented by the fourth option for question A64 and B28.) As mentioned earlier, the clearly evolutionist Dobzhansky[53] appeared to have a theistic evolutionist belief system. These results suggest that theistic evolutionists constitute a substantial group among South African teachers.

More importantly, our results illustrate the conceptions of evolution among teachers who follow religions such as Hinduism or African Independent Churches (Zionist and Shembe). Reddy[62] found that a sample of South African Hindu teachers and students experienced no conflict between their religious beliefs and evolution. As shown in Table 1, Hindu beliefs refer to several creation stories, which reduces the possibility of conflict between religious texts and acceptance of 'descent with modification'. The beliefs of Hindu teachers in this study were more evolutionist than those of their Protestant and Roman Catholic colleagues, but less evolutionist than the beliefs of agnostic or atheist teachers. This result can be compared to that observed in South Korea, where the concepts of Buddhist teachers were as evolutionist as those of their agnostic and atheist colleagues.[37] This finding can be explained by the fact that Buddhists, like Hindus, do not have a creation story as one of the central tenets of their religion.

With regard to the African Independent Churches, our results contribute to understanding many teachers' beliefs, because this is the largest church group in South Africa.[37] Although such churches are officially Christian (for instance, ZCC has a strong Pentecostal orientation), the views of adherents of these churches are more evolutionist compared with other Christian teachers'. A possible explanation is the influence of African traditional beliefs and the decreased emphasis on a literal acceptance of the Bible.

Conclusion

This research points to some important findings regarding the conceptions of South African teachers from a broad range of religious affiliations, with regard to evolution. In terms of the influence on their beliefs about evolution, religious affiliation is more important than their subject of specialisation for teaching, or their level of teaching experience. These results contribute to the Biohead–Citizen Project by collecting data for South Africa. In most of the 30 countries already investigated through the Biohead–Citizen questionnaire, there was little difference (within a country) that was related to the teachers' religions. There were a few exceptions, such as the relatively higher level of creationist views of Protestant teachers in Brazil, Burkina Faso and South Korea than their Non-Protestant colleagues. The results presented here show a great diversity in South Africa, possibly linked to a more heterogeneous socio-cultural context than in other countries. Understanding the relationship between teachers' religious affiliation and their positions on the creationist–evolutionist continuum is important if we wish to reduce the conflict teachers may experience when teaching evolution. Interventions could be designed to include pre-service teacher education programmes, and professional development programmes for in-service biology teachers.

Authors' contributions

P.C. was the project leader who initiated the project in South Africa. He was instrumental in the design of the questionnaire that was used in all countries. He contributed to statistical analysis of the data and wrote sections of the manuscript, including the results section. M.S. and A.J. identified possible participants, collected the data, wrote the introduction, background and literature review, and parts of the methods and discussion sections. E.D. adapted some of the questions for the South African context, and refined the results and discussion section.

References

1. Inter-Academy Panel. Statement on the teaching of evolution [homepage on the Internet]. c2006 [cited 2015 Apr 20]. Available from: http://www.interacademies.net/File.aspx?id-6150
2. Gould SJ. Hen's teeth and horse's toes. New York: W.W. Norton & Company; 1983.
3. Lovely EC, Kondrick LC. Teaching evolution: Challenging religious preconceptions. Integ Comp Biol. 2008;48(2):164–174. http://dx.doi.org/10.1093/icb/icn026
4. Cavallo AML, McCall D. Seeing may not mean believing: Examining students' understandings and beliefs in evolution. Am Biol Teach. 2008;70(9):522–529. http://dx.doi.org/10.2307/27669336
5. Blancke S, Boudry M, Braeckman J, De Smedt J, De Cruz H. Dealing with creationist challenges. What European biology teachers might expect in the classroom. J Biol Educ. 2011;45(4):176–182. http://dx.doi.org/10.1080/00219266.2010.546677
6. Forrest B, Gross PR. Creationism's Trojan Horse. The wedge of intelligent design. Oxford: Oxford University Press; 2001.
7. Miller D, Scott E, Okamoto S. Public acceptance of evolution. Science. 2006;313:765–766. http://dx.doi.org/10.1126/science.1126746
8. Numbers RL. The Creationists: From scientific creationism to intelligent design. Cambridge, MA: First Harvard University Press; 2006.
9. Curry A. Creationist beliefs persist in Europe. Science. 2009;323:1159. http://dx.doi.org/10.1126/science.323.5918.1159
10. Allgaier J. Scientific experts and the controversy about teaching creation/evolution in the UK press. Sci Educ. 2010;19:797–819. http://dx.doi.org/10.1007/s11191-009-9195-5
11. Großschedl J, Konnemann C, Basel N. Pre-service biology teachers' acceptance of evolutionary theory and their preference for its teaching. Evol Educ Outreach. 2014;7, Art. #18, 16 pages. http://dx.doi.org/10.1186/s12052-014-0018-z
12. Clément P. Teachers' conceptions of evolution in some Muslim countries. Pub Understand Sci. 2015;24:388-399. http://dx.doi.org/10.1177/0963662513494549
13. Hameed S. Bracing for Islamic creationism. Science. 2008;322:1637–1638. http://dx.doi.org/10.1126/science.1163672
14. Martin JW. Compatibility of major US Christian denominations with evolution. Evol Educ Outreach. 2010;3:420–431. http://dx.doi.org/10.1007/s12052-010-0221-5

15. BouJaoude S, Asghar A, Wiles JR, Jaber L, Sarieddine D, Alters B. Biology professors' and teachers' positions regarding biological evolution and evolution education in a Middle Eastern society. Int J Sci Educ. 2010;1–22.

16. Clément P. Muslim teachers' conceptions of evolution in several countries. Pub Understand Sci. 2015;24(4):400–421. http://dx.doi.org/10.1177/0963662513494549

17. Asghar A. Canadian and Pakistani Muslim teachers' perceptions of evolutionary science and evolution education. Evol Educ Outreach. 2013;6, Art. #10, 12 pages. http://dx.doi.org/10.1186/1936-6434-6-10

18. BouJaoude S, Wiles JR, Asghar A, Alters B. Muslim Egyptian and Lebanese students' conceptions of biological evolution. Sci Educ. 2011;20:895–915. http://dx.doi.org/10.1007/s11191-011-9345-4

19. Department of Education. National curriculum statement, Grades 10–12 for the life sciences. Pretoria: Government Printing Works; 2003.

20. Lever J. Science, education and schooling in South Africa. In: James W, Wilson L, editors. The architect and the scaffold: Evolution and education in South Africa. Cape Town: HSRC Press; 2002. p.10–44.

21. Sanders M, Ngxola N. Identifying teachers' concerns about teaching evolution. J Biol Educ. 2009;43:121–128. http://dx.doi.org/10.1080/00219266.2009.9656166

22. Kirsten F. Life Sciences teachers' understanding of the nature of science within the context of teaching evolution [MEd dissertation]. South Africa: University of KwaZulu-Natal; 2013.

23. Stears M. Exploring biology education students' responses to a course in evolution at a South African university: Implications for their role as future teachers. J Biol Educ. 2011;46(1):12–19. http://dx.doi.org/10.1080/00219266.2011.560171

24. Department of Basic Education. National curriculum statement: Curriculum and assessment policy, Grades 10–12 for the life sciences. Pretoria: Government Printing Works; 2011.

25. Abrie AL. Student teachers' attitudes towards and willingness to teach evolution in a changing South African environment. J Biol Educ. 2010;44(3):102–107. http://dx.doi.org/10.1080/00219266.2010.9656205

26. Kyriacou XS. A design-based study of the effect of an evolution education short learning programme on the conceptual development of participants with a view to theory building and improvement of practice [PhD thesis]. Johannesburg: University of Johannesburg; 2013.

27. Mpeta M. The influence of the beliefs of teachers and learners on the teaching and learning of evolution [PhD thesis]. Pretoria: University of Pretoria; 2013.

28. Keke B. Understanding Life Sciences teachers' engagement with ongoing learning through continuous professional development programmes [PhD thesis]. Pietermaritzburg: University of KwaZulu-Natal; 2014.

29. Kyriacou XS, De Beer J, Ramnarain U. Evolutionary ideas held by experienced South African biology teachers. Afr J Res Math Sci Technol Educ. 2015;19(2):118–130. http://dx.doi.org/10.1080/10288457.2015.1014231

30. Pillay CM. The difficulties faced by some teachers with strong religious beliefs when they teach evolution [MEd research report]. Johannesburg: University of the Witwatersrand; 2011.

31. Yalvac G. Barriers in the teaching and learning of evolutionary biology amongst Muslim teachers and learners in South African Muslim schools [MEd dissertation]. Johannesburg: University of Johannesburg; 2011.

32. Naidoo T. No problem. Naptosa–Insight. 2008;2(2):16.

33. Carvalho G, Clément P, Bogner F, Caravita S. BIOHEAD–Citizen: Biology, health and environmental education for better citizenship, final report. FP6 Priority 7, Project No CITC-CT-2004-506015. Brussels: European Commission; 2008.

34. Clément P, Quessada M-P. Les conceptions sur l'Evolution d'enseignants de différentes confessions chrétiennes dans 16 pays [The conceptions of evolution of teachers from different Christian denominations in 16 countries]. Actes 7èmes Rencontres de l'ARDIST; 2012 Mar 14–16; Bordeaux, France. p. 95–104. French.

35. Clément P, Quessada M-P, Castéra J. Creationism and innatism of teachers in 26 countries. Paper presented at: IOSTE XV 2012; 2012 October 30–November 02; Yasmine Hammamet, Tunisia.

36. Caldeira AM de A, De Araujo ESNN, Carvalho GS. Creationism and evolution views of Brazilian teachers and teachers-to-be. J Life Sci. 2012:6;99–109.

37. Seo H-A, Clément P. Teachers' views on evolution: Religion matters in South Korea. Procedia Soc Behav Sci. 2015;167:96–102. http://dx.doi.org/10.1016/j.sbspro.2014.12.649

38. Statistics South Africa (StatsSA). Census 2001: Primary tables South Africa. Pretoria: StatsSA; 2004.

39. West M. Bishops and prophets in a black city: African Independent Churches in Soweto, Johannesburg. London: Rex Collings / Cape Town: David Phelps; 1975.

40. Anderson AH. The Lekganyanes and prophecy in the Zion Christian Church. J Relig Africa. 1999;29(3):285–312. http://dx.doi.org/10.2307/1581527

41. African Independent churches [homepage on the Internet]. No date [cited 2015 Mar 05]. Available from: http://www.zapmeta.co.za.

42. Gotquestions.org [homepage on the Internet]. No date [cited 2015 Sep 22]. Available from: http://www.gotquestions.org/agnosticism.html

43. Catholics.com [homepage on the Internet]. Catholic answers to explain and defend the faith. c2004 [cited 2015 Dec 07]. Available from: http://www.catholic.com/tracts/adam-eve-and-evolution

44. Religious tolerance.org [homepage on the Internet]. Beliefs of different faith groups about origins. c1995 [cited 2015 Dec 07]. Available from: http://www.religioustolerance.org/ev.denom.htm.

45. Gotquestions.org [homepage on the Internet]. What is Islam and what do Muslims believe? No date [cited 2015 Dec 07]. Available from: http://www.gotquestions.org/Islam.html.

46. Iskon Educational Services [homepage on the Internet]. The heart of Hinduism. c2004 [cited 2015 Dec 07]. Available from: http://hinduism.iskcon.org/concepts/112.htm.

47. Hexham I, editor. The scripture of the amaNazaretha of EkuphaKamene. Calgary: University of Calgary Press; 1993.

48. Clément P. Didactic transposition and the KVP model: Conceptions as interactions between scientific knowledge, values and social practices. In: Proceedings of ESERA Summer School, IEC, Braga, Portugal; 2006. p. 9–18.

49. Clément P. Conceptions, représentations sociales et modèle KVP [Designs, social representations and model KVP]. *Skholê: cahiers de la recherche et du développement*. Marseille: IUFM de l'académie d'Aix-Marseille; 2010. p. 55–70. French.

50. Fuller FF. Concerns of teachers: A developmental conceptualization. Am Educ Res J.1969;6(2):207–226.http://dx.doi.org/10.3102/00028312006002207

51. Hall G, Loucks S. Teacher concerns as a basis for facilitating and personalizing staff development .The Teachers College Record. 1978;80(1):36–53.

52. Scott EC. Antievolution and creationism in the United States. Ann Rev Anthropol. 1997;26:263–289.

53. Dobzhansky T. Nothing in biology makes sense except in light of evolution. Am Biol Teach. 1973;35:125–129.

54. R_Development_Core_Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2006.

55. Dolédec S, Chessel D. Rythmes saisonniers et composantes stationnelles en milieu aquatique, II – Prise en compte et élimination d'effets dans un tableau faunistique. [Seasonal rhythms and components stational aquatic II – Recognition and elimination of effects in a table of fauna]. Acta OEcol Gen. 1989;10:207–232. French.

56. Sabatier R, Lebreton JD, Chessel D. Principal component analysis with instrumental variables as a tool for modelling composition data. In: Coppi R, Bolasco S, editors. Multiway data analysis. Amsterdam: Elsevier; 1989. p. 341–352.

57. Romesburg H-C. Exploring, confirming, and randomization tests. Comp 7 Geosc. 1985;11(1):19–37. http://dx.doi.org/10.1016/0098-3004(85)90036-6

58. Munoz F, Clément P. Des méthodes statistiques originales pour analyser les conceptions d'enseignants de plusieurs pays à partir d'un questionnaire sur des questions vives [Original statistical methods to analyze conceptions of teachers from several countries using a questionnaire on controversial issues]. Strasbourg: Actes Colloque AREF (Actualité de la Recherche en Education et en Formation); 2007. p. 1–11. French.

59. Munoz F, Bogne F, Clément P, Carvalho GS. Teachers' conceptions of nature and environment in 16 countries. J Env Psych. 2009;29:407–413. http://dx.doi.org/10.1016/j.jenvp.2009.05.007

60. Clément P. Creationism, science and religion: A survey of teachers' conceptions in 30 countries. Procedia Soc Behav Sci. 2015;167:279–287 http://dx.doi.org/10.1016/j.sbspro.2014.12.675

61. El-Hani CN, Sepulveda C. The relationship between science and religion in the education of protestant biology preservice teachers in a Brazilian university. Cultural Stud Sci Educ. 2010;5(1):103–125. http://dx.doi.org/10.1007/s11422-009-9212-7

62. Reddy C. The lived experiences of Hindu teachers and learners in the teaching and learning of evolution in life sciences in the FET phase [MEd dissertation]. Johannesburg: University of Johannesburg; 2012.

Climate change threats to two low-lying South African coastal towns: Risks and perceptions

AUTHORS:
Jennifer M. Fitchett[1]
Bronwyn Grant[1]
Gijsbert Hoogendoorn[2]

AFFILIATIONS:
[1]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa
[2]Department of Geography, Environmental Management and Energy Studies, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Jennifer Fitchett

EMAIL:
Jennifer.m.fitchett@gmail.com

POSTAL ADDRESS:
School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Private Bag 3, Wits 2050, South Africa

DATES:
Received: 17 July 2015
Revised: 15 Oct. 2015
Accepted: 19 Jan. 2016

KEYWORDS:
digital elevation model; perceptions; tourism; tourism climate index

HOW TO CITE:
Fitchett JM, Grant B, Hoogendoorn G. Climate change threats to two low-lying South African coastal towns: Risks and perceptions. S Afr J Sci. 2016;112(5/6), Art. #2015-0262, 9 pages. http://dx.doi.org/10.17159/sajs.2016/20150262

Climate change poses a considerable threat to low-lying coastal towns. Possible risks include flooding induced by sea-level rise, increased discomfort from changes in temperature and precipitation, more frequent extreme events, biodiversity shifts, and water shortages. For coastal towns that attract many tourists, these threats can have far-reaching economic effects and may compromise the continued viability of the tourism sector. A growing number of studies are being published on the inter-relationship between climate change and tourism in the global North. As yet, little equivalent research has been conducted in developing countries with economically significant tourism sectors. This paper presents a mixed-method pilot study on two adjacent coastal towns, St Francis Bay and Cape St Francis, in the Eastern Cape Province of South Africa. We explored the climate change threats in this region, and perceptions of these threats within the tourism sector. The tourism climate index results showed that the towns are climatically well suited to tourism, but a decrease in these index scores between 1978 and 2014 suggests that climate change experienced in recent decades has detrimentally affected tourist comfort. A digital elevation model sea-level projection for the towns indicated a high risk of sea-level induced flooding by 2050, particularly for properties along the coastline. Interviews with tourism establishment respondents showed that people are aware of climate change threats, yet little adaptation is forthcoming. Rather the government is deemed responsible for adaptation, despite its limited capacity. A disjuncture therefore exists between the perceived severity of risk and the risk that is evident from scientific analyses. This gap results in poor planning for the costs associated with adaptation.

Introduction

The success of tourism in a particular location depends largely on climatic conditions.[1] Climate is commonly used as a descriptive factor when marketing destinations, and thus selection between locations often includes consideration of their relative climate suitability for preferred tourism activities.[2] Intra-annual patterns in temperature, cloud cover, wind, and rainfall influence the length and timing of peak tourist seasons.[3] Climate is also one of the factors that control the availability of activities, and it influences the overall tourist satisfaction with a destination.[4-6] Climate change is increasingly identified as a significant threat to tourism, because of the reliance of tourism on predictable contemporary climatic conditions and seasonality.[6,7] Climate change is likely to affect the seasonality of tourism, and may threaten the infrastructure and natural environment required to accommodate and entertain tourists.[8] Predictable and ameliorable weather is critical to the feasibility of outdoor activities.[9]

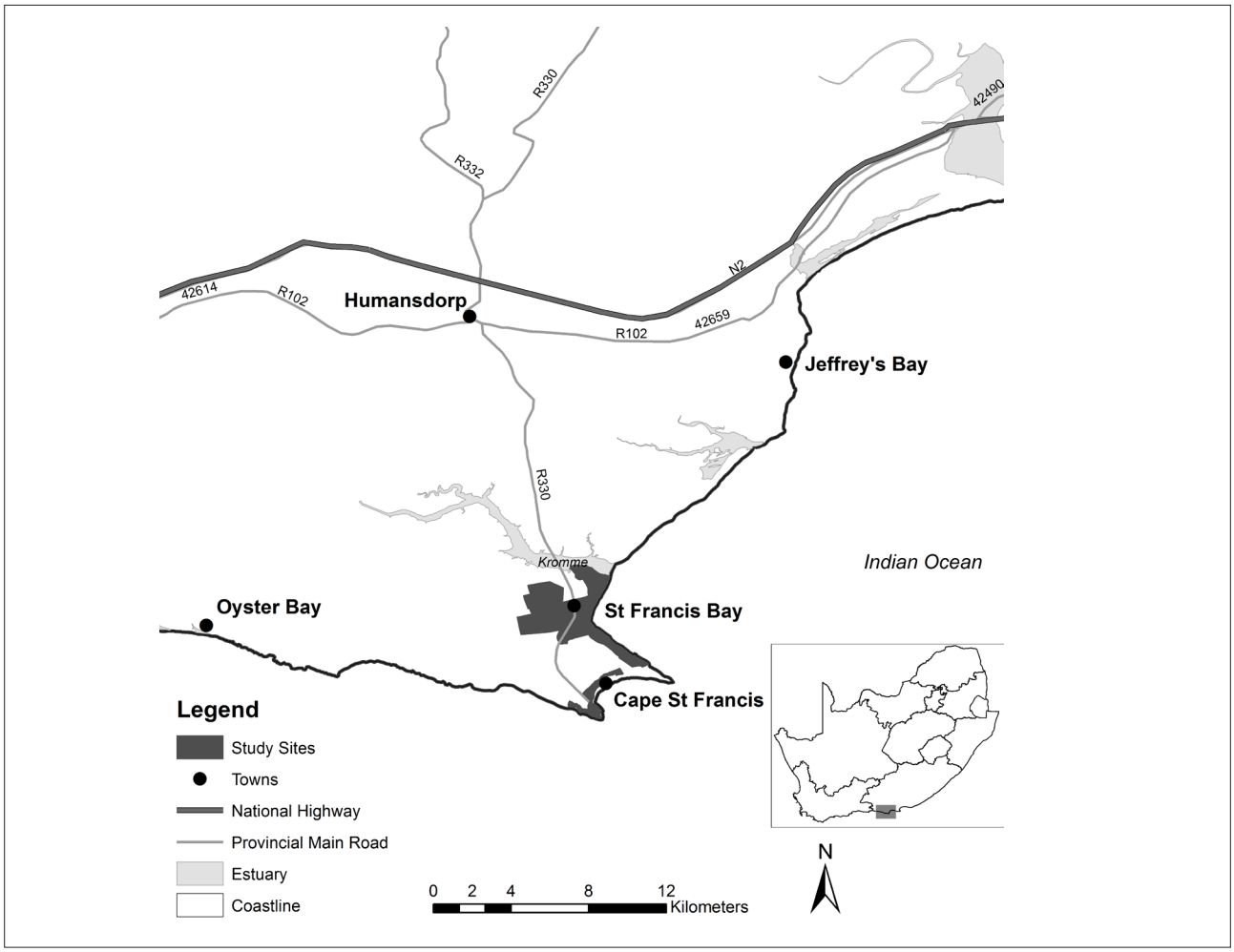
The increased variability in weather projected under climate change scenarios, together with associated risks of sea-level rise, render coastal destinations particularly vulnerable.[9] It is predicted that by 2050, the global tourism sector will experience substantial climate change impacts, with fluctuations in tourist numbers as tourists opt for better-suited locations.[10] The issue of climate change is therefore being integrated into sustainable tourism development initiatives and research.[11,12] To attain long-term sustainable development, the tourism sector must adapt to a range of projected climate change impacts.[13]

In the global South, a lack of integration between climate change and tourism research is of concern because it limits future adaptation potential.[14] 'Adaptation' can be defined for the purposes of this study as '…the ability of a unit (e.g. a tourism operator or a community) to transform its structure, operations or organisation to survive under changes threatening its existence and success'[15].

South Africa has a growing tourism sector, which contributes considerably to the local economy.[16] The country has a wide range of tourism attractions, many of which rely on the comparatively good climate of the region,[17] characterised by warm weather and clear skies. A large proportion of the sector involves beach and nature based tourism, which rely heavily on continued ameliorable climatic conditions.[18] In addition, South Africa has numerous small towns that are largely economically dependent on tourism. Although there is some variation in tourist offerings, many involve outdoor attractions – including beaches, game drives, water sports and fishing.[19]

South Africa is predicted to experience temperature increases as high as 4 °C by the year 2100, as well as changes in the timing, amount, and severity of precipitation; changes in wind direction and strength; and rising sea levels.[20,21] Given that tourists are highly mobile, even a few weeks of poor weather during the tourist season can affect the long-term viability of the sector in terms of destination attractiveness.[22] Information on climatic conditions is easily accessible for a range of possible tourist locations, both locally and internationally, and tourists inevitably select destinations that are relatively less risky in terms of the weather.[23] Climate change therefore has the potential to economically cripple tourist establishments and activities, as well as tourism-led towns – and indeed the South African tourism economy – if sufficient adaptation is not achieved.

It is insufficient to implement adaptation measures only after climate change impacts become evident; forward planning is essential.[24,25] Policy documents indicate that such planning and adaptation to climate change threats are not forthcoming in the study region. We explored these climate change threats and people's perceptions within the tourist sector regarding the level of risk, to determine the levels of preparedness. This study is of value in determining the likely impacts of climate change on the tourism sector in the study areas, and broader patterns of response in terms of climate change adaptation in the tourism sector regionally.



Source: Authors' map drawn for the study

Figure 1: Map indicating the location of study sites.

Study site

Our study examined two adjacent coastal towns, St Francis Bay and Cape St Francis, in the Eastern Cape Province of South Africa (Figure 1). These towns provide examples of tourism-reliant locations that face multiple climate change threats. Located on a gently sloping coastal plain, the towns are at risk of sea-level rise, which could encroach on the beach and eventually inundate much of the developed area. This area would include tourism establishments and roads and bridges, prohibiting access to the towns.[26,27] Changes in temperature, rainfall and wind are also of concern. Both towns provide predominantly outdoor tourist attractions and have distinct peak tourist seasons associated with warm temperatures and clear skies.[20] Additional concerns for these towns include the Kromme River Estuary and Canal, which pose flood risks, and the position of nearby sand dunes, which could result in dust storms.[20,27]

St Francis Bay covers a land area of 10.04 km^2 at an altitude ranging from 0 masl to 34 masl.[28] Cape St Francis comprises a smaller area of 4.38 km^2 at altitudes from 0 masl to 15 masl.[28] Census 2011 data give the population of St Francis Bay as 4933 people, and 342 in Cape St Francis.[29]

A series of canals and waterways, ideal for water sports, attracts tourists to St Francis Bay. A large number of holiday homes in both towns are rented out to tourists. Many are located along the canal system, an area of St Francis Bay commonly known as 'Little Venice'.[30] A nature reserve connects the two towns and provides a popular eco-tourism destination.[31]

The towns are situated in a zone that receives year-round rainfall, although more rain falls during the winter months.[32,33] Mean air temperatures range from 12.3 °C to 14.2 °C in winter and 18.9 °C to 20.8 °C in summer.[32,34] There is marked seasonality in wind trajectories. In winter, the wind blows from west to southwest, driven by the northward trajectory of the westerly belt, and in summer it blows easterly to southeast, when atmospheric circulation is dominated by the tropical easterlies.[33,35]

Methodology

We adopted an interdisciplinary mixed-method approach in this study. We wanted to determine the climate change risks to the towns, and to explore accommodation establishment proprietors' perceptions of those threats.

Climate change risks to the towns were classified on the basis of the towns' proximity to the sea and the reliance of tourism on ameliorable weather to encourage outdoor activities. The risk of sea-level rise was determined using a digital elevation model (DEM) with overlaid sea-level projections. The climatic suitability of the region for tourism was determined using the tourism climate index (TCI).

We probed the perceptions of accommodation establishment proprietors through individual face-to-face interviews in which open-ended questions were posed, and respondents were also asked to rank their concerns on a Likert-type scale. Qualitative data from these interviews were interpreted using thematic analysis. All respondents were presented with a consent letter before the interview commenced, indicating the purpose of the study and assuring them of confidentiality. The ethics requirements of the University of the Witwatersrand were adhered to prior to the commencement of this study.

Table 1: Components of tourism climate index (TCI)

Sub-index	Climatic variable	Influence on TCI	Weight (%)
Daytime Comfort Index (CID)	Maximum daily air temperature (°C) Minimum daily relative humidity (%)	Thermal comfort when maximum tourist activity occurs	40
Daily Comfort Index (CIA)	Mean daily air temperature (°C) Mean daily relative humidity (%)	Thermal comfort over 24-hour period including night time	10
Precipitation (R)	Total precipitation (mm)	A negative factor on overall experience	20
Sunshine (S)	Total hours of sunshine (hours)	A positive factor on overall experience	20
Wind (W)	Average wind speed (m/s)	Dependent on air temperature: evaporative cooling effect in hot climates; wind chill in cold climates	10

Source: after Perch-Nielsen et al.[48]

Digital elevation model

To model sea-level rise for St Francis Bay and Cape St Francis, and to plot regions likely to be affected by consequent flooding, a DEM for the region was created. The input data were acquired from Viewfinder Panoramas using the DEM World Coverage Map (2014). Raw data were acquired from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), an imaging instrument on board the Terra satellite.[36,37] ASTER Global DEM Version 2, produced at a 3 arc-second resolution (~90 m), was used.

To perform an initial sea-level rise vulnerability assessment, we used a simple inundation model known as the 'bathtub approach'.[38] This method maps land that is adjacent to the sea and is situated below a given elevation threshold to project inundated areas.[39] The mapping process uses a deterministic line to indicate inundation, relative to the sea-level rise and ground elevation.[40] This approach does not account for coastal dynamics including wind, waves or tide, and therefore excludes the heightened risk of storm surges.[40,41] However, it facilitates the visual identification of locations at highest risk for flooding directly from sea-level rise, and provides a projection of the 'best-case scenario'.

The inundation levels were created using the raster calculator function in the spatial analyst tool in ArcGIS®. Sea-level rise scenarios were projected for the years 2050 and 2100, using sea-level rise projections of 0.4 m for 2050 and 1.6 m for 2100,[42-45] based on an average rate of change of 0.3 mm/year along the south coast of South Africa.[46] This tested the Boolean statement at each cell of the DEM (interrogating whether the value within the cell was less than or equal to 0.4 m and 1.6 m, respectively).

Tourism climate index

The TCI is a numerical model that was developed to evaluate and compare tourist destinations, based on the suitability of their climates for tourists.[47,48] Daily climate data for the closest registered weather station, Port Elizabeth, were sourced from the South African Weather Service for the period 1978–2014 (the total period for which complete datasets were available).

To calculate the TCI, five main sub-indices were used, comprising seven climatic factors[47] (Table 1). These factors were each given a weighting (based on work by Perch-Nielsen et al.[48]) and each climate variable was rated.[48] We then used the standard TCI formula[47,48]:

$$TCI = 2(4CID + CIA + 2R + 2S + W) \quad \text{Equation 1}$$

Mean annual and mean monthly TCI scores were calculated for the period 1978–2014, based on mean climate variables. This calculation determined the long-term climatic suitability of the region, whilst accounting for inter-annual climate variability for the southwest Indian Ocean because of El Niño Southern Oscillation, Quasi-Biennial Oscillation, and the Indian Ocean Dipole.[49] In addition, changes in annual TCI scores were explored for the period 1978–2014 using linear regression, to detect progressive changes in the climatic suitability for tourism. The TCI scores were classified according to climate suitability (Table 2).

Table 2: Rating categories of the tourism climate index

TCI Score	Descriptive Category
90-100	Ideal
80-89	Excellent
70-79	Very good
60-69	Good
50-59	Acceptable
40-49	Marginal
30-39	Unfavourable
20-29	Very Unfavourable
10-19	Extremely Unfavourable
< 10	Impossible

Source: Mieczkowski[47]

Interviews

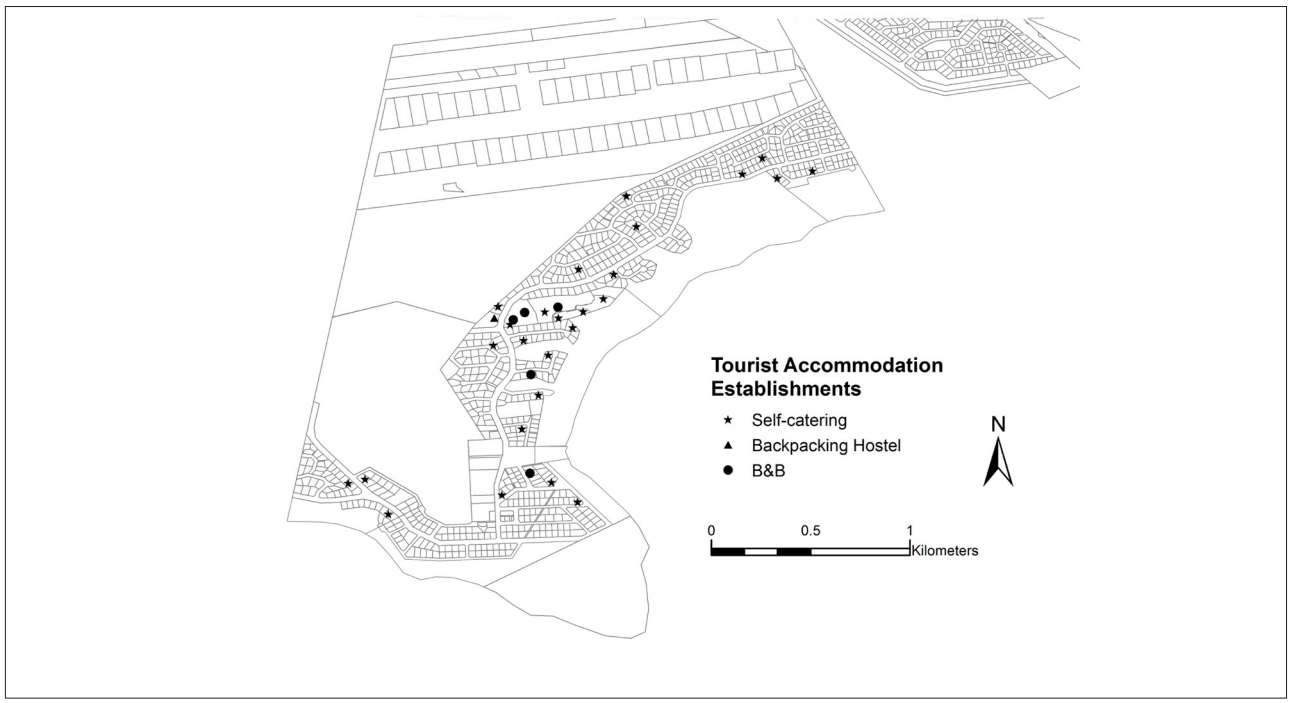
Semi-structured interviews with respondents, who were all tourism accommodation proprietors, were conducted in St Francis Bay and Cape St Francis. During the period of fieldwork, 57 establishments existed in St Francis Bay (Figure 2) and 31 in Cape St Francis (Figure 3).

Among these 88 establishments, we interviewed 36 respondents in St Francis Bay and 17 respondents in Cape St Francis (a total of 53 people). Interviews were conducted during the off-season period of September 2014. The interview schedule included questions relating to the proprietors' understanding of climate change, their concerns regarding climate change risks to their establishments, their experiences of weather-related damage in recent years, and their climate change adaptation strategies. 'Perception' is defined in this study as 'the process of interpreting sensory information received through environmental stimuli and actions taken in responding to stimuli'[50].

Accommodation establishments were contacted telephonically prior to the interview to arrange a suitable time, and the interviews were later conducted in person. The interviews were transcribed and coded according to key themes. The frequency of the appearance of these key themes was then explored in relation to the geographic locations of each of the accommodation establishments, mapped using ArcGIS®.

Source: Authors' map drawn for the study

Figure 2: Map indicating the locations and types of tourism accommodation establishments in St Francis Bay.



Source: Authors' map drawn for the study

Figure 3: Map indicating the locations and types of accommodation establishments in Cape St Francis.

Results and analysis

Digital elevation model sea-level rise projection

Many accommodation establishments are situated at low elevations along the coastline. Sea-level rise is thus a particular concern for the towns we studied, both directly through the flooding of accommodation establishments and indirectly through damage to local infrastructure, which impedes access for tourists. The results of our DEM predict a considerable reduction of the beach area, with extensive coastal squeeze, by 2050, with the worst effect being predicted for the Sea Vista area of St Francis Bay (Figure 4).

Sea-level rise projected for 2050 is modelled to result in the permanent opening of the Kromme River Estuary, with a resultant increase in the salinity of the estuary and a heightened flood risk for the Kromme River. The artificial spit which currently protects the canal system from salt water incursion is projected to be inundated by 2050, compromising freshwater supply to the towns. By 2100, the DEM projects inundation of low-lying regions of the two towns, including areas in which tourism accommodation establishments are located. The Marina Glades area is projected to be at the highest risk area of inundation by 2100.

Our DEM projections indicate that the Santareme area of St Francis Bay has the lowest threat of inundation. Of direct relevance to the tourism accommodation establishments, despite the projected high risk for Sea Vista, only one establishment is located in that region. By contrast, Marina Glades has 12 establishments that may be affected; however, only two of these are at high risk as they are located near the spit.

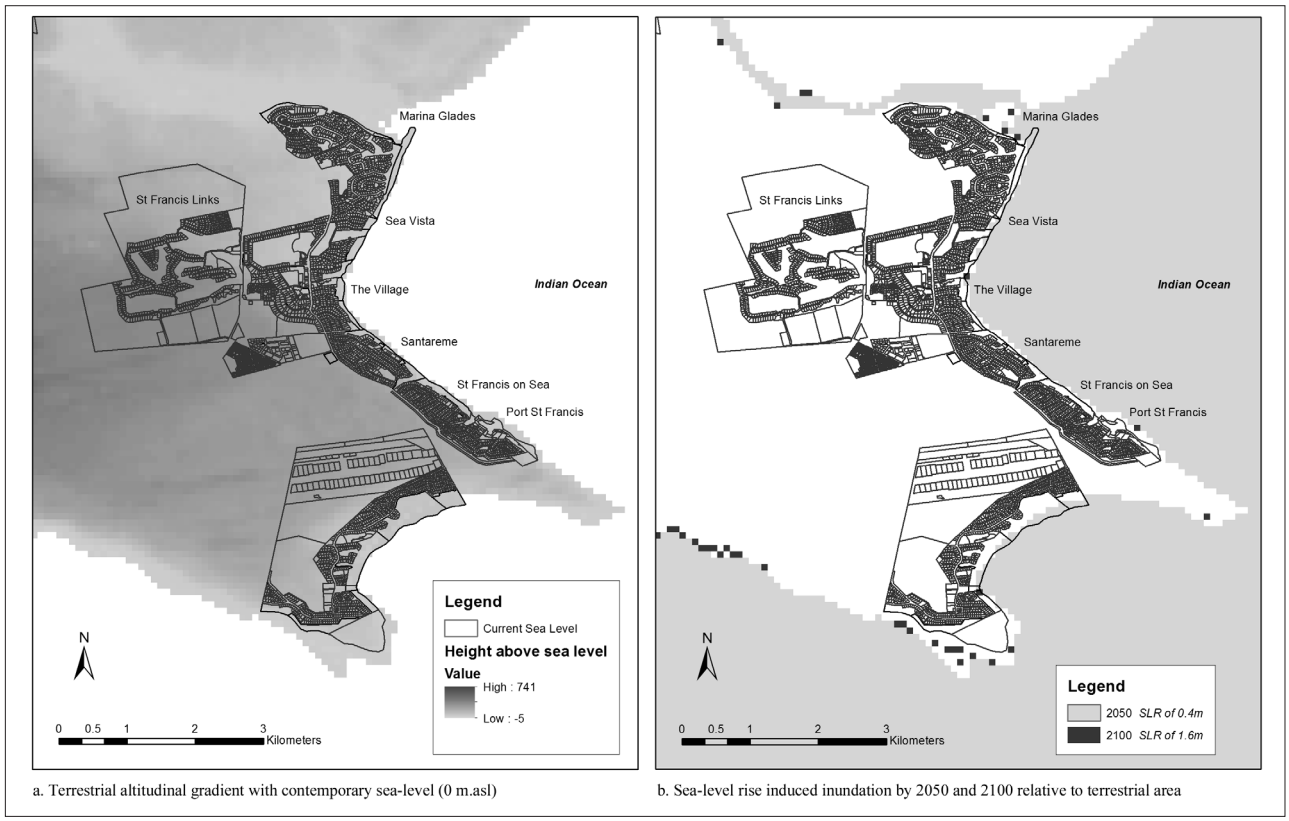
Cape St Francis is similarly projected to lose almost half of its beach area by 2050 owing to coastal squeeze. The town is situated relatively further inland, thus none of its accommodation establishments appear to be at risk under the 2050 sea-level rise projections. DEM sea-level projections for 2100 suggest partial inundation of the town, centred behind the Seal Point Headland. Three accommodation establishments are located behind this headland, and according to our projections are likely to be affected. Our DEM projections also indicated that by 2100 the beach will be completely inundated, which would remove the primary tourist attraction of the town. It would also heighten flood risk for the 11 beachfront accommodation establishments.

Sea-level rise additionally results in an increased risk of storm surges for low-lying areas.[25,51] Furthermore, climate change projections indicate changing wind patterns, and a southward trajectory of mid-latitude cyclones, both of which would further heighten the risk of storm surges.[20,49] Storm surges pose the greatest threat to coastal accommodation establishments, but severe events could also result in considerable damage to town infrastructure. Flooding related to storm surges has already resulted in considerable damage to St Francis Bay during 1996[26] and 2007[52]. Hence, our DEM projections likely provide an underestimate of the flooding potential for the region and can be regarded as a 'best-case' scenario.

Tourism climate index scores

The annual TCI scores that we calculated for the region over the period 1978–2014 ranged from 73 to 86. These scores represent 'very good' to 'ideal' climatic suitability for tourism, respectively. The long-term annual average TCI score for the period 1978–2014 was 80, categorised as 'excellent' (see Table 2). An identical score was obtained for the short-term average for the period 2009–2014.

The monthly TCI scores for the region, averaged between 1978 and 2014, peaked during the local summer months (December to February). This pattern is consistent with the accommodation establishments' recorded periods of peak tourist arrivals in the two towns (Figure 5). The pattern is notable from a climatic perspective, as the region is located within the year-round rainfall zone and is thus less likely to experience a unimodal distribution.[53] The variation does, however, occur within a small range of TCI scores, all categorised as either 'very good' or 'ideal'.



a. Terrestrial altitudinal gradient with contemporary sea-level (0 m.asl)

b. Sea-level rise induced inundation by 2050 and 2100 relative to terrestrial area

Figure 4: Digital elevation model illustrating predicted sea-level rise for both towns for the years 2050 and 2100. (a) Terrestrial altudinal gradient with contemporary sea level (0 masl) and (b) sea-level rise (SLR) induced inundation relative to terrestrial area.

At both the annual and monthly scales, the daytime comfort index (CID) and wind achieved consistently high scores. The CID reflects the maximum daily temperature and minimum daily relative humidity. Both these variables thus contributed positively to the high TCI scores, whereas the effects of rainfall and sunshine hours resulted in more moderate TCI scores (i.e. closer to the central tendency).

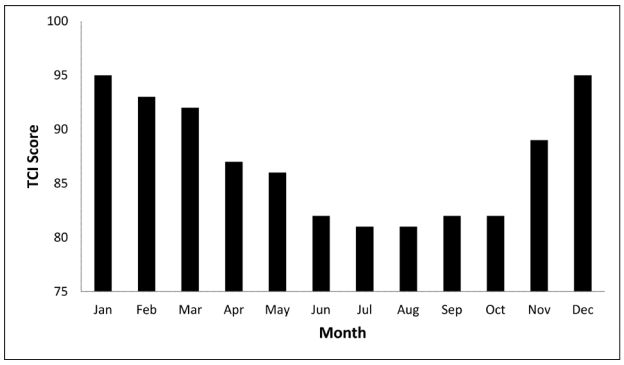


Figure 5: Mean monthly tourism climate index (TCI) scores for 1978–2014 for the region (St Francis Bay and Cape St Francis).

As evidenced by the equal TCI scores for the 35-year and 5-year calculation periods, there has been no considerable change in the climatic suitability of the region we studied. Over the period 1978–2014, the inter-annual variability in TCI score was more pronounced than any consistent long-term trend. Notably low TCI scores were evident for 1996, 2011 and 2012, and particularly high scores were evident in 1987 and 2013 (Figure 6). A weak trend towards lower TCI scores for the region, at a drop of 0.3 units per decade, was observed.

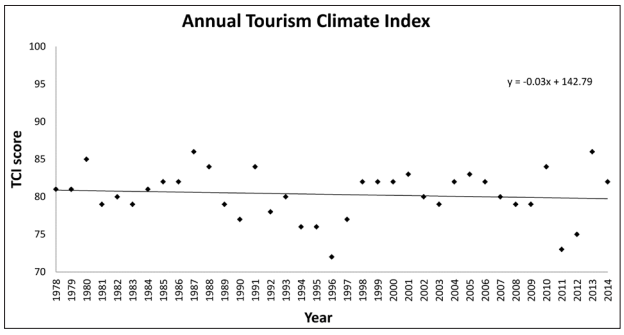


Figure 6: Mean annual TCI scores for 1978–2014, indicating long-term trend.

The climate suitability of the region for tourism is therefore not as great a threat as sea-level rise. However, the trend of a slow drop in TCI scores indicates the importance of continual monitoring of changes in meteorological variables in the region. This would help in making adaptation plans to address these changes before they detrimentally affect the region's ability to attract tourists.

Climate change perceptions

The interview data showed that 35 proprietors of tourism accommodation establishments believed that climate change was a cause for concern. Among all 53 respondents, 42 proprietors believed they were already experiencing the effects of climate change, and the remaining 11 felt these effects will be experienced only in 100 years' time. Respondents from establishments located closest to the coast showed more concern for climate change than those further inland.

Respondents from establishments situated along the Kromme River estuary and in the Marina Glades canal system had a medium to high level of concern regarding the threat of climate change. This finding was understandable given the risk of these water bodies flooding. Respondents whose establishments were relatively protected from possible coastal flooding reported fairly low levels of concern. They were situated at a higher altitude than the rest of the town, with large boulders along the beach which could provide defence against flooding. Thirteen respondents, with accommodation establishments scattered across the study area, did not believe that climate change would negatively affect their establishments.

More than half of the respondents (35) believed that climate change would negatively affect the towns themselves. Many of these respondents had been affected by or been within close proximity to the fires which had destroyed numerous houses on the canals in 2012. Some had also been affected by floods that occurred near the St Francis Links Golf Course in 2011. Owners of properties located close to the coast in Cape St Francis largely agreed that the town is vulnerable to the negative effects of climate change.

Fourteen respondents had experienced damage to infrastructure as a result of extreme weather events during the past 5 years. These events included a bridge in St Francis Bay being washed away four times within a 2-year period, compromising access to some of the tourism establishments. The establishments that had experienced flooding were located along the Marina Glades canal system and along the coast near the Santareme area of St Francis Bay. Cape St Francis respondents had also experienced flooding along the main road, which washed away the Sand River Bridge, hindering access to the town. Certain establishments had experienced direct problems from heavy rainfall, with water leaks causing damage to their accommodation rooms. The majority of respondents (44) felt that climate change would affect their establishments in the future. Most respondents mentioned flooding, with the threat of bridges being frequently washed away. This raised concerns about a loss of revenue.

Sea-level rise was perceived to be the largest threat for beachfront properties, with owners fearing flood damage or the infrastructure collapsing entirely if foundations become damaged. Beach erosion was mentioned, with respondents worried about coastal retreat. At the town level, concerns about infrastructure were centred on damage to roads. Overall, respondents said that the greatest impact of climate change would be a decrease in the number of guests and changes in reservation patterns. Concern was raised by respondents that climate change would lessen the ability of the towns to offer tourism activities, including fishing and hiking, which would reduce their competitive advantage relative to other coastal towns.

When asked who should be responsible for mitigating the negative effects of climate change, many respondents felt that government should be responsible for the future provision of climate change adaptation plans. They felt these plans should include the provision of information about climate change adaptation and covering the costs involved. They believed that the lack of information from local government was an indication of the minimal extent of climate change risks. A number of respondents were concerned about hotter, more humid weather, and had installed air-conditioning systems to make their establishments more comfortable for tourists under such conditions. The promotion of tourism during winter months or during adverse weather conditions was being pursued by some establishments, through increasing the range of indoor activities available, such as lounge areas with digital satellite television, DVD players, Wi-Fi, and a selection of books and board games.

Although respondents demonstrated a good understanding of climate change threats, they appeared to underestimate the severity. None of the respondents mentioned relocating their business in response to sea-level rise. Rather, sea-level rise was classified as a town-level issue, to be addressed by government through the use of sea defences, including *dolosse* (interlocking concrete blocks) in the port of St Francis Bay. Some respondents believed that the effects of climate change are distant, and that they would no longer be alive by the time such change occurs. Others believed that adaptations would only become necessary when climate change directly affects their businesses. This view was interesting to us, because of the regular flooding which participants said they have already experienced, with considerable damage to individual establishments. However, such opinions might be influenced by the high cost of infrastructural change, or a belief that insurance will pay for flood damage.

Discussion

Our research identified climate change threats and risks to coastal tourism, and compared this reality with the perceptions of proprietors of tourism accommodation establishments. We examined the vulnerability of the coastal area to climate change in the context of the broader tourism sector. Our DEM sea-level rise projections indicated considerable risk to beach attractions in both towns by 2050, and risk to the infrastructure and to accommodation establishments in the towns by 2100. Accommodation establishments located closest to the coastline or inland water are at the highest risk, although a combination of natural and man-made barriers protects certain properties within these areas.

The TCI results confirmed that the climate of the region is presently very well suited to tourism, particularly during the summer months. The small range of scores stems from the use of the international TCI model, which does not account for the relatively ideal climates throughout South Africa compared with much of the northern hemisphere. The greatest detriment to the TCI scores are rainfall and cloud cover, both of which predominate in the winter months. Notably, temperatures during winter do not detract from the TCI score. However, a weak downward trend in scores from 1978 to 2014 suggests that climate change might already be altering the climate suitability of the region for tourism. This trend is predominantly driven by changes in precipitation in the region over the 35-year period.

Although summer temperatures are currently within the bracket of climate suitability, continued increase would further detriment the TCI scores for the region if the comfort threshold is exceeded.[53] Because of the towns' proximity to the ocean, as well as warm water from the Agulhas current retroflection[54] and a large area of inland water,[55] such temperature increase would be associated with a proportional increase in humidity. This would further decrease the region's TCI scores.

Comparing the perceptions of respondents to the results of our TCI and DEM analyses, a low level of concern amongst respondents regarding sea-level rise is notable. Based on our findings and other studies,[24] we strongly recommend that accommodation establishment proprietors in St Francis Bay and Cape St Francis should consider implementing their own mitigation measures as a matter of urgency.

Little information on climate change adaptation plans for the study sites is available from the local municipality.[56,57] This lack of concern bodes ill for future adaption and mitigation, particularly given the perception by respondents that government is responsible for and will provide adaptation that is sufficient to protect the local tourism sector from the threat of climate change.[58,59]

Respondents were particularly concerned about changes to the environmental comfort of their patrons. This finding is interesting, as the DEM and TCI results indicate that sea-level poses a greater threat in a shorter time than changes in the climatic suitability of the region for tourism. Respondents acknowledged that sea-level rise may pose a threat, but they did not perceive the immediacy or extent of the problem – namely that their establishments could be severely damaged or entirely destroyed within the next 50 to 100 years.[60,61] Furthermore, many respondents said that in 50 to 100 years they will no longer be active in the tourism sector and therefore this is not an issue they will personally face.

Little consideration is made for the long-term sustainability of the tourism sector in the region. Respondents appear to be making cost–benefit assessments regarding the implementation of adaptation measures, by comparing the cost of specific forms of mitigation against the probability of a climate threat occurring, whilst considering additional benefits or detriments of such measures should the climate threat not be realised. They demonstrate willingness to adapt to changing meteorological conditions by installing air-conditioning to improve comfort during hot days, and satellite television and Wi-Fi to provide alternative activities on increasingly frequent rainy days. These climate events are more tangible than sea-level rise, given the recent occurrence of heat waves, cold fronts, and rainy days. Such adaptation measures also improve the amenities of establishments under any climatic conditions, which will afford visitors greater relaxation and business facilities. Finally, they are relatively low-cost adaptations. Therefore, the proprietor is likely to recoup the costs of such adaptation measures over a short period, with no harm should climate change not occur in the projected period.

Our study showed that the respondents were not yet investing in infrastructural changes to address flooding induced by sea-level rise. Although floods have occurred in the recent past, they were discrete events that are often assumed to be freak accidents rather than indicative of progressive change. Owing to substantial fluctuations in the return rates of flood events and the heightened variability under climate change, there is no certainty that a flood could occur the next year – or even in the next 10 years. Therefore, the cost outlay to make infrastructure changes might not be offset by a direct or immediate benefit.[25] Infrastructural changes, such as building retaining walls, are considerably more expensive than installing satellite television, with potentially no payback for the current proprietor. Furthermore, unlike satellite television and Wi-Fi, which contribute positively to the establishment under current conditions, infrastructural changes could negatively affect the amenities of an establishment by blocking sea views. Finally, as discussed above, many respondents said the government is responsible for infrastructural changes to prevent flooding induced by rising sea levels. Such action by the state would render personal retaining walls and other infrastructural adaptations unnecessary and would mean poor strategic spending by individuals.

Although such decisions can be understood from the perspective of proprietors' perceptions of the contemporary climate, they are not consistent with modelled climate change risks for the region. The results of our DEM sea-level rise projections indicate that sea water incursion, with further flood risk from storm surges, is likely to occur by 2050. This requires urgent action to improve infrastructure to adapt to the rising sea levels. Climate models for the region[20,21] and the results of our TCI time-trend analysis suggest that decreases in climatic suitability of the region will take longer to be detected by tourists, and thus require less urgent adaptation. Although the adaptation measures that the respondents are already implementing are relatively cheap and provide immediate additional benefits, they are not a substitute for infrastructural adaptation. They would be insufficient if an establishment is flooded, if the beach retreats, or if the arterial roads are flooded and this prevents tourist access.[24]

Conclusion

This research quantified the nature and extent of two predominant climate change threats and risks to the tourism sector in St Francis Bay and Cape St Francis. These main threats are sea-level rise and climatic suitability for tourism. Climatic suitability for tourism was explored over the historical period of the three most recent decades, using the TCI, with future trends inferred on the basis of climate forecasts. For sea-level rise, the threats were predicted using DEMs. The results showed that sea-level rise poses a serious threat to sustained tourism in the region, whereas a drop in the climatic suitability of the region is a less immediate concern.

Tourism accommodation establishments in the region have made small-scale adaptations to these threats, but these predominantly focus on climatic suitability. The reasons for this approach are the relatively low cost of such adaptations and the perception that government is responsible for larger-scale infrastructural changes to mitigate the negative effects of sea-level rise. Many respondents said they will not witness in person the future negative effects of climate change in the region. In addition, none appeared to be considering relocating because of any current risks or threats. Should mitigatory actions not be undertaken sufficiently by government and tourism accommodation establishments, the tourism sector in the region will likely be negatively affected. The current adaptation measures are inadequate to protect the beaches and prevent infrastructural damage. Over time, this will result in individual establishments losing their competitive advantage to regions with a lower flood risk or those with more robust and proactive adaptation plans.

South Africa has a varied climate and natural environment, attracting a range of tourists with a large spectrum of interests. The findings

of this pilot study provide initial insight into climate change risks and perceptions in low-lying coastal towns that rely heavily on beach and outdoor attractions. Although inland tourism destinations experience different climate change threats, the ambivalence of respondents to climate change threats in St Francis Bay and Cape St Francis might well be shared. In general, the sustainability of the tourism sector in South Africa relies on adequate and timely adaptation to climate change threats, so as not to lose tourists to destinations that become better climatically suited. Such adaptation requires an understanding of the nature and extent of climate change threats, and appropriate measures to protect individual establishments and towns. There is an urgent need for further studies in this line, especially given the dominance of the tourism sector in the South African economy. Our pilot study provides a valuable methodological framework for continued research.

Authors' contributions

J.F., B.G. and G.H. conducted the research. J.F and G.H. conceptualised the project and developed the experimental design. J.F. compiled and wrote the sections on the digital elevation model and tourism climate index methods and results, and compiled the first draft. G.H. wrote the section on perceptions. B.G. collected the raw data.

References

1. Hernandez AB, Ryan G. Coping with climate change in the tourism industry: A review and agenda for future research. Journal of Tourism and Hospitality Management. 2011;17(1):79–90.
2. Berritella M, Bigano A, Roson R, Tol RSJ. A general equilibrium analysis of climate change impacts on tourism. Tourism Management. 2006;27:913–924. http://dx.doi.org/10.1016/j.tourman.2005.05.002
3. Valls JF, Sarda, R. Tourism expert perceptions for evaluating climate change impacts on the Euro-Mediterranean tourism industry. Tourism Review. 2009;64(2):41–51. http://dx.doi.org/10.1108/16605370910963518
4. Becken S. Harmonizing climate change adaptation and mitigation: the case of tourist resorts in Fiji. Global Environmental Change. 2005;15(4):381–393. http://dx.doi.org/10.1016/j.gloenvcha.2005.08.001
5. Kyriakidis A, Felton J. Too hot to handle? The hospitality industry faces up to climate change. In: Blanke J, Chiesa T, editors. The Travel & Tourism Competitiveness Report. Geneva: World Economic Forum; 2008.
6. Gössling S, Scott D, Hall CM, Ceron JP, Dubois G. Consumer behaviour and demand response of tourists to climate change. Annals of Tourism Research. 2012;39(1):36–58. http://dx.doi.org/10.1016/j.annals.2011.11.002
7. Patterson T, Bastianoni S, Simpson M. Tourism and climate change: Two-way street, or vicious/virtuous circle? J Sustain Tour. 2006;14(4):339–348. http://dx.doi.org/10.2167/jost605.0
8. Moreno A. Climate change and tourism: Impacts and vulnerability in Coastal Europe [thesis]. Maastricht: Maastricht University; 2006.
9. Zaninović K, Matzarakis A. The bioclimatological leaflet as a means conveying climatological information to tourists and the tourism industry. Int J Biometeorol. 2009;53(4):369–374. http://dx.doi.org/10.1007/s00484-009-0219-2
10. Berritella M, Bigano A, Roson R, Tol RSJ. A general equilibrium analysis of climate change impacts on tourism. Working Papers 127. Hamburg: Centre for Marine and Atmospheric Science; 2004.
11. Williams PW, Ponsford IF. Confronting tourism's environmental paradox: Transitioning for sustainable tourism. Futures. 2009;41:396–404. http://dx.doi.org/10.1016/j.futures.2008.11.019
12. Hall CM, Amelung B, Cohen S, Eijgelaar E, Gössling S, Higham J, et al. On climate change skepticism and denial in tourism. J Sustain Tour. 2015;23(1):4–25. http://dx.doi.org/10.1080/09669582.2014.953544
13. Turton S, Hadwen W, Wilson R. The impacts of climate change on Australian tourism destinations: Developing adaptation and response strategies – a scoping study. Gold Coast: CRC for Sustainable Tourism; 2009.
14. Simpson MC, Gössling S, Scott D, Hall CM, Gladin E. Climate change adaptation and mitigation in the tourism sector: Frameworks, tools and practices. Paris: UNEP; 2008.
15. Kaján E, Saarinen J. Tourism, climate change and adaptation: A review. Curr Issues Tourism. 2013;16(2):167–195. http://dx.doi.org/10.1080/13683500.2013.774323
16. Rogerson CM. Reviewing Africa in the global tourism economy. Dev South Afr. 2007;24(3):361–379. http://dx.doi.org/10.1080/03768350701445350
17. Rogerson CM. The tourism-development nexus in sub-Saharan Africa – progress and prospects. African Insight. 2012;42(2):28–45.
18. Preston-Whyte RA, Watson HK. Nature tourism and climate change. Clevedon: Channel View Publications; 2005. p. 130–142.
19. Van Staden JW, Marais L. The tourism potential of Beaufort West: A study based on visitor demand. Dev South Afr. 2005;22(2):233–249. http://dx.doi.org/10.1080/03768350500163147
20. Theron AK, Rossouw M. Analysis of potential coastal zone climate change impacts and possible response options in the southern African region. Pretoria: CSIR; 2008.
21. Davis C. Climate change handbook for north-eastern South Africa. Pretoria: CSIR; 2010.
22. Saarinen J, Tervo K. Perceptions and adaptation strategies of the tourism industry to climate change: The case of Finnish nature-based tourism entrepreneurs. Int J Innovation and Sustainable Development. 2006;2(2):214–228. http://dx.doi.org/10.1504/IJISD.2006.012423
23. Saarinen J, Hambira WL, Atlhopheng J, Manwa H. Tourism industry reaction to climate change in Kgalagadi South District, Botswana. Dev South Afr. 2012;29(2):273–285. http://dx.doi.org/10.1080/0376835X.2012.675697
24. Kwadijk JCJ, Haasnoot M, Mulder JPM, Hoogvliet MMC, Jeuken ABM, Van Der Krogt RAA, et al. Using adaptation tipping points to prepare for climate change and sea-level rise: A case study from the Netherlands. WIREs Climate Change. 2010;1(5):729–740. http://dx.doi.org/10.1002/wcc.64
25. Nicholls RJ, Cazenave A. Sea-level rise and its impact on coastal zones. Science. 2010;382:1517–1520. http://dx.doi.org/10.1126/science.1185782
26. Illenberger WK. Nuclear power station ('nuclear 1') and associated infrastructure: Addendum to dune geomorphology impact assessment: Debris flows in the Sand River and potential for flood damage to the R330. Port Elizabeth: Arcus GIBB; 2010.
27. James N, Hermes J. Insights into impacts of climate change on the South African marine and coastal environment. Pretoria: SAEON; 2011.
28. Eliassen A. Saint Francis Bay, Eastern Cape (South Africa) [article on the Internet]. c2013 [cited 2014 March 7]. Available from: http://www.yr.no/place/South_Africa/Eastern_Cape/Saint_Francis_Bay/
29. Frith A. Census 2011 [article on the Internet]. c2014 [cited 2014 February 14]. Available from: http://census2011.adrianfrith.com/place/268011
30. Mortimer B. Little Venice in the Cape [article on the Internet]. c2014 [cited 2014 April 12]. Available from http://www.safricavoyage.com/stfrancis.htm
31. Elsden M, Rossoux S. Exploring South Africa: A coastal odyssey. Vol. 1. Pretoria: Modus Associates Limited; 2009.
32. Cowling RM. The occurrence of C3 and C4 grasses in fynbos and allied shrublands in the South Eastern Cape, South Africa. Oecologia. 1983;58:121–127. http://dx.doi.org/10.1007/BF00384551
33. Lubke RA. Erosion of the beach at St Francis Bay, Eastern Cape, South Africa. Biol Conserv. 1985;31:99–127. http://dx.doi.org/10.1016/0006-3207(85)90080-1
34. Buzter KW, Helgren DM. Late Cenozoic evolution of the Cape coast between Knysna and Cape St Francis, South Africa. Quaternary Res. 1972;2:143–169. http://dx.doi.org/10.1016/0033-5894(72)90035-X
35. Schumann EH, Martin JA. Climatological aspects of the coastal wind field at Cape Town, Port Elizabeth and Durban. S Afr Geogr J. 1991;73(2):48–51. http://dx.doi.org/10.1080/03736245.1991.9713548
36. J-Space Systems. ASTER GDEM [article on the Internet]. c2016 [cited 2016 May 25]. [First accessed 30 Oct 2014 from: http://gdem.ersdac.jspacesystems.or.jp/index.jsp] Available from: http://jspacesystems.or.jp/library/archives/ersdac/GDEM/E/index.html
37. National Aeronautics and Space Administration (NASA). ASTER [article on the Internet]. c2014 [cited 2014 September 21]. Available from: http://asterweb.jpl.nasa.gov/

38. Mcleod E, Poulter B, Hinkel J, Reyes E, Salma R. Sea-level rise impact models and environmental conservation: A review of models and their applications. Ocean Coast Manage. 2010;53(9):507–517. http://dx.doi.org/10.1016/j.ocecoaman.2010.06.009

39. Brunel C, Sabatier F. Potential influence of sea-level rise in controlling shoreline position on the French Mediterranean Coast. Geomorphology. 2009;107:47–57. http://dx.doi.org/10.1016/j.geomorph.2007.05.024

40. Schmidt K, Hadley B, Waters K. Mapping and portraying inundation uncertainty of bathtub-type models. J Coastal Res. 2013;30(3):548–561.

41. Levine NS, Kaufman CC, Sutherland M, Renaud L. Adapting HAZUS-MH to assess sea-level rise. Geological Society of America. 2008;40(6):84.

42. Stainforth DA, Allen MR, Tredger ER. Smith LA. Confidence, uncertainty and decision-support relevance in climate predictions. Phil Trans R Soc A. 2007;365(1857):2145–2161.

43. Breetzke T, Parak O, Cilliers L, Mather AA, Colenbrander D. Living with coastal erosion: A short term best practice guide. In: Proceedings of the International Association of Impact Assessors South Africa conference. Bloemfontein: University of the Free State; 2008.

44. Mather AA, Garland GG, Stretch DD. Southern African sea-levels: Corrections, influences and trends. Afr J Mar Sci. 2009;31(2):145–156. http://dx.doi.org/10.2989/AJMS.2009.31.2.3.875

45. Rouault M, Pohl B, Penven P. Coastal oceanic climate change and variability from 1982 to 2009 around South Africa. Afr J Mar Sci. 2010;32(2):237–246. http://dx.doi.org/10.2989/1814232X.2010.501563

46. Goschen WS, Schumann EH, Bernard KS, Bailey SE, Deyzel SHP. Upwelling and ocean structures off Algoa Bay and the south-east coast of South Africa. Afr J Mar Sci. 2012;34(4):525–536. http://dx.doi.org/10.2989/1814232X.2012.749810

47. Mieczkowski Z. The tourism climatic index: A method of evaluating world climates for tourism. Can Geogr. 1985;29(3):220–233. http://dx.doi.org/10.1111/j.1541-0064.1985.tb00365.x

48. Perch-Nielsen S, Amelung B, Knutti R. Future climate resources for tourism in Europe based on the daily tourism climatic index. Climatic Change. 2010;103(3–4):363–381. http://dx.doi.org/10.1007/s10584-009-9772-2

49. Fitchett JM, Grab SW. A 66-year tropical cyclone record for south-east Africa: Temporal trends in a global context. Int J Climatol. 2014;34(13):3604–3615. http://dx.doi.org/10.1002/joc.3932

50. Gössling S, Bredberg M, Randow A, Sandström E, Svensson P. Tourist perceptions of climate change: A study of international tourists in Zanzibar. Curr Issues Tourism. 2006;9(4–5):419–435. http://dx.doi.org/10.2167/cit265.0

51. Tandross M, Davis C, Engelbrecht F, Joubert A, Archer van Garderen E. Regional scenarios of future climate change over southern Africa. Pretoria: CSIR; 2011. http://dx.doi.org/10.1175/WCAS-D-11-00026.1

52. Thorpe H. Submission on Sand River bridge report. St. Francis Chronicle. 2014; p.10.

53. Roshan G, Yousefi R, Fitchett JM. Long-term trends in tourism climate index scores for 40 stations across Iran: The role of climate change and influence on tourism sustainability. Int J Biomet. 2015; http://dx.doi.org/10.1007/s00484-015-1003-0

54. Walker ND, Mey RD. Ocean/atmosphere heat fluxes within the Agulhas retroflection region. J Geophys Res. 1988;93(12):15473–15483. http://dx.doi.org/10.1029/JC093iC12p15473

55. Baird D, Heymans JJ. Assessment of ecosystem changes in response to freshwater inflow of the Kromme River Estuary, St. Francis Bay, South Africa: A network analysis approach. Water SA. 1996;22:307–318.

56. Carter A, McMaster A, Morris G, Whittington-Jones K, Constable L, Parramon M. Eastern Cape climate change response strategy. Eastern Cape: Department of Economic Development and Environmental Affairs; 2011.

57. Cartwright A. Coastal vulnerability in the context of climate change: A South African perspective. Cape Town: Centre for Criminology; 2011.

58. Mukheibir P, Ziervogel G. Developing a municipal adaptation plan (MAP) for climate change: The city of Cape Town. Environ Urban. 2007;19(1):143–158. http://dx.doi.org/10.1177/0956247807076912

59. Faling W, Tempelhoff JW, Van Niekerk D. Rhetoric or action: Are South African municipalities planning for climate change? Dev South Afr. 2006;29(2):241–257. http://dx.doi.org/10.1080/0376835X.2012.675695

60. Morton TA, Rabinovich A, Marshall D, Bretschneider P. The future that may (or may not) come: How framing changes responses to uncertainty in climate change communications. Global Environ Chang. 2011;21(1):103–109. http://dx.doi.org/10.1016/j.gloenvcha.2010.09.013

61. Miceli R, Sotgiu I, Settanni M. Disaster preparedness and perception of flood risk: A study in an alpine valley in Italy. J Environ Psychol. 2008;28(2):164–173. http://dx.doi.org/10.1016/j.jenvp.2007.10.006

Two-dimensional mapping of scientific production of nations in the fields of physics and astronomy

AUTHOR:
Seyyed Mehdi Hosseini Jenab[1]

AFFILIATION:
[1]Centre for Space Research, Faculty of Natural Sciences, North-West University, Potchefstroom, South Africa

CORRESPONDENCE TO:
Mehdi Jenab

EMAIL:
mehdi.jenab@yahoo.com

POSTAL ADDRESS:
Centre for Space Research, Faculty of Natural Sciences, North-West University, Hoffman Street, Potchefstroom 2520, South Africa

DATES:
Received: 22 Mar. 2015
Revised: 20 July 2015
Accepted: 20 Jan. 2016

KEYWORDS:
national-level analysis; quality; quantity; ranking; scientometrics

HOW TO CITE:
Hosseini Jenab SM. Two-dimensional mapping of scientific production of nations in the fields of physics and astronomy. S Afr J Sci. 2016;112(5/6), Art. #2015-0119, 8 pages. http://dx.doi.org/10.17159/sajs.2016/20150119

The quantity and quality of scientific production in the fields of physics and astronomy over a period of 16 years (1996–2012) was studied. The level of analysis was national, with the scientific output of 108 countries being analysed. The measurement unit was the number of papers published in peer-reviewed journals, as listed on the Scopus database. Modified versions of the number of publications (Pm) and citations (Cm) were employed as indicators of quantity and quality, respectively. A two-dimensional method, the Pm-Cm diagram, was adopted to provide a coherent and simultaneous approach to study the positions, rankings, and temporal evolution of countries in the global context. A static approach to studying the Pm-Cm diagram resulted in countries being grouped into five main categories based on average positions. A dynamic approach to analysing the Pm-Cm diagram also resulted in five groups (i.e. when considering the temporal evolution patterns of the countries during the studied years). The rank and temporal-evolution group associated with each country are listed in two tables in this paper. These tables, together with the Pm-Cm diagrams (showing different scales) present a general view of the scientific activity in the field of physics and astronomy for each country. This methodology allows each country's output to be compared with that of other countries or the world average.

Introduction

Scientometrics is the systematic measurement of science, or more precisely, a method of objectively measuring scientific activity. This approach was pioneered by the American psychologist James McKeen Cattell in 1906, when the first edition of the *American Men of Science* was published. Cattell stated that 'There is here given for the first time a fairly complete survey of the scientific activity of a country at a given period'.[1] In scientometrics, contemporary researchers in North America are ranked and their output is statistically analysed based on productivity and merit.[2,3]

After the Second World War, two main organisations began to analyse and measure the scientific activity of the industrial world. These were the US National Science Foundation (NSF) in the 1950s, and intergovernmental organisations such as the Organization for Economic Co-operation and Development (OECD) in the 1960s. Their analyses were important for countries to show their development.[4] This concern was expressed in the OECD report of 1963 as follows:

> *… the strength, progress and prestige of countries are today measured in part by their achievements in S&T [science and technology], scientific excellence is more and more becoming an important national goal. National resources are therefore increasingly devoted in research and development.*[5]

In 1960, Eugene Garfield started a systematic archive of scientific activity. He founded the well-known Institute for Scientific Information (ISI), which specialised in indexing scientific papers and their citations in peer-reviewed journals. Since then, enormous effort has been dedicated to developing methods for ranking and analysing scientific activity at various levels, namely the individual, group, university, and national levels. Much of this effort, especially recently – owing to easier access to bibliometric data – has been focused on introducing indicators for measuring various aspects of scientific activity. Such measures include excellence, impact, quality, quantity, effectiveness of the process, and even economical value ('return on investment' or ROI). Rostaing states that 'The goal in our days is to create indicators that assess research activity, measure research actors' productivity and their strategic position'.[6]

The definitions and measuring techniques for assessing various aspects of scientific activity remain a challenge. Nonetheless, the most important and common aspects of scientific activity, which have been widely studied, include the quality and quantity of scientific production. The most typical and tangible measure of scientific production refers to outputs such as papers, patents, or offshoot technologies. These measures are especially pertinent when considering science and technology (S&T) activity at the national level.[7]

The number of publications, as an index of quantity, was one the early indicators introduced in scientometrics. Modified adaptations of this measure have been presented for the different levels.[8] However, considering the number of publications individually means assessing only one aspect of scientific activity, which results in a relatively simplistic understanding. Hence a few techniques have been introduced and developed to present a broader view of scientific production, by combining or adding other aspects – especially quality – to the indicator. Vinkler has discussed some of these modifications as 'composite indicators'.[9]

A well-known index of this kind on the individual level was introduced by J.E. Hirsch.[10] Hirsch combined both quantity and quality indicators, namely the number of papers and number of citations respectively, for a single researcher. The resulting score was termed the 'H-index'. The two main aspects are often shown in a two-dimensional diagram called the H-index diagram. Although the H-index is relatively new, it has been so successful that nowadays it is used officially in assessing individual researchers in some countries. Databases such as Scopus and ISI provide it as the main index for authors of scientific papers. Since its introduction, a prolific sequence of customised versions

of the H-index have also been introduced, such as the G-index[11] and the 'H-index for journals'[12]. Variations of the H-index have been compared by Bornmann et al.[13] In defiance of its initial definition, the H-index has been employed in studying and ranking other levels of scientific activity, such as countries.[7]

By its nature, the H-index is a static indicator that ignores the temporal evolution of scientific activity, and the H-index associated with a researcher or country cannot be reduced. In other words, an H-index value will stay the same even if the researcher's career or the country's scientific activity declines or stops altogether.[14]

The time evolution of scientific activity should be considered. As the work of Radicchi[15] on individual physicists showed, rankings that take account of time evolution differ from those that are obtained from the H-index alone.

Two-dimensional diagrams have been employed in various fields, for example the well-known Carnot cycle in thermodynamics, as proposed by Sadi Carnot in 1824. This kind of diagram provides an inclusive understanding of two aspects of a specific topic, while preserving the chronological order of events. In the case of the Carnot cycle, volume and pressure are presented in a PV diagram.

We studied the further development of H-index diagrams to represent the national level, based on a two-dimensional method. The resulting diagrams showed the coherent and simultaneous temporal evolution of two aspects of the work of a country, namely the quantity and quality of its science.[16] We chose to consider scientific production at the national level in fundamental areas of science – physics and astronomy – because of the role these fields play as a driving force of scientific activity. In the modern era of big science, physics and astronomy receive large amounts of funding and hence acquire large numbers of researchers. This range of activity renders these fields a good example for our study because statistical data are available for many countries. Furthermore, physics and astronomy, as two main fields in science, may reflect the overall scientific activity of the studied countries in general. However, the results of our study should be cautiously interpreted and are not necessarily generalizable to other scientific activity.

The quality of scientific activity is regarded as a complex and to some extent ambiguous concept to define. The most common measure is that of the number of citations. This measure is widely accepted as an indicator of the quality of scientific production, that is, the 'quality' criterion for scientific activity.[17-19] In the period we studied, 1996–2012, papers had been published in the fields of physics and astronomy in peer-reviewed journals by 108 countries worldwide. The peer-reviewed journals were indexed in Scopus. Hence, although only 108 counties are considered here, the data represent the entire world's scientific activity in physics and astronomy during the studied period.

According to Macilwain[20], 'Science has been invested even during the last economic recession as part of stimulus packages designed to aid troubled economies around the globe'. Understanding patterns of quality and quantity in scientific activity at the national level plays an important role in countries' investment in science. The struggle to maintain a balance between these two aspects has been a great issue for policy-makers and scientific advisors around the world. Most countries, especially in the early stage of scientific development, focus their resources on increasing the quantity of scientific activity, which leads to an unbalanced pattern of quantity versus quality. These countries emphasise an increase in quality indicators, which is coined 'poor Scientometrics' by Van Raan[21]. A two-dimensional approach contributes a more comprehensive perspective on countries' patterns of scientific activity in terms of the balance between quality and quantity. This can also help in understanding their relative positions compared to each other and to the world as a whole.

Method

We selected as the 'quantity' indicator the number of publications in peer-reviewed journals, represented here as $P(y_i, c_j)$. In this term, y_i and c_j stand for year and country respectively. Index i for y_i (year) runs from 1 to 16, presuming $y_1 = 1996$ and $y_{16} = 2012$. Index j in c_j (country) varies from 1 to 108, as 108 countries featured in our study. We ordered them alphabetically, hence $c_1 = Albania$ and $c_{108} = Yemen$. For example, in 2010, China published $P(2010,China) = P(y_{14},c_{18}) = 27700$ papers.

The number of citations is considered the 'quality' indicator, and is represented by $C(y_i\text{-}y_{16},c_i)$. Note that a citation window spans from the year of publication until the final year of our study period, that is $y_{16} = 2012$. For instance, publications by China in 2010 received $C(2010\text{-}2012,China) = C(y_{14}\text{-}y_{16},c_{18}) = 116143$ citations.

The global growth in the number of publications during the studied period makes a comparison of publication numbers misleading, especially considering the ranking of countries. However, we can also compare the number of publications of a country to the world average by introducing a modified formula, as follows:

$$Pm(y_i,c_j) = \frac{P(y_i,c_j)}{\sum_{j=1}^{108} P(y_i,c_j)/108} \qquad \text{Equation 1}$$

The numbers belonging to different years will have the same meaning, and can have the same value as they refer to the world average. For instance, comparing the number of publications of Japan in 1996, that is $P(y_1,c_{47}) = 10776$, with the number of publications of France in 2010, that is $P(y_{14},c_{32}) = 10793$, one might conclude that both these countries attained the same level for quantity in scientific activity. But the global intensity of scientific activity in the fields of physics and astronomy differed considerably between 1996 and 2010. In other words, these two numbers belong to different contexts. Scaled to the world average of the respective years, France in 2010 produced $Pm(y_{14},c_{32}) = 5.34$ times the world average, whereas Japan in 1996 published $Pm(y_1,c_{47}) = 10.51$ times the world average. Consequently, using the Pm measure, the quantity of scientific activity of Japan in 1996 was twice that of France in 2010 (see Table 1 below).

Table 1: Number of publications by Japan in 1996 and France in 2010

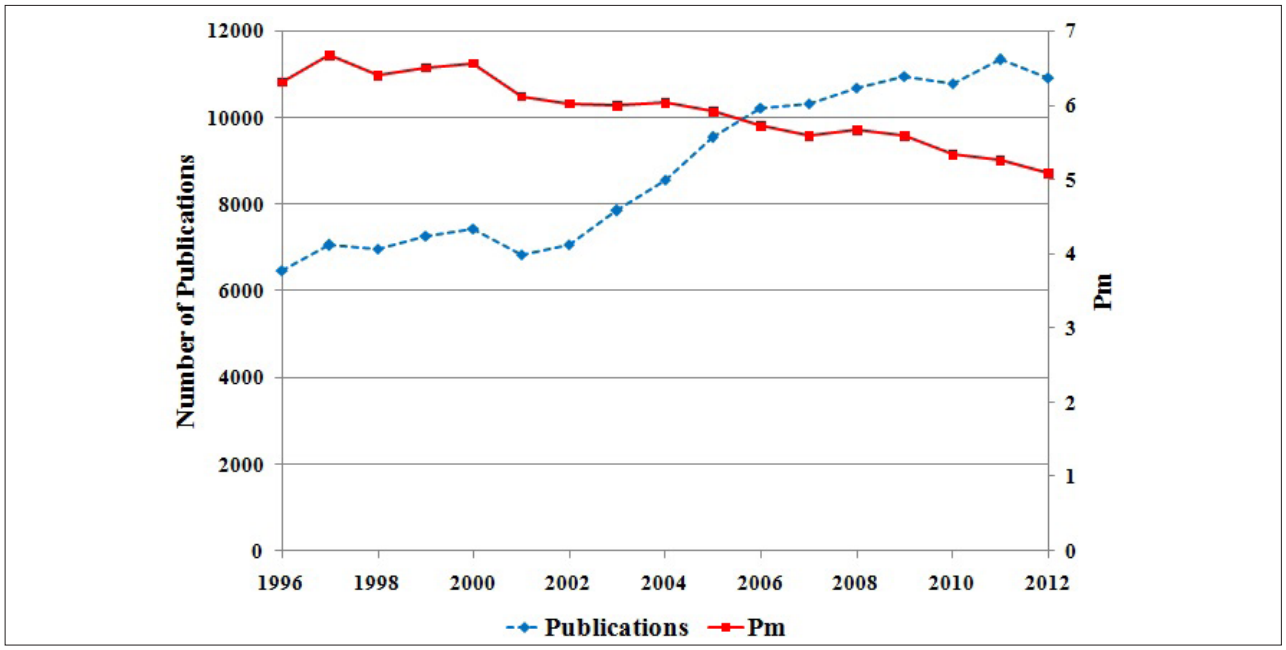
Country	Year	Publication (P)	Pm
Japan	1996	10 776	10.51
France	2010	10 793	5.34

Temporal evolution should be placed in the context of the world average, even if only one country is analysed. For instance, France increased its number of annual publications during the study period, almost doubling its output from 6475 to 10924. However, when these numbers are compared with the world average, the position of France has actually declined, with its Pm dropping from $Pm(y_1,c_{32}) = 6.32$ to $Pm(y_{16},c_{32}) = 5.08$ (see Figure 1).

The same logic justifies the application of a modified number of citations, as defined below.

$$Cm(y_i - y_{16},c_j) = \frac{C(y_i - y_{16},c_j)}{\sum_{j=1}^{108} P(y_i - y_{16},c_j)/108} \qquad \text{Equation 2}$$

There is a gap between the year of publication of a paper and the point at which it has been cited enough times to achieve statistical saturation for citations. This lag period lasts roughly 3 to 5 years.[22] There is also a downward trend, over time, for the number of citations of a published study. These two patterns distort the temporal evolution lines of countries. The number that results from a comparison with the world average (Cm) can remove these effects, which is another reason to use Cm instead of the number of citations.[22]



Blue dashed line = number of publications; red solid line = Pm value

Figure 1: Temporal evolution of number of publications and Pm value for France compared with world average.

To produce the diagram, each country's pair number (Pm,Cm) for every year is shown by a dot on the diagram. The Pm and Cm values are represented on the vertical and horizontal axes respectively. By connecting these dots based on chronological order, the temporal evolution line of a country can be drawn. The data needed for this study were obtained from the SCImago Journal & Country Rank portal,[23] which provides Scopus data arranged according to country, branch of science, and year.

Results and discussion

Logarithmic scale diagram

Figure 2 shows a two-dimensional diagram of all countries in logarithmic scale, to provide a holistic and integrated view of the world's scientific activity pattern in the fields of physics and astronomy. A striking difference among countries with regard to scientific activity is evident. The figure shows that countries do not cluster around the average value of Pm=Cm=1, but rather a few countries cluster around the average points and most countries fall in the low-rank area of (Pm,Cm)<0.5 (see Table 2).

Activity

Three main areas can be distinguished in the Pm-Cm diagram:

- lower than average: Pm,Cm<0.5
- average: 0.5<Pm,Cm<1.5
- higher than average: 1.5<Pm,Cm

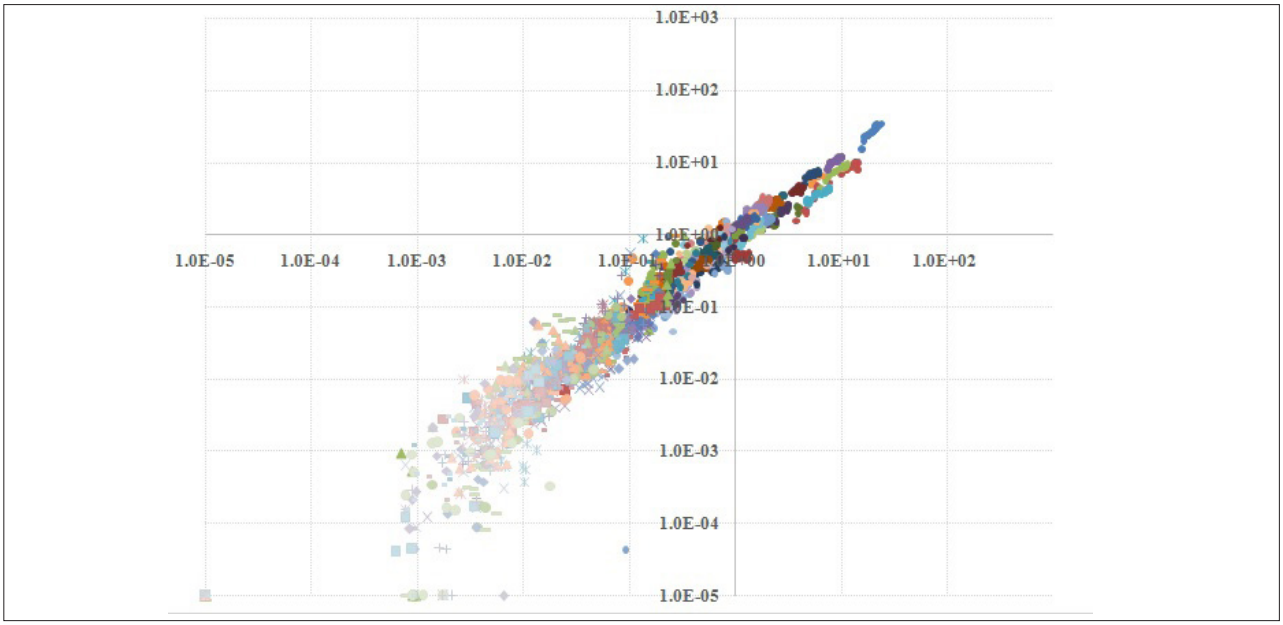


Figure 2: Logarithmic scale of the diagram (Pm,Cm) for all countries in the study period (1996–2012).

Table 2: Ranks of countries based on their positions in Pm-Cm diagram

Rank	Definition	Countries
Exceptional	Pm,Cm>15	USA, China
Top	4<Pm,Cm<15	Germany, Japan, UK, France, Russia, Italy, India
High	1.5<Pm,Cm<4	Spain, South Korea, Canada, Switzerland, Poland, Netherlands, Australia, Taiwan, Brazil, Sweden
Average	0.5<Pm,Cm<1.5	Austria, Belgium, Iran, Israel, Czech Republic, Ukraine, Denmark, Turkey, Portugal, Singapore, Finland, Romania, Malaysia, Hong Kong
Low	Pm,Cm<0.5	Greece, Mexico, Hungary, Argentina, Ireland, South Africa, Norway, Chile, Serbia, Colombia, Slovenia, Armenia, Pakistan, Slovakia, Belarus, Egypt, Bulgaria, Saudi Arabia, Georgia, Croatia, New Zealand, Lithuania, Morocco, Cyprus, Estonia, Azerbaijan, Algeria, Thailand, Tunisia, Vietnam, Jordan, Indonesia, Puerto Rico, Cuba, Latvia, Uzbekistan, Ecuador, Moldova, Kazakhstan, Venezuela, Iceland, Bangladesh, United Arab Emirates, Peru, Nigeria, Luxembourg, Iraq, Uruguay, Lebanon, Qatar, Oman, Cameroon, Palestine, Philippines, Macedonia, Syrian Arab Republic, Kuwait, Sri Lanka, Bahrain, Mongolia, Sudan, Bosnia and Herzegovina, Senegal, Montenegro, Kenya, Costa Rica, Ethiopia, Yemen, Tajikistan, Albania, Ghana, State of Libya, Kyrgyzstan, Côte d'Ivoire, North Korea

In the last area, three subgroups can be detected based on clusters they have formed, by considering their average position. The definition and borders of these above-average subgroups are to some extent arbitrary, because by changing the time-span of the study the clusters might be altered. Consequently, countries are divided into five ranks of activity level. Table 2 below presents the results of this categorising, with the names of countries written in order based on their average positions. Countries at the bottom or top of a ranking list should be seen as falling between two ranks (i.e. borderline). In the case of countries that have shown a pattern of either great decline or great progress, the last three years of their temporal evolution was considered when ranking them.

In the following section, each category is discussed and analysed, and diagrams (drawn to appropriate scale) are presented in Figures 3 to 7.

Exceptional activity: USA and China

For years, USA has held the position of being by far the most active nation in physics and astronomy. However, the picture is changing, and because of the recent growth in scientific activity across the world, the output of USA now shows a rapid decline. By contrast, China is showing rapid progress and a high level of activity, and seems likely to catch up with USA in the next five years. The USA has been warned about this by senior scientists from US National Academies.[24] These two countries currently stand out from other nations because of their exceptional level of activity – almost 15 times more than the global average (see Figure 3).

Top active countries

Figure 4 shows the rankings of the seven most active countries, defined as 4<(Pm,Cm)<15. All except India show a decline pattern as a result of the global growth in scientific activity. China is shown in the figure for the sake of comparison. The second fastest decline after USA is noted for Japan, which clearly has lost its position in comparison with China and even Germany.

Countries with high activity

The third category consisted of ten countries labelled as having a 'high' activity level. These countries are shown in Figure 5, with India (a 'top' country) included for the sake of comparison. Four countries, mostly new-emerging economies, show a growth pattern – that is, Korea, Spain, Taiwan, and Brazil.

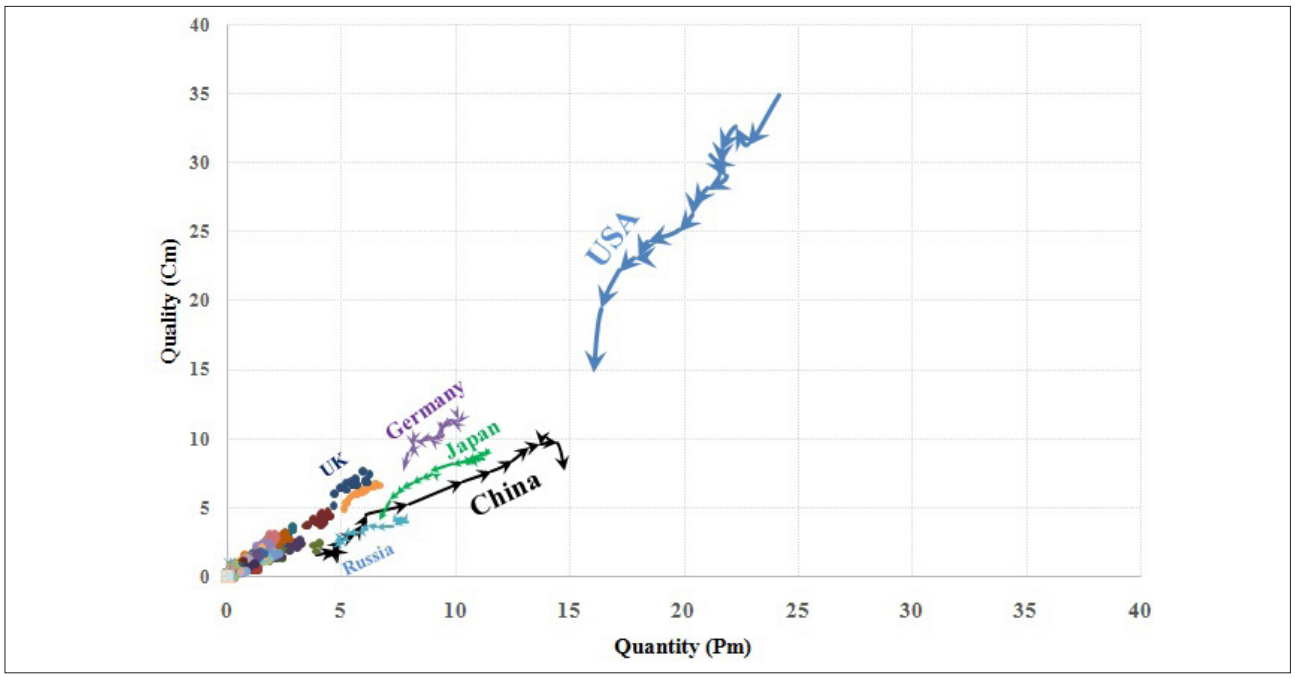


Figure 3: Two-dimensional diagram for most active countries, including USA.

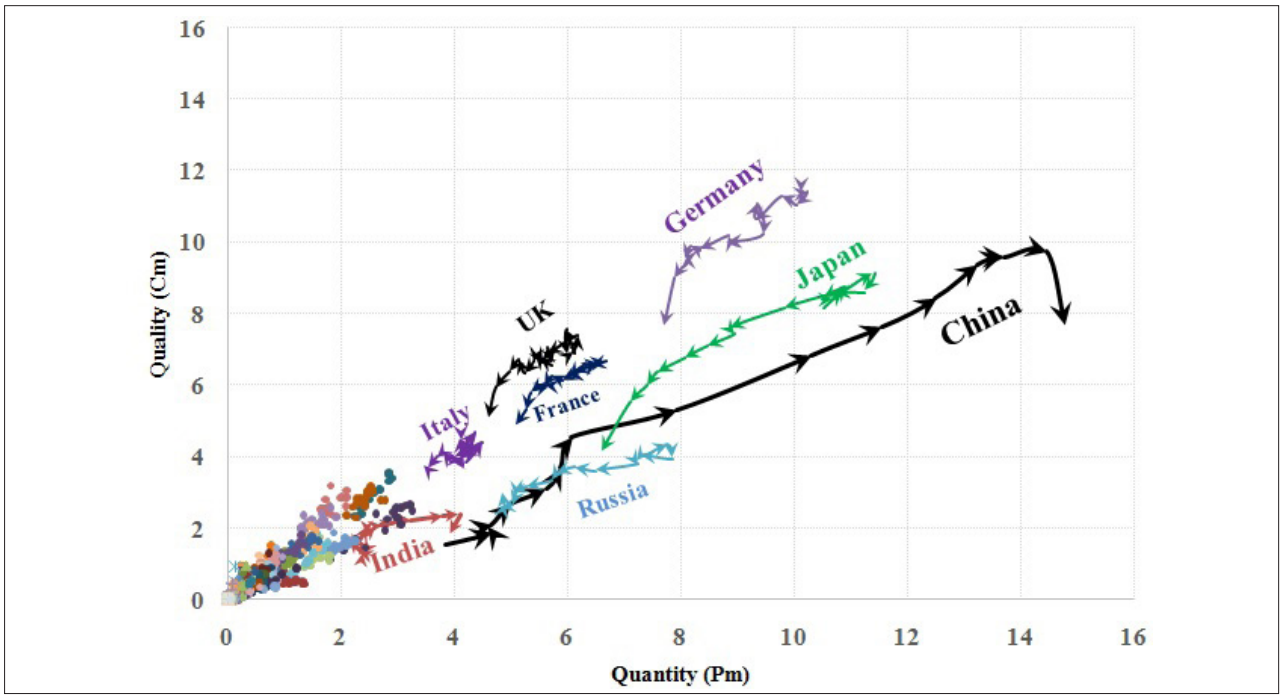


Figure 4: Two-dimensional diagram for most active countries, excluding USA.

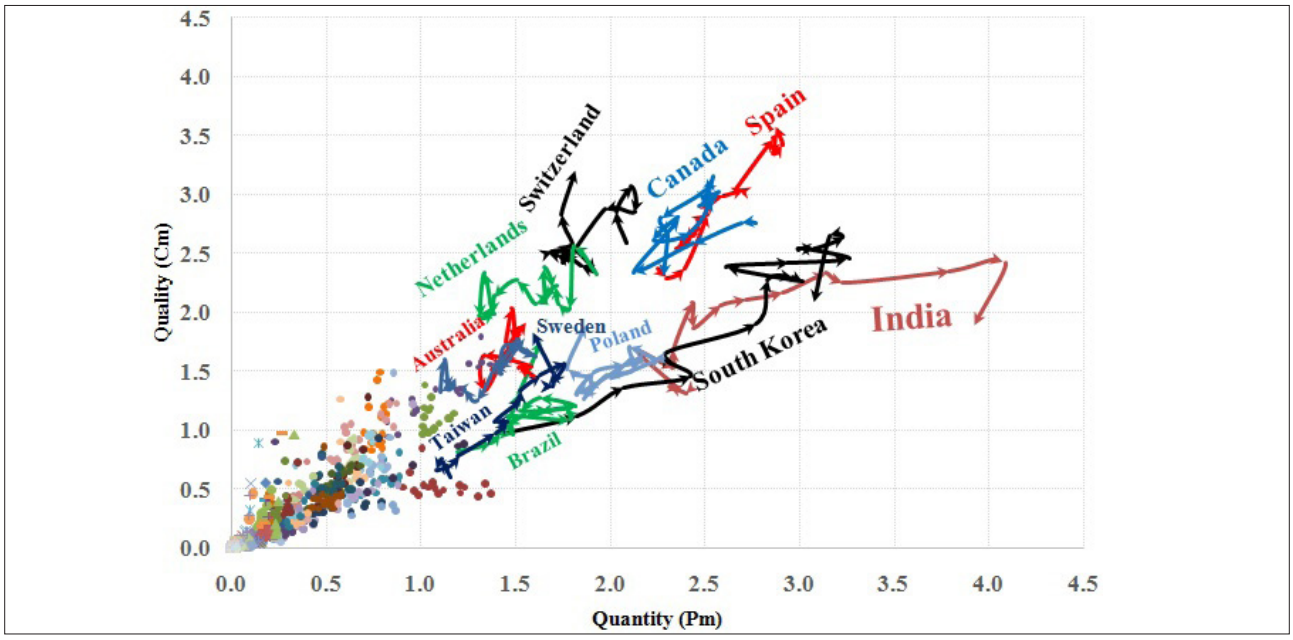


Figure 5: Pm-Cm diagram for ten highly active countries.

Countries with average activity

Twelve counties with 'average' activity are shown in Figure 6. Two separate figures are presented to minimise the complexity and confusion that arises from temporal evolution lines crossing each other. The rapid growth of Iran, Turkey, and Malaysia are clearly visible, as well as a steep decline for Israel.

Countries with low activity

Most of the countries in this study, namely 74 countries (representing 68% of the included countries) were categorised as having a 'low' level of activity. This means their activity fell below half of the world average for the studied period. However, certain countries in this group have shown great improvement and fall on the border between 'low' and 'average' rankings. South Africa and Greece, for example, are catching up with the 'average' countries in physics and astronomy (see Figure 7 and Table 2).

Temporal evolution

Categories of temporal evolution

With regard to temporal evolution, countries usually show one of three main patterns: growth, decline, or stasis (stationary). Countries showing a decline or growth pattern are divided into two subgroups each, based on the amplitude of their movement in the Pm-Cm diagram, which is basically defined as the distance between a country's averaged initial and final positions. Hence in temporal terms, countries are divided into five groups as shown in Table 3. Many countries are labelled as stationary because they have negligible movement (i.e. the distance between their averaged initial and final positions is less than 0.5).

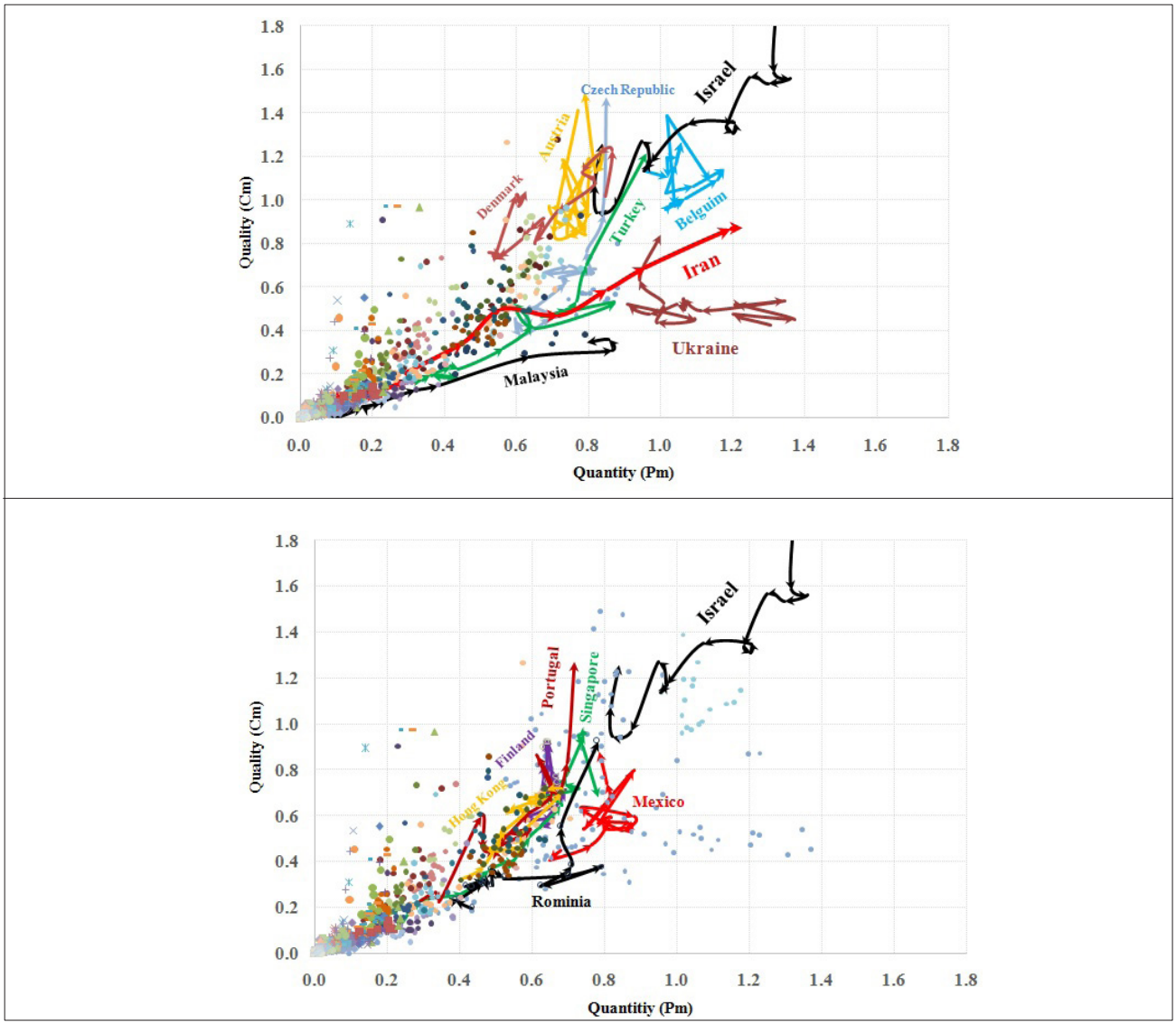


Figure 6: Countries with 'average' rank (shown in two pictures with the same scale).

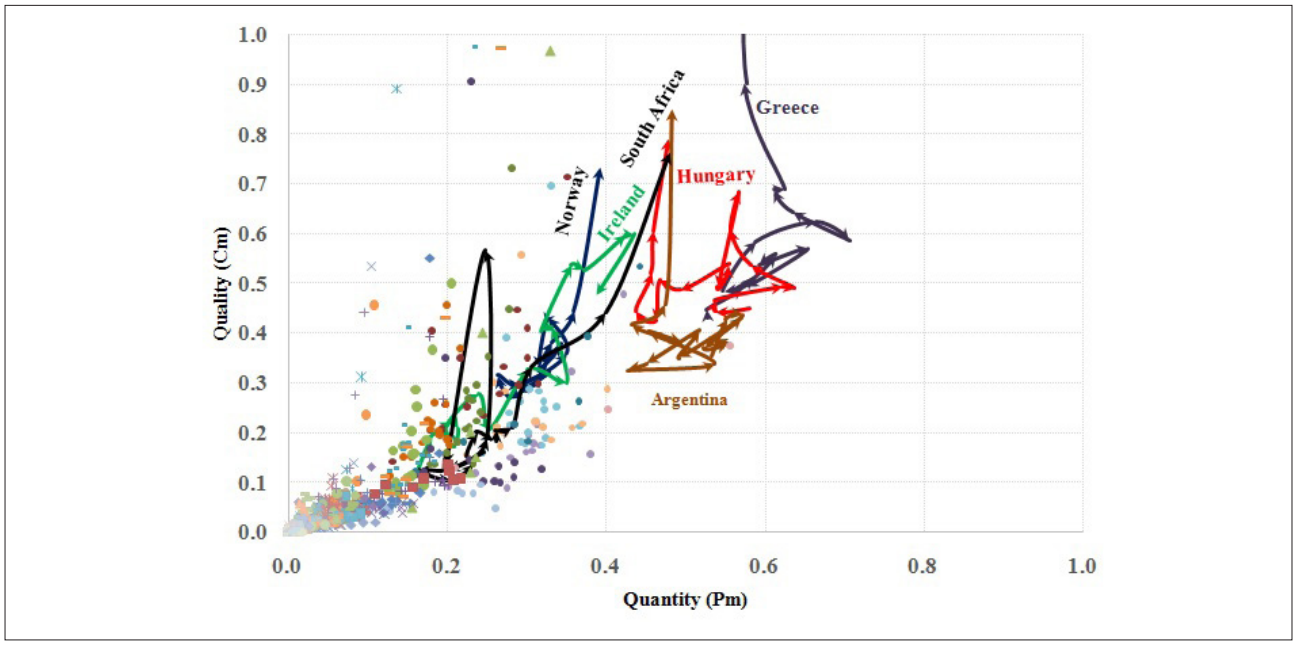


Figure 7: Selected countries of 'low' ranking (note the rapid growth for Greece and South Africa).

Table 3: Groups of countries based on temporal evolution in Pm-Cm diagram

Temporal evolution group	Countries
Rapid growth	China 12.51, South Korea 1.98, India 1.60, Iran 1.43, Spain 1.37, Taiwan 1.31, Turkey 1.23, Portugal 1.13, Czech Republic 1.10, Brazil 1.04, Serbia 1.02
Growth	Colombia 0.94, Armenia 0.92, Georgia 0.85, Greece 0.84, Singapore 0.82, Romania 0.81, Malaysia 0.78, South Africa 0.65, Pakistan 0.63, Chile 0.61, Saudi Arabia 0.59, Slovenia 0.59, Azerbaijan 0.52
Stationary	Mexico, Cyprus, Egypt, Ireland, Croatia, Norway, Estonia, Lithuania, New Zealand, Thailand, Algeria, Hong Kong, Vietnam, Tunisia, Jordan, Austria, Indonesia, Ecuador, United Arab Emirates, Iceland, Iraq, Luxembourg, Peru, Qatar, Lebanon, Palestine, Cuba, Cameroon, Nigeria, Sudan, Sri Lanka, Montenegro, Syrian Arab Republic, Kenya, Senegal, North Korea, Yemen, Ghana, State of Libya, Côte d'Ivoire, Tajikistan, Bahrain, Bangladesh, Kazakhstan, Oman, Bosnia and Herzegovina, Ethiopia, Philippines, Argentina, Australia, Slovakia, Poland, Bulgaria, Finland, Denmark, Belgium, Moldova, Kyrgyzstan, Sweden, Macedonia, Latvia, Venezuela, Uzbekistan, Puerto Rico, Uruguay, Mongolia, Albania, Kuwait, Costa Rica, Morocco
Decline	Ukraine 0.5, Switzerland 0.67, Canada 0.67, Italy 0.68, Netherlands 0.70, Israel 0.72, Belarus 0.80, Hungary 0.80
Rapid decline	France 2.07, UK 2.36, Russia 2.78, Germany 4.64, Japan 5.59, USA 21.55

Note: The number after a country represents the amplitude of movement.

China possesses the largest amplitude, and is by far the fastest-growing country in the fields of physics and astronomy. But although China shows the most *rapid* growth in the studied period (see Table 3), the Pm-Cm diagram reveals a deviation toward *quantity* when China is compared with other countries in the 'top' category. The same bias is evident when China's pattern is compared with the line that represents the world's average temporal evolution line – which runs at a 45° angle (see Figure 2). The other rapidly-growing countries are represented by amplitude in Table 3, with their temporal evolution shown in Figures 3 to 6.

The USA and most other developed countries fall into the groups labeled 'decline' or 'rapid decline'. The USA possesses the largest amplitude, and stands out as by far the fastest-declining country in the fields of physics and astronomy. However, considering Figure 1, it remains ahead of all other nations.

Changing rankings

A few countries having growth patterns have managed to change their ranks in the diagram. This can be clearly witnessed by analysing their positions in the early years versus recent years of the studied period. The countries that have managed to improve their rankings are as follow:

1. From high to exceptional: China
2. From high to top: India
3. From average to high: Brazil, Taiwan
4. From low to average: Turkey, Iran, and Malaysia

No country dropped in ranking during the studied years, although a few countries moved from the upper area of a category to the lower area. In the coming years, such countries might drop down to a lower category.

Conclusion

Pm-Cm diagrams were produced by employing a two-dimensional method. The results were analysed based on two approaches, namely static and dynamic analyses. Countries were divided into five ranks according to their positions, using the static approach. They were also divided into five groups based on their temporal evolution patterns, using the dynamic approach. For each country, the rank and temporal group (see Tables 2 and 3) allowed a comparison of that country's scientific activity in the fields of physics and astronomy with that of the world – or with any other country or group of countries.

A discussion about the scientific activity of a few countries with notable temporal evolution was presented together with the Pm-Cm diagrams (Figures 3 to 7) drawn to appropriate scale. Countries with patterns of fast decline or growth were mentioned, such as the USA – which remains the most active country but shows the fastest decline evolution among all the studied nations. China achieved the fastest-growing evolution; however, it shows considerable deviation from the balanced line where quantity and quality are roughly equal.

The diagrams were produced by focusing on published papers as a main form of scientific production. However, the results shown in the diagrams, in terms of both ranking and temporal evolution, are a fair reflection of the overall scientific activity of countries (in the fields of physics and astronomy). Scientific activity is a complex process that is not easily measured, especially at national level. Many factors ranging from social phenomena to economic trends affect scientific production. Nonetheless, the Pm-Cm diagrams provide a general view of the quality and the quantity of scientific activity in the fields of physics and astronomy for each country.

The findings of this study are likely to interest science policy-makers. The study provides a global context for scientific activity in the fields of physics and astronomy for countries of interest, and the quality and quantity of scientific output are considered simultaneously. The two-dimensional approach thus also provides insight on the balance between quality and quantity for science policy advisors at the national level.[27]

This study might also inspire other researchers to pay attention to the potential of a two-dimensional approach in scientometrics. The two-dimensional approach is an easy and effective method of analysing scientific activity at different levels, while preserving the chronological order of a country's progress. This study focused only on the national level, whereas the H-index diagram has focused on individual researchers. Future studies could consider other levels of analysis such as groups, journals, and universities.

Finally, the results of this study are of interest to a broader audience as they provide an understanding of the position of countries in comparison to each other and to the world average. The paper presents these results both visually (Figures 3 to 7) and as data categories (shown in Tables 2 and 3).

References

1. Cattell J. American men of science. New York: Science Press; 1906.
2. Godin B. On origins of bibliometrics. Scientometrics. 2006;68:109–133. http://dx.doi.org/10.1007/s11192-006-0086-0
3. Godin B. From eugenics to scientometrics: Galton, Cattell, and men of science. Soc Stud Sci. 2007;37:691–728. http://dx.doi.org/10.1177/0306312706075338
4. Godin B. Measurement and statistics on science and technology. London: Routledge; 2005.
5. Organization for Economic Co-operation and Development (OECD). Science and the policies of government. Paris: OECD; 1963.

6. Rostaing H. La bibliométrie et ses techniques [Bibliometrics and its techniques]. Paris: Centre Scientifique de Saint Jérôme; 1996. French.

7. Godin B. The value of science: Changing conceptions of scientific productivity, 1869–circa 1970. Soc Sci Inform. 2009;48:547–586. http://dx.doi.org/10.1177/0539018409344475

8. Vinkler P. The evaluation of research by scientometric indicators. London: Chandos; 2010.

9. Vinkler P. Composite scientometric indicators for evaluating publications of research institutes. Scientometrics. 2006;68:629–642. http://dx.doi.org/10.1007/s11192-006-0123-z

10. Hirsch JE. An index to quantify an individual's scientific output. Proc Natl Acad Sci USA. 2005;102:16569–16572. http://dx.doi.org/10.1073/pnas.0507655102

11. Egghe L. Theory and practise of the g-index. Scientometrics. 2006;69:131–152. http://dx.doi.org/10.1007%2Fs11192-006-0144-7

12. Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals. Scientometrics. 2006;69:169–173. http://dx.doi.org/10.1007/s11192-006-0147-4

13. Bornmann L, Mutz R, Daniel HD. Are there better indices for evaluation purposes than the h index? A comparison of nine different variants of the h-index using data from biomedicine. J Assoc Inf Sci Technol. 2008;59:830–837. http://dx.doi.org/10.1002/asi.20806

14. Van Noorde R. A profusion of measures. Nature. 2010;465:864. http://dx.doi.org/10.1038/465864a

15. Radicchi F, Fortunato S, Markines B, Vespignani A. Diffusion of scientific credits and the ranking of scientists. Phys Rev E Stat Nonlin Soft Matter Phys. 2009;80:056103. http://dx.doi.org/10.1103/PhysRevE.80.056103

16. Nejati A, Hosseini Jenab SM. A two-dimensional approach to evaluate the scientific production of countries (case study: the basic sciences). Scientometrics. 2010;84:357–364. http://dx.doi.org/10.1007/s11192-009-0103-1

17. Harvey L, Green D. Defining quality. Assess Eval High Educ. 1993;18:9–35. http://dx.doi.org/10.1080/0260293930180102

18. Lamont M. How professors think: Inside the curious world of academic judgment. Cambridge, MA: Harvard University Press; 2009.

19. Czellar J, Lanarès J. Quality of research: Which underlying values? Scientometrics. 2013;95:1003–1021. http://dx.doi.org/10.1007/s11192-012-0928-x

20. Macilwain C. What science is really worth. Nature. 2010;465:682. http://dx.doi.org/10.1038/465682a

21. Adam D. The counting house. Nature. 2002;415:726–729. http://dx.doi.org/10.1038/415726a

22. Hosseini Jenab SM. Temporal study of countries' scientific evolution based on two-dimensional approach. J Sci Res. 2013;2:169–172. http://dx.doi.org/10.4103/2320-0057.135406

23. SCImago Journal and Country Rank [homepage on the Internet]. No date [cited 2014 Sep 10]. Available from: http://www.scimagojr.com

24. Committee on Prospering in the Global Economy of the 21st Century. Rising above the gathering storm. Washington DC: The National Academies Press; 2007.

25. Kahn M. A bibliometric analysis of South Africa's scientific outputs – some trends and implications. S Afr J Sci. 2011;107(1/2), Art. #406, 6 pages. http://dx.doi.org/10.4102/sajs.v107i1/2.406

26. Sooryamoorthy R. Scientific research in the natural sciences in South Africa: A scientometric study. S Afr J Sci. 2013;109(7/8), Art. #0001, 11 pages. http://dx.doi.org/10.1590/sajs.2013/20120001

27. Kumar A, Prakasan ER, Mohan L, Kademani BS, Kumar V. Bibliometric and scientometric studies in physics and engineering: Recent ten years analysis. Paper presented at: National Conference on Putting knowledge to work: Best practices in librarianship; 2009 May 1–2; Mumbai, India. http://eprints.rclis.org/14829/1/bibliometric_and_scientometric_bosla-cdac_conf__2009-2.pdf

Multimodal spatial mapping and visualisation of Dinaledi Chamber and Rising Star Cave

AUTHORS:
Ashley Kruger[1]
Patrick Randolph-Quinney[1,2]*
Marina Elliott[1]

AFFILIATIONS:
[1]Evolutionary Studies Institute, School of Geosciences, University of the Witwatersrand, Johannesburg, South Africa
[2]School of Anatomical Sciences, University of the Witwatersrand, Johannesburg, South Africa
*Current address: School of Forensic and Applied Sciences, University of Central Lancashire, Preston, Lancashire, United Kingdom

CORRESPONDENCE TO:
Ashley Kruger

EMAIL:
ashleykruger@gmail.com

POSTAL ADDRESS:
Evolutionary Studies Institute, School of Geosciences, University of the Witwatersrand, Private Bag 3, Wits 2050, South Africa

DATES:
Received: 29 Jan. 2016
Revised: 18 Mar. 2016
Accepted: 09 Apr. 2016

KEYWORDS:
Cradle of Humankind; *Homo naledi*; laser scanning; vertebrate taphonomy; archaeological survey

HOW TO CITE:
Kruger A, Randolph-Quinney P, Elliott M. Multimodal spatial mapping and visualisation of Dinaledi Chamber and Rising Star Cave. S Afr J Sci. 2016;112(5/6), Art. #2016-0032, 11 pages. http://dx.doi.org/10.17159/sajs.2016/20160032

The Dinaledi Chamber of the Rising Star Cave has yielded 1550 identifiable fossil elements – representing the largest single collection of fossil hominin material found on the African continent to date. The fossil chamber in which *Homo naledi* was found was accessible only through a near-vertical chute that presented immense practical and methodological limitations on the excavation and recording methods that could be used within the Cave. In response to practical challenges, a multimodal set of recording and survey methods was thus developed and employed: (1) recording of fossils and the excavation process was achieved through the use of white-light photogrammetry and laser scanning; (2) mapping of the Dinaledi Chamber was accomplished by means of high-resolution laser scanning, with scans running from the excavation site to the ground surface and the cave entrance; (3) at ground surface, the integration of conventional surveying techniques as well as photogrammetry with the use of an unmanned aerial vehicle was applied. Point cloud data were used to provide a centralised and common data structure for conversion and to corroborate the influx of different data collection methods and input formats. Data collected with these methods were applied to the excavations, mapping and surveying of the Dinaledi Chamber and the Rising Star Cave. This multimodal approach provides a comprehensive spatial framework from individual bones to landscape level.

Introduction

The field recovery of hominin fossils and other skeletal remains is a process which owes much of its practical and epistemological background to prehistoric archaeology and archaeological field techniques. As such, searching for, recovering and collecting evidence from a sub-surface or surface scatter of bones are disruptive processes that occur during field investigation. Field archaeology is, by its very nature, a destructive process[1] – once a fossil (or stone tool) has been removed, it can never be put back in its original place, and the link to its context of deposition is broken. As a result, the goal of spatial field recording is to ensure that all items and samples taken from a site can be sourced to their particular context and modelled in three dimensions. Adequate survey and mapping techniques allow investigators to maintain a record of where every item of evidence was found within a defined area, allowing the investigator to recreate the order of events that took place from when the deposit was first exposed, to final recovery. The recording process of any site usually follows the traditional archaeological field recording and requires at least two fixed reference points from which the fieldworker can extrapolate the position of any bone or sample within a deposit, using basic spatial geometry. In many respects the practical fundamentals of field recording of bones, artefacts, dating and sediment samples, is a largely unchanged component of archaeological field skills, and can use items as simple as tape measures, plumb-bobs, and pencil and drafting film, up to the level of electronic distance measuring devices, the Global Positioning System (GPS), photogrammetry and non-contact surface laser scanning.[1] Regardless of the level of technology utilised, it is critical that any recording system allows the archaeologist or palaeoanthropologist to accurately record, plot and model the contexts and contents of a site as precisely as possible in three dimensions. Whether the expected outcome is hand-drawn maps and plans, or three-dimensional computerised models of a site, the fundamental principles are the same.

Historically, work conducted in the Cradle of Humankind with three-dimensional recording methods aimed for visualisation of geological structures, taphonomic associations, and linkages between in-situ and ex-situ fossils. With respect to early hominin sites, mapping the cave systems with reference to the excavation process is essential to understanding the development and movement of sediments and geological units, and has been advanced by several researchers in Africa and Europe.[2-20] Because of the breccified nature of the primary fossil deposits of the Cradle of Humankind (Taung, Gladysvale, Sterkfontein, Kromdraai, Swartkrans, Makapansgat and Malapa), conventional archaeological recovery techniques using traditional methods of recording and planning of discrete contexts (production of single-context plans, overlays and sections, e.g. those used by fieldworkers in the UK) have been difficult to apply in South Africa. Many historical excavations, such as the central component of Sterkfontein[21,22] or Cave of Hearths (Makapansgat)[23,24], have traditionally used an above-site fixed-grid system from which metal poles or markers are suspended, allowing for an approximation of spatial location to the level of the nearest cubic yard or cubic metre, or using offsets from the grid to determine *x-y-z* location within a square or cubic unit with greater precision. With the advent of laser-rangefinder technology later South African excavators have often sought to record the general position of fossils and fossil blocks within the cave by means of broader coordinate systems, using instruments such as electronic distance measuring devices, laser theodolite or total stations (although the adoption of this technology has been relatively late in South African palaeoanthropology as a result of the prohibitive cost of purchasing such equipment).

Three-dimensional digitisation of caves has been undertaken by many researchers since the late 1990s, and has often been used in conjunction with other archaeological and palaeontological data collection methods. Some of the earliest uses of three-dimensional recording within the Cradle of Humankind were undertaken by Quinney and Calabrese in mapping the 'Little Foot' StW 573 skeleton at Sterkfontein in 1998[25], using a Leica total station in conjunction with PenMap surveying software. This undertaking allowed for the precise three-dimensional collection of coordinate data (*x*, *y* and *z* axes) and the reconstruction of contour maps of the exposed geological units within which the skeleton was entombed. Other work (1997 to 2001) by Quinney, Sinclair and colleagues at sites within the Makapansgat Valley

(including the Limeworks australopithecine site and the Cave of Hearths) used the same Leica/PenMap system to assist in the construction of archaeological site plans within individual excavation units, and to combine these with landscape-level surveys of river terraces, sinkholes and fossil-bearing caves, in order to build up broadscale spatial data of the landscape and karstic features.[26,27] In combination with aerial survey data from microlight aircraft, photogrammetry and Geographical Positioning System (GPS) data, the spatial data was integrated into a geographical information system (GIS) model of landscape archaeology in IDRISI.[27]

Later work in 2002 by Lacruz and colleagues[28] at the site of Gladysvale provided a three-dimensional surface (mesh) visualisation of the cave system extrapolated from total station coordinate data, and, more importantly, situated the cave system in direct spatial reference to the current landscape surface. This was followed by Nigro and colleagues[13], who developed an integrated GIS framework in order to map and analyse the fossil deposits at Swartkrans.

Methods of spatial coordinate capture and three-dimensional visualisation of hominin-bearing deposits are now routinely applied at palaeoanthropological fieldschools such as those undertaken at Drimolen and Sterkfontein. All excavations at the site of Malapa (including work on ex-situ dump deposits) have utilised three-dimensional recording from total station equipment (Figure 1). Recent research excavations by Stratford at the site of Sterkfontein have focused on the use of three-dimensional coordinate systems for the mapping of geological units, artefactual and fossil deposits, integrated into a GIS framework.[29] This work has resulted in significant methodological advances compared to earlier approaches and has led to practical frameworks comparable to those applied at European cave excavations.[30] Most recently, Subsol and colleagues[31] applied surface laser scanning to capture the in-situ three-dimensional morphology of the StW 573 skeleton. This work included a test of different laser scanners in order to judge the efficacy and efficiency of different technical set-ups, with resulting recommendations for best practice in such environments.

Such work within the Cradle of Humankind has been mirrored elsewhere in South Africa and beyond, with laser scanning in particular being used in an increasing capacity to record fragile cultural heritage, or to facilitate public understanding of the past through three-dimensional reconstruction. Within cave environments, laser scanning (combined with GIS data) has been applied to spatial reconstructions and mapping, often in conjunction with conventional spatial survey.[2-20,29-32]

Three-dimensional data can also be used in the reconstruction of previously worked sites which were excavated before technology such as terrestrial laser scanning existed. In one such case study, Puchol and colleagues[5] virtually recreated an archaeological excavation site using three-dimensional laser scan data and integrating the data with previously recorded archaeological data, to analyse the spatial context from over 60 years earlier. This amalgamation of 'old' and 'new' data allowed the researchers to extract entire data sets which might not have even been considered in the past. This increases the need for accurate three-dimensional recording of sites, as the better the quality and resolution of recorded data, the easier it will be to relate to new and developing methods of data capture in the future.

Rising Star excavations and Homo naledi

The Rising Star Cave system is located in the Cradle of Humankind World Heritage Site, some 50 km outside of Johannesburg, South Africa (Figure 2). Whilst amateur cavers had been periodically visiting the Cave for a number of years, an incursion by a caving team from the Evolutionary Studies Institute in September 2013 was the first to formally investigate the system for the fossil remains of early hominins. Excavations in November 2013 and March 2014 yielded 1550 identifiable fossil elements representing a minimum of 15 individuals; 300 bone specimens were collected from the surface of the Dinaledi Chamber and 1250 fossil specimens were recovered from a small excavation pit in the chamber floor. This assemblage is the largest single collection of fossil hominin material found on the African continent to date, and has been assigned to the *de novo* hominin taxon, *Homo naledi.* The taxonomy of this new species is detailed by Berger and colleagues[33], with the geological and taphonomic context by Dirks and colleagues[34]. The latter includes a detailed description of the recovery and recording methods applied to the excavation, with aspects detailed in additional commentaries on inferred mortuary behaviours by Randolph-Quinney[35,36].

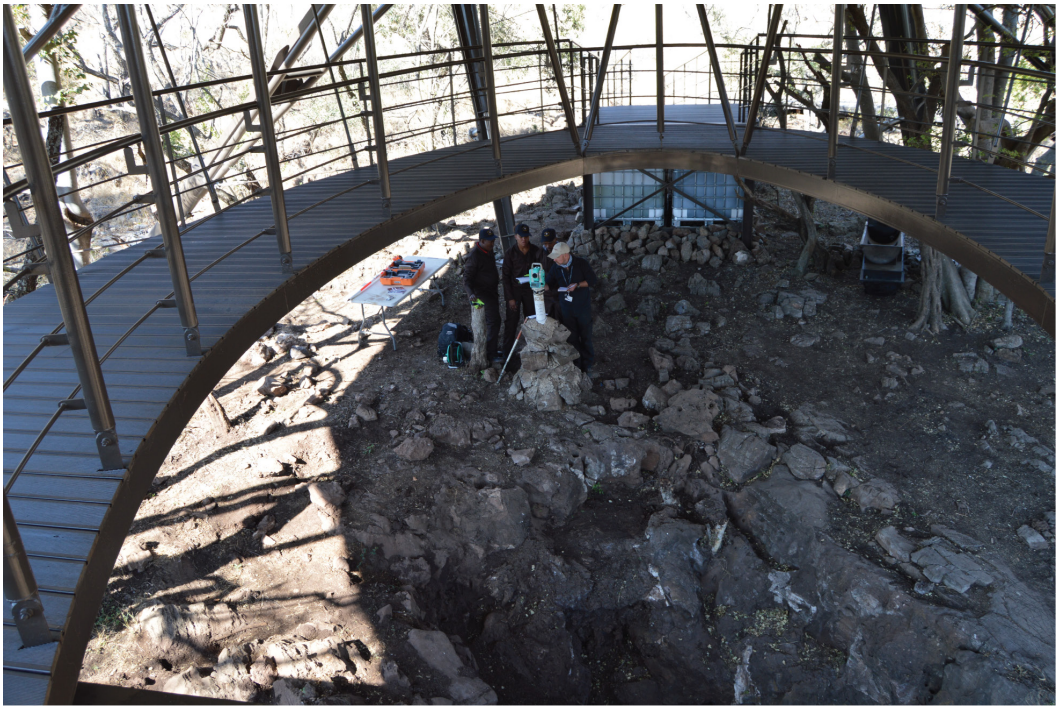


Figure 1: The use of conventional coordinate capture methods during excavations at Malapa Cave, Cradle of Humankind. The excavators are using a Nikon NPR-352 total station mounted on a known-position survey point to record the position of fossil material recovered during the excavation process. Site coordinates (*x, y* and *z*) are converted to geo-reference coordinates (easting, northing and height above sea level in the WGS 84 map reference system) by the total station.

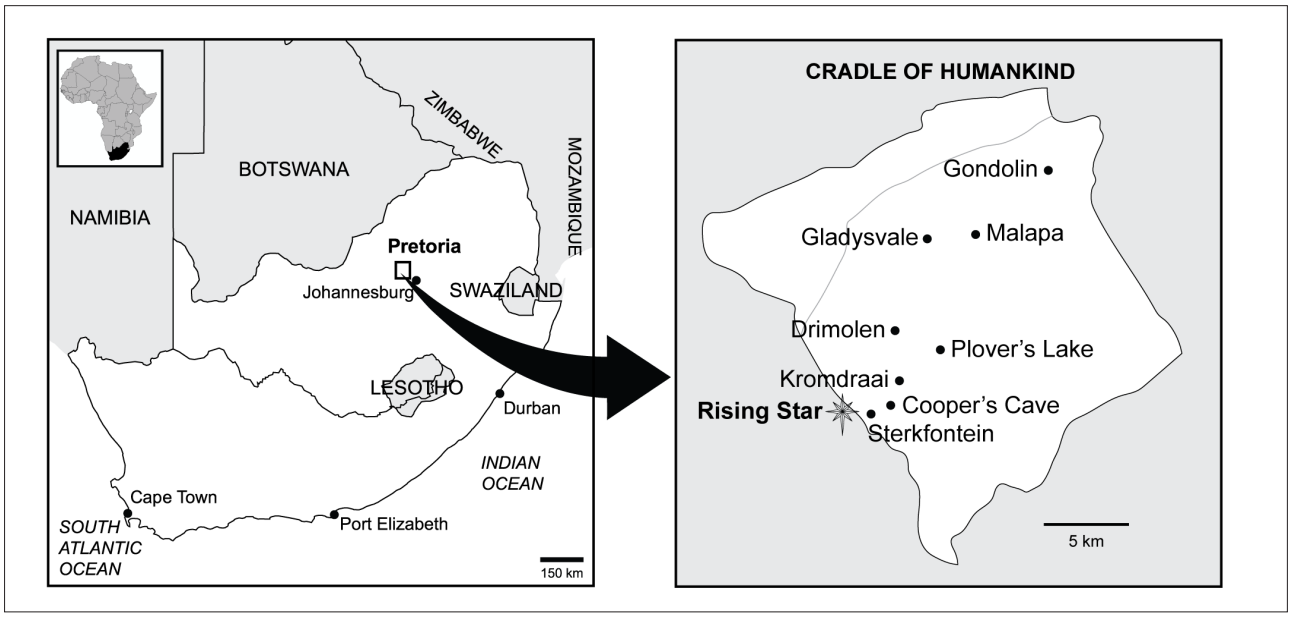


Figure 2: Location of the Rising Star Cave within the Cradle of Humankind, UNESCO World Heritage Site. Other major palaeontological sites are indicated.

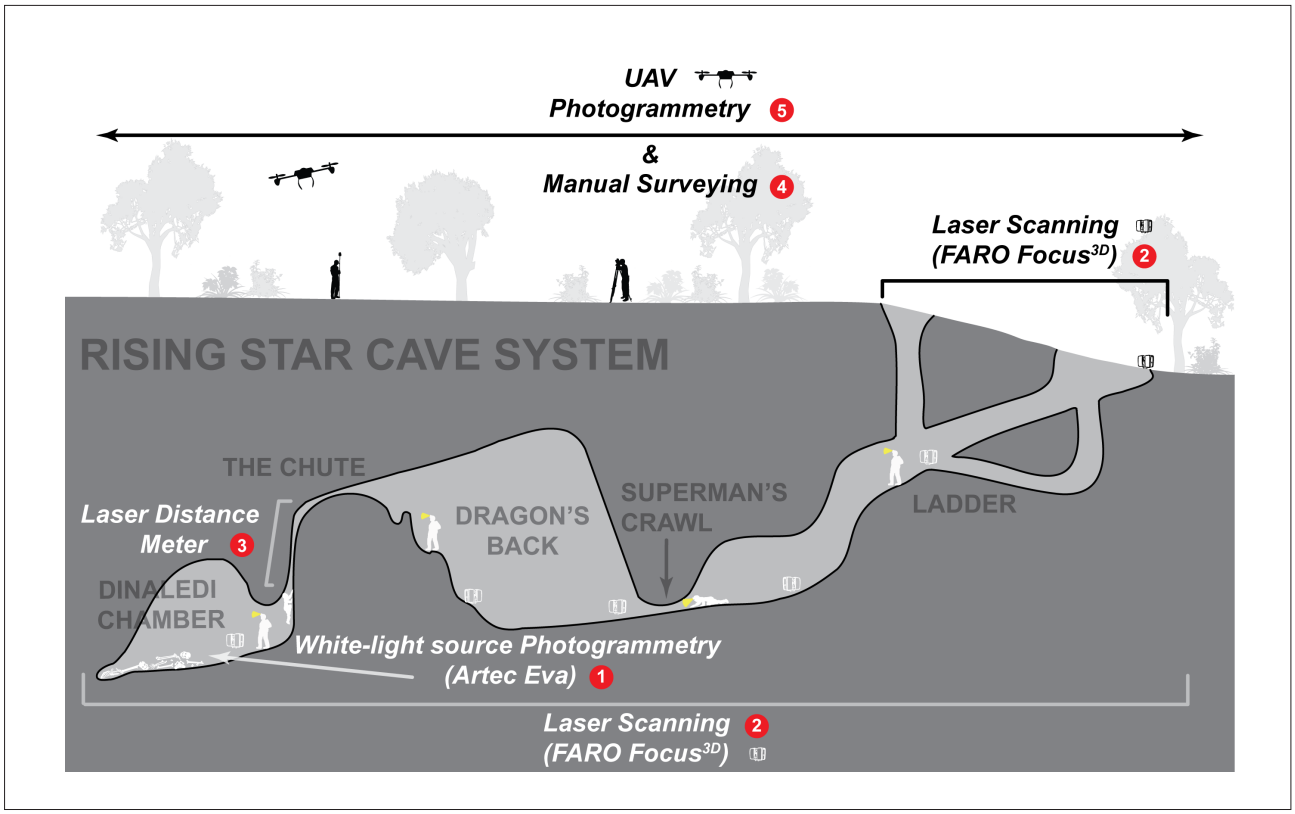
Taphonomically, the assemblage presents a unique situation compared to other South African early hominin-bearing karstic cave sites for three reasons: (1) the hominin remains are numerous and concentrated in a very small area; (2) with the exception of a small number of recent intrusive rodent and bird bones, only hominin bones are found within the Chamber; and (3) the assemblage is unique by what it lacks – there is no evidence of carnivore modifications, sub-aerial exposure or weathering, peri-mortem breakage or trauma, cut marks, or fluvial transport, and no evidence of burning or charring of remains.

Furthermore, the Dinaledi Chamber presents an anomalous depositional environment in comparison to the primary historical hominin sites of the Cradle of Humankind in Gauteng – Sterkfontein, Kromdraai and Swartkrans.[34,37] The latter sites are noted for fossil remains contained in lithified breccia or found in decalcified sedimentary units derived ultimately from clastic breccias. The Dinaledi Chamber is unique in that the fossils were recovered from unconsolidated sediments, and appear never to have been lithified during their depositional history. The basic stratigraphic development of the cave infill is divided into two facies (Facies 1 and 2, with 1 being the oldest) subdivided into three stratigraphic units (1–3, from old to young). Unit 1 comprises reworked laminated mudstones preserved as erosion remnants within the Chamber. Unit 2 is a composite unit that consists of remnant outcrops of variably consolidated sediments containing hominin bones. Unit 3 is the youngest stratigraphic unit, composed of largely unconsolidated sediment derived from weathering and erosion of Units 1 and 2. The majority of the hominin bones was derived from Unit 3. The clay-rich sediments making up the units were derived from in-situ weathering, and from exogenous clays and silts, which entered the Chamber through fractures that prevented passage of coarser-grained materials. Thus, the infill of the Dinaledi Chamber is largely endogenous, fine grained and unconsolidated in nature.[34] Because of the sedimentary environment, recovery of fossils could proceed relatively rapidly using methods more akin to bioarchaeology, rather than the traditional palaeontology of the Cradle in which fossils are extracted from calcified breccia blocks. Unfortunately, the size of the Chamber and the difficulty of the access route precluded the use of conventional surveying equipment such as total stations and necessitated a trade-off in conventional excavation recording strategies. This situation presented immense practical and methodological limitations on the excavation and recording methods that could be used within the cave, where conventional methods of high-resolution spatial data retrieval proved impossible.

Materials and methods

As historical work at sites such as Gladysvale, Makapansgat and Sterkfontein has demonstrated, three-dimensional coordinate data are seen as essential in placing fossil remains in context. At Rising Star, this contextualisation includes collecting three-dimensional data, visualising and understanding the location of fossil material in relation to (1) other material in the Dinaledi Chamber, (2) other chambers within the Rising Star Cave system and (3) the position of the Chamber relative to the landscape surface and surrounding structural geology. Collection of such data demanded high-resolution acquisition, and necessitated the use of equipment of considerable monetary value. Data collection techniques used at the Rising Star Cave ranged from landscape level, down to scanning in-situ fossils (Figure 3). A centralised data structure, in the form of point cloud data, was fundamental to integrating these different formats and data types, many of which have not been integrated in this way at any palaeontological site before.

In response to the practical challenges noted above, a multimodal set of recording and survey methods was developed and employed for use at Rising Star: (1) recording of spatial distribution of the fossils (spatial taphonomy) and the excavation process was achieved through the use of white-light photogrammetry and laser scanning; (2) mapping of the Dinaledi Chamber in landscape context, with scans running from the excavation site to the ground surface and the cave entrance, was accomplished by means of high-resolution laser scanning; (3) at surface level, the integration of conventional surveying techniques as well as photogrammetry with the use of an unmanned aerial vehicle (UAV) was applied. Point cloud data were used to provide a centralised and common data format for conversion and to corroborate the influx of different data collection methods and input formats. Data collected with these methods pushed the boundaries of both technology and image reconstruction as applied to the excavations, mapping and surveying of the Dinaledi Chamber and the Rising Star Cave. This multimodal approach represents the first time that such a seamlessly integrated three-dimensional survey has been applied to a South African hominin site, and provides a comprehensive spatial framework from individual bones within the excavation context, up to the landscape level.



UAV, unmanned aerial vehicle

Figure 3: Graphic representation of the data collection techniques and specific locations in which they were employed to acquire high-resolution data at the Rising Star Cave.

White-light source photogrammetry

The constraints of the Rising Star Cave and Dinaledi Chamber affected both excavators and equipment and demanded an innovative approach to collecting high-resolution spatial data. These data are important for documenting spatial orientation, but can also be used for mapping and surveying. During excavations in November 2013 and March 2014, a handheld Artec Eva (Luxembourg City, Luxembourg; http://www.artecthree-dimensional.com/hardware/artec-eva/) white-light source photogrammetry device was employed for this purpose. 'Painting' or 'sweeping' the desired area with the Artec Eva produces a real-time visualisation of the scan area on an accompanying laptop computer. The Eva has its own white-light source, which is activated in a series of strobe-like flashes during data capturing. Documenting the excavations within the Dinaledi Chamber with the Eva served two purposes. Firstly, spatial data on the location and orientation (axial and surface) of every bone were collected, with a surface resolution of approximately 0.5 mm and a three-dimensional point accuracy of approximately 0.1 mm. Secondly, the data recorded during this process were transported to the surface after every scan, allowing the scientists above ground to provide guidance and direction to the excavators concerning excavations and recovery.

Fossil material from Unit 3 of the Dinaledi Chamber[31] was scanned at approximately 50-mm intervals, or before and after large amounts of material were removed from the sediments. This meant that the excavation recordings were limited to a maximum interval of 50 mm or less. To date, the main excavation area has been excavated to an average depth of 100 mm, and is represented by 151 Artec scans. Fossil material from other parts of the Dinaledi Chamber (i.e. not from the main excavation area, and found mostly at surface level) was also scanned before and after recovery using the same equipment and method.

The post-scan process was managed in Artec Studio 9 (and more recently Artec Studio 10 Professional). Each scan 'sweep' comprised a number of separate images compiled into a single layer. Each scan was then registered to acquire three-dimensional triangulated points. Once registration was complete, the separate layers were manually aligned, using a minimum of three reference points. These reference points were a combination of fixed survey markers within the Dinaledi Chamber, and physical features of the excavation area, which were captured in each scan to provide optimal registration. Once alignment was complete, a global registration process allowed for the scan data to be merged accurately. This produced a three-dimensional mesh representation of the scanned area. The three-dimensional mesh was then overlain with a photographic texture map, captured by the Artec Eva concurrently at the time of scanning. The final product resembles a realistic, photo-like, overlay of the three-dimensional scan. After compiling each three-dimensional scan, the surface scan was exported as a PLY (polygon file format, also known as the Stanford triangle format), which was then stitched together to create a composite surface scan from multiple scans. This composition allowed for the excavation areas and Chamber to be visualised in a three-dimensional environment, through time. Each individual three-dimensional scan, once registered, aligned and fused, was then converted to the E57 file format[38] for three-dimensional imaging data exchange for effortless visualisation and coordinate system manipulation in AutoDesk Recap.

High-resolution laser scanning

Data capture of large areas proved to be cumbersome with the Eva as even small areas (e.g. <1 m^2) of the cave walls or floor, could produce gigabytes of scan data, and in turn many hours of data rendering, fusion and meshing. As a result, a FARO Focus^{3D} X330 laser scanner (Lake Mary, FL, USA; http://www.faro.com/products/three-dimensional-surveying/laser-scanner-faro-focus-three-dimensional/) was used to scan the Dinaledi Chamber, as well as the route between the Chamber and the entrance of the Rising Star Cave system. The Focus^{3D} scanner is physically larger than the Eva and is capable of scanning areas or objects from 0.6 m to a maximum of 330 m. This range was advantageous as large portions of the Rising Star Cave could be scanned relatively quickly.

With a distance accuracy of about 2 mm, the Focus^{3D} scanner was able to produce ultra-high resolution, accurate scans of the cave.

To allow the best possible registration and geo-referencing while scanning in the Rising Star Cave, each scan with the Focus^{3D} scanner was undertaken with the use of 145-mm diameter spheres. These spheres have a magnetic base that attaches to a steel bracket and allows for secure placement. As the cave surfaces are extremely rough and inconsistent, the team modified a number of JOBY GorillaPods (Petaluma, CA, USA; http://joby.com/gorillapod) for positioning the spheres throughout the cave. In this manner, it was possible to attach the spheres securely and safely to various formations and features in the cave. The spheres were placed strategically throughout the cave to complete the scanning as efficiently and accurately as possible, with a minimum of 3 and a maximum of 10 spheres present in each scan.

The Dinaledi Chamber was scanned in its entirety with six scans, each of which took 10–12 min to complete at a resolution of 3.086 mm per 10 m scanned. At this resolution, a point cloud of 173 million points per scan was achieved. A set of 37 scans was completed on a slightly lower resolution of 6.136 mm per 10 m scanned, producing a point cloud of 43 million points per scan. A total of 43 scans was completed, to capture the entire path from the entrance of the Rising Star Cave system to, and including, the Dinaledi Chamber, with the exception of the very confined 'Chute' area (discussed later). Each scan totalled approximately 300 MB and all 43 were combined to create an ultra-accurate point cloud, map and digitisation of the Rising Star Cave. Scans were processed with FARO's SCENE software and were then exported to the Pointools POD file format (.pod) with a compression accuracy of 1 mm. The .pod files were visualised and given a coordinate system with the utilisation of AutoDesk Recap.

UAV photogrammetry

A UAV was used to capture data of the land surface above the Rising Star Cave system. The UAV used was the Aerialtronics ATX8 (Den Haag, the Netherlands) with a full-frame Nikon D610 digital single-lens reflex (DSLR) camera attached. A total of 1031 time-lapse images were taken over three flights, at a height of 103 m. Each flight lasted approximately 8 min. Calibration and corroboration of the data were accomplished using four ground control points, as well as a Nikon GP-1A GPS unit. The program Agisoft Photoscan was used to generate the three-dimensional spatial data, which were then exported as .tiff and .txt data formats. The data were imported to AutoDesk Recap (.rcp) and assigned a coordinate system, which was then overlaid on the Eva and FARO scanner data. In this way, it was possible to accurately locate the Dinaledi Chamber in relation to the ground above it.

Other scanning techniques

The 'Chute' is an extremely difficult and confined 12-m fissure leading down into the Dinaledi Chamber (Figure 4). Despite several attempts, the nature of this passageway proved to be too difficult to scan using either the Artec Eva or FARO Focus^{3D} scanners. Consequently, in order to link the three-dimensional spatial data of the two systems, the Chute was manually mapped and measured using a Leica DISTO S910 laser distance meter. Points were measured at intervals between the top and bottom of the Chute, in the form of *x*, *y* and *z* coordinates at a resolution of approximately 1 mm. Known points collected with the Focus^{3D} were shot in with the S910 and used to align the Dinaledi Chamber scan data with the rest of the laser scan data acquired from the top of the Chute all the way back to the entrance of the Cave.

Manual surveying of the ground above the cave system was also conducted, using traditional surveying techniques. Data acquired in this way were compiled in AutoDesk AutoCad (.dwg file format) and once again overlaid on the data collected with the white-light source photogrammetry, high-resolution laser scanning and UAV photogrammetry. Above ground, two fixed beacons were erected, and assigned UTM coordinates.

Figure 4: Photograph of the small confined nature of the 'Chute' within the Rising Star Cave.

Workflow for three-dimensional data

An integral part of combining the many different formats and scan data presented in this paper was ensuring a centralised and common data structure for conversion. In this case, point cloud data served as an amenable arrangement for data processing, as sets of local, global or arbitrary coordinates could be used to tie data sets together. In addition, a number of different data formats such as .txt, .xyz, .e57, .pod or .rcp can be configured as point cloud data.

A workflow was introduced to handle all the data collected from the excavations and surveys of the Dinaledi Chamber and Rising Star Cave (Figure 5). As noted previously, scan data using the Artec Eva white-light source photogrammetry were processed and rendered in Artec studio. Once the final scan of each layer of excavations was completed and ready for export, a positioning tool was used to set the horizontal plane of the surface. These data were exported to the .e57 file format and imported into AutoDesk Recap as a Recap project (.rcp). Now comprising point cloud data, these data were then given arbitrary, local position coordinates. Coordinates were all set to millimeters and the origin of the coordinates was set to a feature in the cave. This process allowed the excavation area, as well as features and fossil material in the Dinaledi Chamber, to be located and expressed as points, each one with an *x*, *y* and *z* coordinate. By simply selecting a point in Recap, the *x*, *y* and *z* coordinate for a specific point is shown, and the data recorded for spatial data analyses.

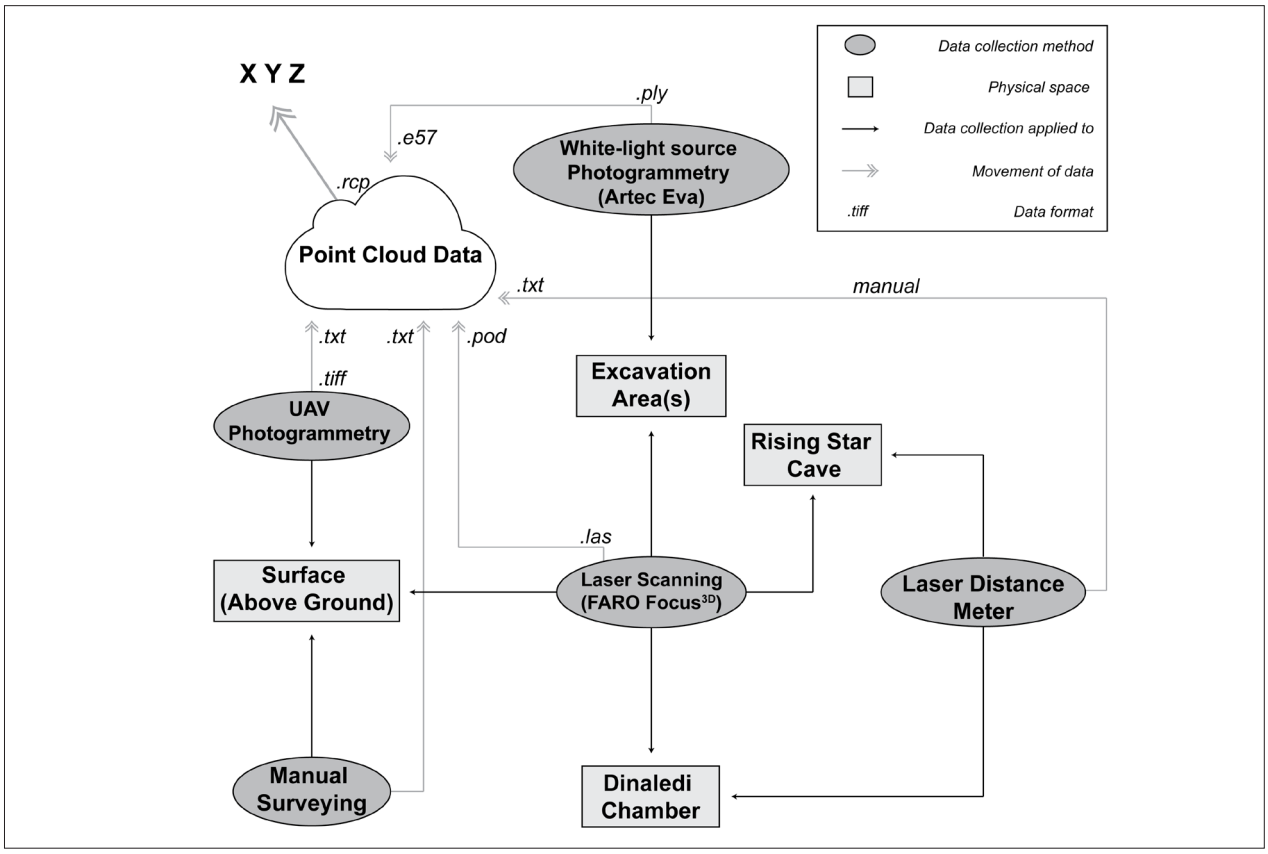


Figure 5: Workflow of the data collection techniques, data formats and structures employed to acquire high-resolution data at the Rising Star Cave.

Data collected with the FARO high-resolution laser scanner were exported as .pod and also imported to Recap as a .rcp project. These data were then linked to the Eva white-light source photogrammetry data by updating a point in the point cloud data, and setting it to the same coordinate as it appeared in the photogrammetry data. In this way, a point in the white-light scans would have the same coordinate *x*, *y* and *z* position as the same point in the laser scans. As data from the excavation area are limited to that region, and are constantly changing as excavations proceed, the linking of this data with the rest of the Dinaledi Chamber and the laser scan data, was essential for alignment and fitting of the coordinates to the larger coordinate system of the Cave's point cloud. Thus, if a point in the white-light source photogrammetry data was 'missing' (for example, a fixed marker was not scanned in a particular scan) it still had a coordinate value in the form of *x*, *y* and *z*.

This procedure allowed fossil material to be assigned an *x*, *y* and *z* value in the form of points. Each in-situ fossil was given a number of coordinates for different spatial attributes, such as proximal, distal or centre of a bone, before excavation. The layered point clouds produced from the white-light and laser scanners were then used to record the fossils' positional attributes in three dimensions while they were still in place.

Laser scan data collected inside the Rising Star Cave system were linked to the ground surface by scanning fixed markers at the entrance to the cave. These reference markers served two purposes. Firstly, they allowed the data collected above ground, by means of manual surveying and UAV photogrammetry, to be synced with data below ground for accurate location of underground features on the surface. Secondly, the reference markers provided a set coordinate system to allow new sections of the cave to be scanned and linked to the existing cave scans, thus expanding the three-dimensional knowledge of the system.

As with the other methods, the UAV photogrammetry data were converted to point cloud data and imported to Recap. Visualisation of these data is more simplified compared to the others as most of it is represented by *x* and *z* coordinates only. By locating the fixed markers at the entrance of the cave, the UAV photogrammetry point cloud data could be assigned the same coordinate system as the other systems, and the coordination of data by the various methods presented in this paper could be achieved. As the UAV captures global coordinates for its data (in the same way as GPS and ground control points are assigned GPS coordinates), it was possible to assign global (GPS) coordinates to the point cloud data, essentially assigning GPS coordinates for fossil material in the Dinaledi Chamber. However, assigning such GPS coordinates to the fossil material offered little benefit above that of assigning arbitrary points and this method was not applied. The GPS coordinate data were, however, instrumental in coordinating the manual surface surveys, in accurately locating the underground chambers, and in visualising underground features in relation to the ground surface. This visualisation is important as the positions of chambers relative to the surface, as well as their relationship to geological features above ground, may bear on how material was deposited or accumulated in the system.

Results

Ongoing analyses are underway to fully describe the context of each fossil element recovered from the excavations in the Dinaledi Chamber. An example of the results obtained by means of the scanning techniques presented here can be seen in Figure 6. Figure 6a is a photograph of the adult 'Hand 1' of *H. naledi*, as described by Kivell and colleagues[39]. Beside it is the hand in situ prior to excavation (Figure 6b). Figure 6c shows the in-situ white-light source photogrammetry mesh accomplished with the Eva after alignment, registration and fusion from four separate scans at an average accuracy of 0.35 mm. The location of the articulated hand is clearly visible, and its spatial context, in relation to the broader excavation area, is captured. Accurate measurements are obtainable through the Artec Studio software, including the distance (in any direction) between specimens within the excavation chamber. Figure 7a shows the broader context of 'Hand 1' in relation to the excavation area and a fixed registration marker ('Pin 2') within the Chamber.

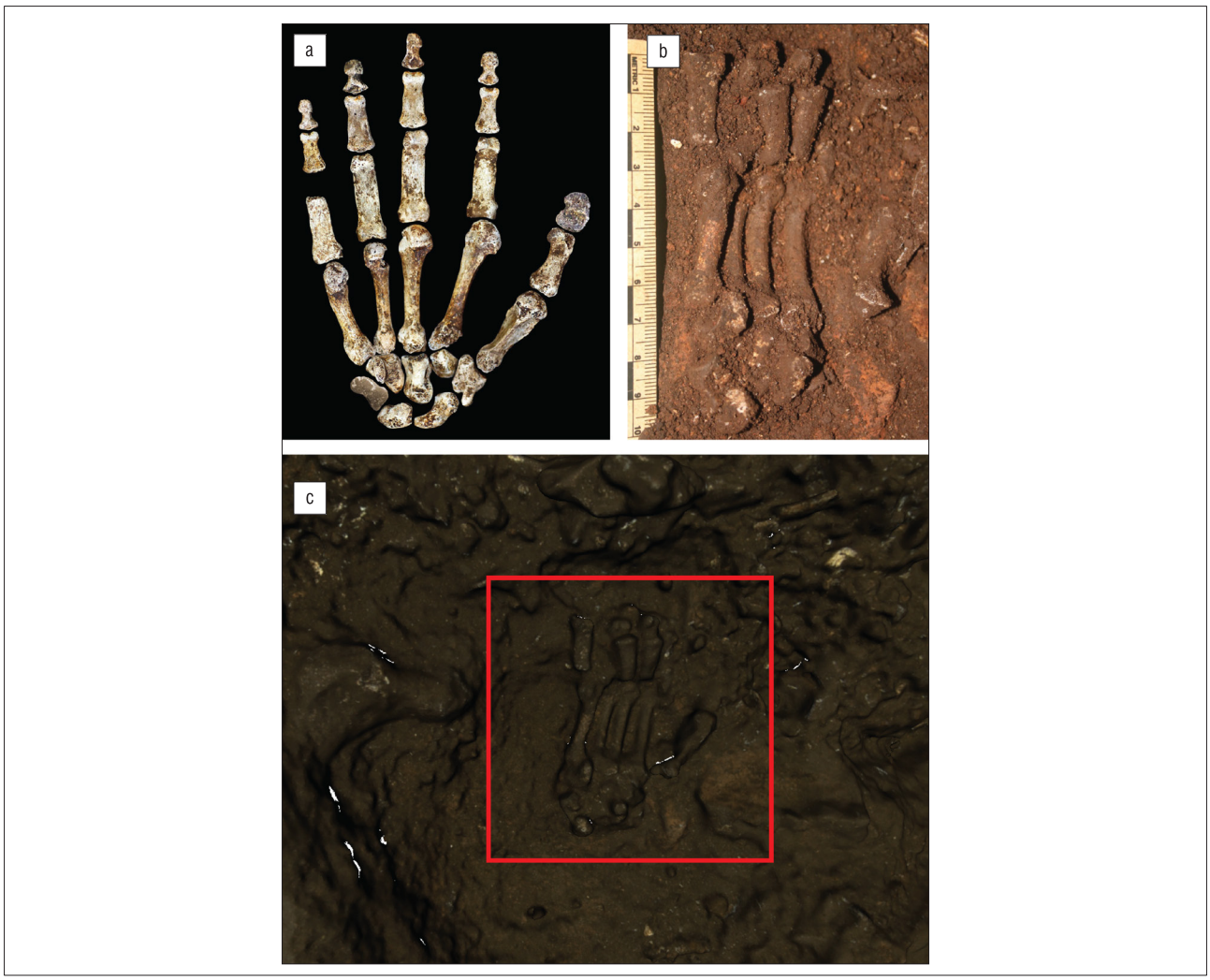


Figure 6: (a) Photograph of the adult 'Hand 1' of *Homo naledi* as described by Kivell et al.[39]; (b) adult 'Hand 1' of *H. naledi* in situ, prior to excavation; and (c) in-situ white-light source photogrammetry mesh of the area surrounding adult 'Hand 1' of *H. naledi* accomplished with the Artec Eva after alignment, registration and fusion of four separate scans at an average accuracy of 0.35 mm.

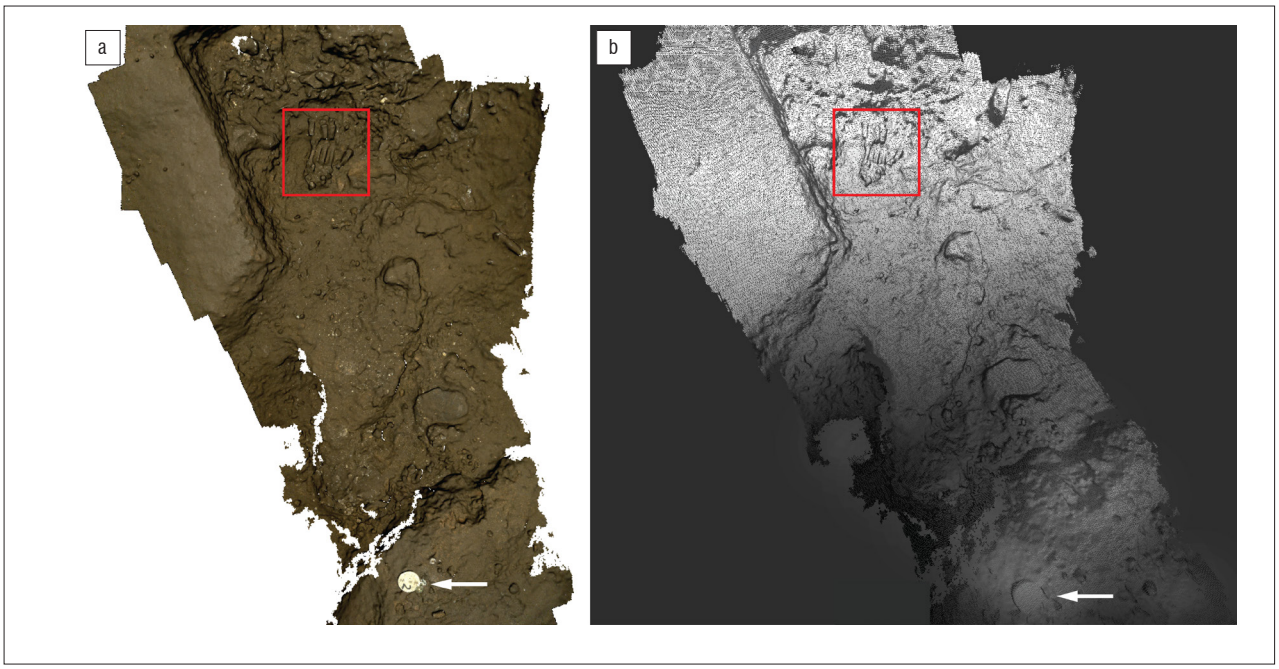


Figure 7: (a) In-situ white-light source photogrammetry mesh of the area surrounding adult 'Hand 1' of *Homo naledi* indicating a reference pin marker (designated with an arrow); and (b) the same mesh converted to the .e57 point cloud data structure, indicating the same pin marker.

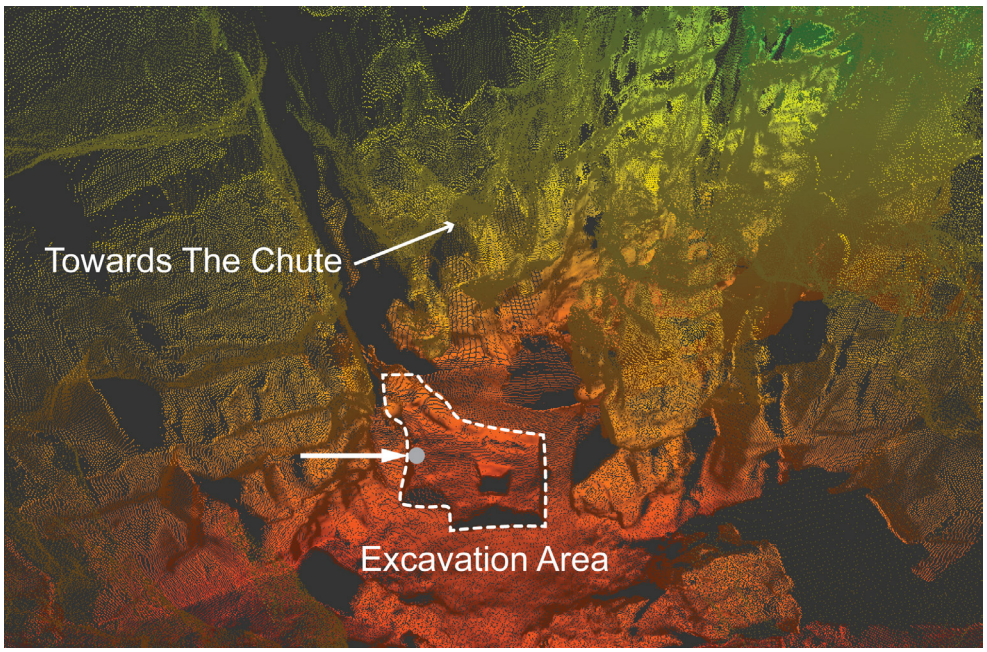


Figure 8: Point cloud data generated from scans conducted with the Faro Focus^{3D} laser scanner indicating the same reference pin marker as shown in Figures 6 and 7 within the excavation area of the Dinaledi Chamber of the Rising Star Cave.

Figure 8 displays one of six point cloud areas generated from the high-resolution laser scanning within the Dinaledi Chamber. As it shows, each point in this cloud has a coordinate, and the location of 'Pin 2' is indicated as in Figure 7. Multiple shared points in each of the different scanning methods allowed for scans and coordinates to be tied together. If a fixed marker within the cave was not represented in an Artec scan, its coordinates were still available in the Focus^{3D} scans. For example, the marker assigned as 'Pin 1' is not present in Figures 6 and 7, which only covers the 'Pin 2' area. However, through the multi-method registration process, fossil material found in scans relating to 'Pin 1' can still be measured and placed in context to fossil material found in scans relating to 'Pin 2'.

Figure 9 shows the accurate placement of the Dinaledi Chamber in relation to the ground surface in high resolution. This image combined data from the Focus^{3D} scans and the point cloud data generated by the UAV photogrammetry and illustrates how the Dinaledi Chamber, and the fossil material recovered from it, relates to the rest of the Rising Star Cave system and the ground above.

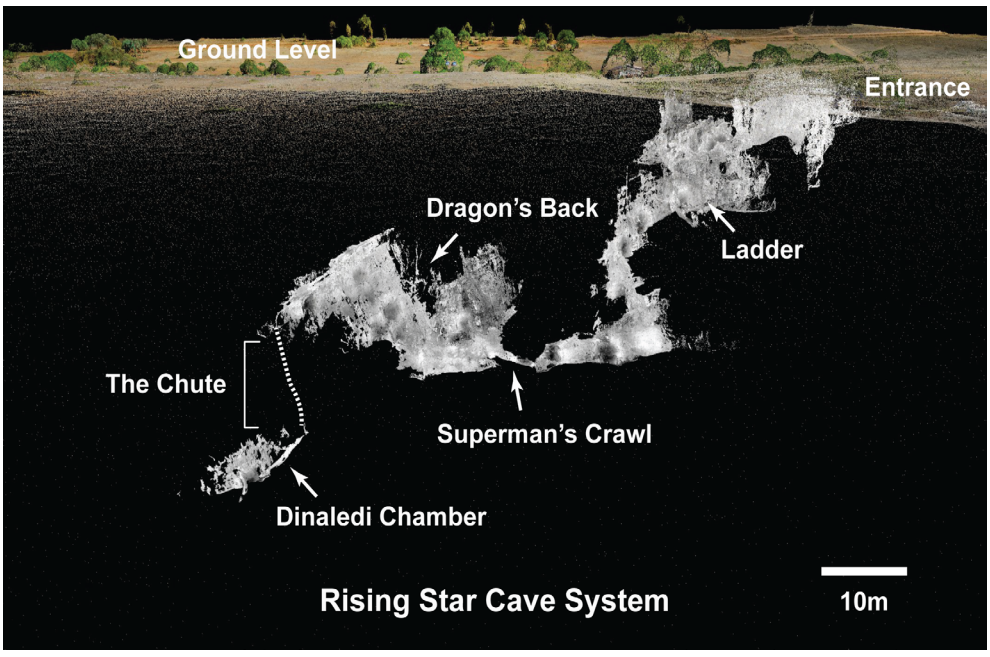


Figure 9: Point cloud data generated from an amalgamation of three-dimensional data collection techniques including unmanned aerial vehicle photogrammetry, high-resolution laser scans and the laser distance meter, showing the accurate position of the Rising Star Cave, including the position of the Dinaledi Chamber, in relation to the above ground level. Scale bar indicates both horizontal and vertical scale.

Discussion

The multimodal approach to the acquisition of three-dimensional data as applied to palaeontology will, we hope, become commonplace, as advancements in technology continue to unfold. As methods for collaborating these data also advance, it would be fortuitous to incorporate these methods into new and existing fieldwork at fossil and archaeological sites. As three-dimensional modelling processes improve, post-processing of extremely large data sets will become much more efficient and less time consuming. At this stage, the requirements for three-dimensional data processing are extremely expensive – in terms of monetary output and processing power and time – and are thus not always attainable. As the advancement of three-dimensional scanners continues, the cost will undoubtedly become more affordable and such equipment will earn its place in archaeological and palaeontological tool kits.

Strengths and limitations of each scanning method

Table 1 highlights the relative format and precision (static accuracy) of each method utilised in the three-dimensional data collection at Rising Star. However, in practice, it was found that each recovery method had a series of inherent strengths and weaknesses in in-field recovery, post-hoc processing, or subsequent data analysis and presentation. The strengths and limitations of the Artec Eva and the FARO Focus^{3D} systems are discussed below.

Artec Eva white-light scanner

Although the small, handheld Artec Eva is a more than adequate scanning device for medium-sized objects (e.g. a human skull), the work resolution and static accuracy is inadequate for rendering smaller objects. For example, isolated human teeth and small fossil fragments were exceptionally hard to visualise in the post-processing stage as they simply did not stand out against the sedimentary 'background'. In addition, the Eva is challenged when presented with right angles, such as the edges of the excavation pit. This inadequacy led to missing data in these areas. The accompanying laptop computer was both a hindrance and an aid during excavations at Rising Star. Because the laptop required an active power source, the Eva unit was unable to operate independently. In addition, the power cabling had to be run into the excavation chamber from the surface because a generator could not be brought into the cave. However, the live display of the scanning area that the laptop provided, allowed the person scanning to make sure the areas or objects scanned were captured correctly. The minimum working distance for the Eva (approximately 400 mm) meant that extremely tight spaces, such as that of the Chute, could not be scanned with this method. The Eva is relatively expensive and perhaps not financially viable for all researchers; however, it is a useful and powerful scanner ideal for creating three-dimensional models of large surfaces as well as smaller objects, not only in the field but also in the laboratory. Rendering larger models takes considerable time, much of which is determined by the computer and graphical processing power at hand.

FARO Focus^{3D} laser scanner

The FARO Focus^{3D} has its own power source, and a replaceable battery pack should extra power be needed. Independent power made for easier scanning in the cave and Dinaledi Chamber, as no cords or cabling were required and the power pack provided power for a substantial period of time. In addition, all data were recorded onto a secure digital (SD) memory card, further reducing its bulk. The Focus^{3D}, however, relies on the use of multiple spheres for registration in the post-processing stage. These spheres are relatively expensive and fragile, which is not ideal in a cave environment. In addition, the requirement for multiple spheres to be visualised in more than one scan presented significant challenges for placement within the tight confines of the cave. The Focus^{3D} is larger than the Eva and requires a tripod and platform in order to conduct its scans. However, its size did not hinder its transportation through the cave, and the tripod–platform assembly was detachable. The need for a line-of-sight connection between the spheres meant that the Focus^{3D} was still incapable of scanning the Chute and *x*, *y* and *z* data for this section of the cave had to be captured with the DISTO S910. This meant that high-resolution data were acquired for the entire path to the Dinaledi Chamber, with the exception of the Chute. The Focus^{3D} is exceptionally expensive, and is usually rented out for a prescribed number of days. Rending the resulting point cloud from scans is remarkably quick.

Combining different survey and recording modalities: A way forward

The large amounts of data collected with the methods presented in this paper, will serve to answer a number of key questions around the formation of the site (UW-101) at Rising Star and the deposition of the material found within the Dinaledi Chamber. Laser scanning has been used successfully in many archaeological sites for GIS[12,13] and heritage preservation[5,9,10,17]. Work conducted at Wonderwerk Cave in the Northern Cape Province of South Africa by Rüther and colleagues[17] shows the extent to which laser scan data can be useful. The main aims of this project were to collect data for the 'African Cultural Heritage and Landscape Database' – a database of information for heritage preservation. Non-contact laser and photogrammetric recovery methods served as an ideal form of data collection as the process is exceptionally rapid, and the data set compiled from it can be used for a variety of other uses, including scientific ones. The data from this project were found to be useful for three-dimensional grid plotting, three-dimensional mapping and the reconstructions of lithic, faunal, botanical and geological associations.[17] As with the work carried out at Rising Star, and the Dinaledi Chamber, the data collected with laser scanning have allowed the creation of a three-dimensional map of the cave system and three-dimensional grid plotting (should it be necessary) and will be used in the mapping of natural and geological features occurring in the Rising Star Cave and their association to features above ground. Accurate placement of the cave system in relation to the ground above it will allow for other technologies to be used on the surface, including ground penetrating radar and geological magnetics studies. We will report on these developments in due course.

Table 1: Comparison of three-dimensional data collection methods used at Rising Star Cave and surrounding landscape

Data collection method	Scanner/ equipment	Applicable area	Data format output (post-processing)	Data format input (point cloud data)	Software for processing	Static accuracy
Laser distance meter	Leica DISTO S910	Chute		*x,y,z* / .txt		1 mm
White-light source photogrammetry	Artec Eva	Dinaledi Chamber	.ply	.e57	Artec Studio 9 / Artec Studio 10 Professional	0.1–0.5 mm
Laser scanning	Faro Focus^{3D} X330	Rising Star Cave; Dinaledi Chamber; above ground	.las	.pod	Faro SCENE	~2 mm
Unmanned aerial vehicle photogrammetry	Aerialtronics ATX8 / Nikon D610	Above ground	.tiff	.txt	Agrisoft	N/A

Acknowledgements

We acknowledge the following for providing funding: DST-NRF Centre of Excellence in Palaeosciences; Lee R Berger Foundation for Exploration; Lyda Hill Foundation; National Geographic Society; National Research Foundation (South Africa); and University of the Witwatersrand. We thank R. Hunter, S. Tucker, D. van Rooyen, W. Crichton, M. Ramalepa and M. Hlophe for assistance in the field during data collection. We thank R. Hugo for his assistance with operating and training of the FARO Focus^{3D}, innovative solutions to positioning calibration spheres and modifying tripods, as well as his dedication to this project. We also thank Group Five Survey Department for the loan of the FARO Focus^{3D} and Aerialtronics ATX8, and processing of laser scan and UAV photogrammetry data.

Authors' contributions

A.K. conducted the research and wrote the original draft of the manuscript. P.R-Q. supervised and developed the original recovery protocols. M.E. supervised the infield recovery of three-dimensional data. All authors contributed equally to manuscript rewrites and editing.

References

1. Hunter J, Cox M. Forensic archaeology: Advances in theory and practice. London: Routledge; 2005.
2. Assefa Z, Pleurdeau D, Duquesnoy F, Hovers E, Pearson O, Asrat A, et al. Survey and explorations of caves in southeastern Ethiopia: Middle Stone Age and Later Stone Age archaeology and Holocene rock art. Quatern Int. 2014;343:136–147. http://dx.doi.org/10.1016/j.quaint.2013.07.132
3. Carvalho M, Fernández S, Pujia L, Rocha C, Rodríguez C, Zelli F. Architecture, archaeology and landscape: An interdisciplinary educational experience in archaeological sites. Procedia Chem. 2013;8:292–301. http://dx.doi.org/10.1016/j.proche.2013.03.036
4. Elliott WR. Cave exploration and mapping in the Sierra de El Abra Region. In: Keene AC, Yoshizawa M, McGaugh SE, editors. Biology and evolution of the Mexican Cavefish. Cambridge, MA: Academic Press; 2016. p. 9–40. http://dx.doi.org/10.1016/B978-0-12-802148-4.00001-3
5. Puchol OG, McClure SB, Blasco Senabre J, Cotino Villa F, Porcelli V. Increasing contextual information by merging existing archaeological data with state of the art laser scanning in the prehistoric funerary deposit of Pastora Cave, Eastern Spain. J Archaeol Sci. 2013;40(3):1593–1601. http://dx.doi.org/10.1016/j.jas.2012.10.015
6. Geiling JM, Marín-Arroyo AB. Spatial distribution analysis of the Lower Magdalenian human burial in El Mirón Cave (Cantabria, Spain). J Archaeol Sci. 2015;60:47–56. http://dx.doi.org/10.1016/j.jas.2015.03.005
7. Gopher A, Parush Y, Assaf E, Barkai R. Spatial aspects as seen from a density analysis of lithics at Middle Pleistocene Qesem Cave: Preliminary results and observations. Quatern Int. 2016;398:103–117. http://dx.doi.org/10.1016/j.quaint.2015.09.078
8. Joukowsky MS. Survey and excavation (field methods) in archaeology. In: Wright, JD, editor. International encyclopedia of the social & behavioral sciences. 2nd ed. Oxford: Elsevier; 2015. p. 742–746. http://dx.doi.org/10.1016/B978-0-08-097086-8.13052-1
9. Leonov AV, Anikushkin MN, Bobkov AE, Rys IV, Kozlikin MB, Shunkov MV, et al. Development of a virtual 3D model of Denisova Cave in the Altai Mountains. Archaeol Ethnol Anthropol Eurasia. 2014;42(3):14–20. http://dx.doi.org/10.1016/j.aeae.2015.04.003
10. Lerma J, Navarro S, Cabrelles M, Villaverde V. Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: The Upper Palaeolithic Cave of Parpalló as a case study. J Archaeol Sci. 2010;37:499–507. http://dx.doi.org/10.1016/j.jas.2009.10.011
11. Melnikova N, Orlov S, Shabrov N, Kiev V, Kuzin A, Resch M, et al. CAVE 3D: software extensions for scientific visualization of large-scale models. Procedia Comp Sci. 2015;66:679–688. http://dx.doi.org/10.1016/j.procs.2015.11.077
12. Neruda P. GIS analysis of the spatial distribution of Middle Palaeolithic artefacts in Kůlna Cave (Czech Republic). Quatern Int. In press 2015. http://dx.doi.org/10.1016/j.quaint.2015.10.028
13. Nigro J, Ungar P, De Ruiter D, Berger L. Developing a geographic information system (GIS) for mapping and analysing fossil deposits at Swartkrans, Gauteng Province, South Africa. J Archaeol Sci. 2003;30:317–324. http://dx.doi.org/10.1006/jasc.2002.0839
14. Núñez M, Buill F, Edo M. 3D model of the Can Sadurní cave. J Archaeol Sci. 2013;40:4420–4428. http://dx.doi.org/10.1016/j.jas.2013.07.006
15. Olivares M, Castro K, Corchón M, Gárate D, Murelaga X, Sarmiento A, et al. Non-invasive portable instrumentation to study Palaeolithic rock paintings: The case of La Peña Cave in San Roman de Candamo (Asturias, Spain). J Archaeol Sci. 2013;40:1354–1360. http://dx.doi.org/10.1016/j.jas.2012.10.008
16. Robu M. The assessment of the internal architecture of an MIS 3 cave bear bone assemblage. Case study: Urşilor Cave, Western Carpathians, Romania. Palaeogr Palaeoclimatol Palaeoecol. 2016;444:115–123. http://dx.doi.org/10.1016/j.palaeo.2015.12.012
17. Rüther H, Chazan M, Schroeder R, Neeser R, Held C, Walker SJ, et al. Laser scanning for conservation and research of African cultural heritage sites: The case study of Wonderwerk Cave, South Africa. J Archaeol Sci. 2009;36(9):1847–1856. http://dx.doi.org/10.1016/j.jas.2009.04.012
18. Sellers W, Chamberlain A. Ultrasonic cave mapping. J Archaeol Sci. 1998;25:867–873. http://dx.doi.org/10.1006/jasc.1997.0232
19. Sellers W, Orton R, Chamberlain A. Computer-aided visualisation of archaeological caves [archive on the Internet]. c2001 [cited 2002 Aug 30]. Available from: http://capra.group.shef.ac.uk/3/sellars.html
20. Stephens M. Cave archaeology and karst geomorphology of Northwest England. In: O'Regan HJ, Faulkner T, Smith IR, editors. Proceedings of the Geologists' Association. 2014;125(3):359–360.
21. Kuman K. The archaeology of Sterkfontein – Past and present. J Hum Evol. 1994;27:471–495. http://dx.doi.org/10.1006/jhev.1994.1065
22. Kuman K, Clarke RJ. Stratigraphy, artefact industries and hominid associations for Sterkfontein, Member 5. J Hum Evol. 2000;38:827–847. http://dx.doi.org/10.1006/jhev.1999.0392
23. Mason RJ. The prehistory of the Transvaal. Johannesburg: Witwatersrand University Press; 1962.
24. Mason RJ. Cave of Hearths, Makapansgat, Transvaal. Johannesburg: Archaeological Research Unit, University of the Witwatersrand; 1983.
25. Clarke RJ. First ever discovery of a well-preserved skull and associated skeleton of *Australopithecus*. S Afr J Sci. 1998;94:460–463.
26. Latham AG, Herries A, Quinney PS, Sinclair A, Kuykendall K. The Makapansgat australopithecine site from a speleological perspective. In: Pollard AM, editor. Geoarchaeology: Exploration, environments, resources. London: Geological Society; 1999. p. 61–77. http://dx.doi.org/10.1144/GSL.SP.1999.165.01.05
27. McNabb J, Sinclair A, Quinney PS. Recent investigations into the Later Acheulean of Makapansgat Region, Northern Province, South Africa. In: Moloney N, Shott MJ, editors. Lithic analysis at the millennium. London: Institute of Archaeology, UCL; 2003. p. 1–16.
28. Lacruz RS, Brink JS, Hancox PJ, Skinner AR, Herries A, Schmid P, et al. Palaeontology and geological context of a Middle Pleistocene faunal assemblage from the Gladysvale Cave, South Africa. Palaeontol Afr. 2002;38:99–114.
29. Stratford D. Cave excavation: Some methodological and interpretive considerations. Cave Karst Sci. 2011;38(3):111–116.
30. Stratford D, Grab S, Pickering TR. The stratigraphy and formation history of fossil- and artefact-bearing sediments in the Milner Hall, Sterkfontein Cave, South Africa: New interpretations and implications for palaeoanthropology and archaeology. J Afr Earth Sci. 2014;96:155–167. http://dx.doi.org/10.1016/j.jafrearsci.2014.04.002
31. Subsol G, Moreno B, Jessel J, Braga J, Bruxelles L, Thackeray F, et al. In situ 3D digitization of the 'Little Foot' *Australopithecus* skeleton from Sterkfontein. Paleoanthropology. 2015:44–53.
32. Birkenfeld M, Avery MD, Horwitz LK. GIS virtual reconstructions of the temporal and spatial relations of fossil deposits at Wonderwerk Cave (South Africa). Afr Archaeol Rev. 2015;32(4):857–876. http://dx.doi.org/10.1007/s10437-015-9209-4
33. Berger LR, Hawks J, De Ruiter DJ, Churchill SE, Schmid P, Delezene LK, et al. *Homo naledi*, a new species of the genus *Homo* from the Dinaledi Chamber, South Africa. eLife. 2015;4, e09560, 35 pages. http://dx.doi.org/10.7554/eLife.09560

34. Dirks PHGM, Berger LR, Roberts EM, Kramers JD, Hawks J, Randolph-Quinney PS, et al. Geological and taphonomic context for the new hominin species *Homo naledi* from the Dinaledi Chamber, South Africa. eLife. 2015;4, e09561, 37 pages. http://dx.doi.org/10.7554/eLife.09561

35. Randolph-Quinney PS. A new star rising: Biology and mortuary behaviour of *Homo naledi*. S Afr J Sci. 2015;111(9/10), Art. #a0122, 4 pages. http://dx.doi.org/10.17159/sajs.2015/a0122

36. Randolph-Quinney PS. The mournful ape: Conflating expression and meaning in the mortuary behaviour of *Homo naledi*. S Afr J Sci. 2015;111(11/12), Art. #a0131, 5 pages. http://dx.doi.org/10.17159/sajs.2015/a0131

37. Dirks P, Berger L. Hominin-bearing caves and landscape dynamics in the Cradle of Humankind, South Africa. J Afr Earth Sci. 2012;78:109–131. http://dx.doi.org/10.1016/j.jafrearsci.2012.09.012

38. Huber D. The ASTM E57 file format for 3D imaging data exchange. In: Beraldin JA, Cheok GS, McCarthy MB, Neuschaefer-Rube U, Baskurt AM, McDowall IE, et al., editors. Proceedings of SPIE 7864: Three-Dimensional Imaging, Interaction, and Measurement; 2011 January 24–27; San Francisco, CA, USA. Bellingham, WA and Springfield, VA: SPIE and IS&T–The Society for Imaging Science and Technology; 2011. http://dx.doi.org/10.1117/12.876555

39. Kivell TL, Deane AS, Tocheri MW, Orr CM, Schmid P, Hawks J, et al. The hand of *Homo naledi*. Nature Commun. 2015;6, Art. #8431, 9 pages. http://dx.doi.org/10.1038/ncomms9431

Does the DHET research output subsidy model penalise high-citation publication? A case study

AUTHORS:
Yolande X. Harley[1]
Esmari Huysamen[1]
Carlette Hlungwani[1]
Tania Douglas[1]

AFFILIATION:
[1]Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Yolande Harley

EMAIL:
yolande.harley@uct.ac.za

POSTAL ADDRESS:
Faculty Research Office, Faculty of Health Sciences, Barnard Fuller Building, University of Cape Town, Anzio Road, Observatory 7925, South Africa

DATES:
Received: 08 Sep. 2015
Revised: 07 Jan. 2016
Accepted: 04 Apr. 2016

KEYWORDS:
higher education subsidy; collaboration; impact; citation; scientometrics

HOW TO CITE:
Harley YX, Huysamen E, Hlungwani C, Douglas T. Does the DHET research output subsidy model penalise high-citation publication? A case study. S Afr J Sci. 2016;112(5/6), Art. #2015-0352, 3 pages. http://dx.doi.org/10.17159/sajs.2016/20150352

South African universities are awarded annual subsidy from the Department of Higher Education and Training (DHET) based on their research publication output. Journal article subsidy is based on the number of research publications in DHET-approved journals as well as the proportional contribution of authors from the university. Co-authorship with other institutions reduces the subsidy received by a university, which may be a disincentive to collaboration. Inter-institutional collaboration may affect the scientific impact of resulting publications, as indicated by the number of citations received. We analysed 812 journal articles published in 2011 by authors from the University of Cape Town's Faculty of Health Sciences to determine if there was a significant relationship between subsidy units received and (1) citation count and (2) field-weighted citation impact. We found that subsidy units had a significant inverse relationship with both citation count (r= -0.247; CI = -0.311 – -0.182; $p<0.0001$) and field-weighted citation impact (r= -0.192; CI= -0.258 – -0.125; $p<0.0001$). These findings suggest that the annual subsidy awarded to universities for research output may inadvertently penalise high-citation publication. Revision of the funding model to address this possibility would better align DHET funding allocation with the strategic plans of the South African Department of Science and Technology, the National Research Foundation and the South African Medical Research Council, and may better support publication of greater impact research.

Introduction

South African universities are awarded annual subsidy from the Department of Higher Education and Training (DHET) based on research publication output – a significant proportion of which is composed of journal article publications. The journal article subsidy is based on the number of journal output units generated by the university, calculated from the number of research publications in DHET-approved journals and the proportional contribution of authors from the university. This subsidy provides financial incentive to increase research output.

Given that the DHET subsidy rewards and intends to stimulate research for the benefit of the country, the manner of awarding subsidy should align with strategies to maintain or improve the impact of South Africa's research. Ideally, university researchers would publish high-quality research that makes an impact in the scientific field, and, where appropriate, work collaboratively with other groups to add value to studies and aid further development and translation of the research.

The goals of South Africa's National Research Foundation (NRF) Strategic Plan[1] incorporate not only research output, but also 'citation intensity', emphasising the importance of the impact of the country's research (not only the volume). The NRF system of rating researchers is also based primarily on the quality and impact of their outputs.[2] The South African Medical Research Council (SAMRC) Strategic Plan 2014/15–2018/19[3] highlights the need to publish in high-impact journals, and includes the number of articles published in the 'top four' journals (*New Engl J Med, Lancet, Science* and *Nature*) as an indicator towards meeting its objectives. The SAMRC also encourages its scientists to work collaboratively, as 'no single group can respond alone to the priorities'[3]. This sentiment is echoed by the NRF Strategic Plan[1], which advocates for 'promoting and enhancing international networks and partnerships', as well as by the South African Department of Science and Technology's (DST's) Ten-Year Innovation Plan[4], which states that greater networking and collaboration (domestic and international) is needed for the country's biotechnology industry to grow.

However, the current subsidy model does not factor in research quality or impact (other than specifying that journals must be DHET accredited). In addition, the greater the co-authorship with other institutions (domestic or international), the lower the subsidy received by a university. This consequence may result in a disincentive to collaboration. It has been argued that the current system may lead to 'non-virtuous practices in research'[5], such as writing short 'salami-sliced' papers, targeting low-tier journals with high acceptance rates, and avoiding collaboration[5,6] to increase subsidy. Cautioning that the drive to increase research volume had come at the expense of the pursuit of excellence, the 2014 Report of the Ministerial Committee for the Review of the Funding of Universities argued it was time to change the funding framework.[5] Yet, the new Research Outputs Policy published in March 2015[7] did not make any changes to the journal article subsidy formula to address these concerns.

Inter-institutional collaboration may affect the scope and quality of research as well as the impact of the resulting publications in the scientific field, as indicated by the number of citations received. Indeed, it has been shown that research that is more collaborative is associated with higher citation rates.[8,9] The inverse relationship between DHET subsidy units received by an institution for a paper and the proportion of authors from outside that university may therefore lead to greater subsidy being awarded to articles of lower citation impact than those of higher citation impact. We hypothesised that greater subsidy (as a result of fewer 'outside' authors) would in fact be associated with lower citation impact. We analysed a set of journal articles published in 2011 to determine if there was a significant relationship between subsidy units received and (1) the number of citations received (citation count) and (2) the field-weighted citation impact. The latter measure is the ratio of citations received by a publication and the average number of citations received by all other similar publications[10] (i.e. with the same publication year, publication type and discipline), and so takes into account differences in citation patterns across disciplines or publication types.

Table 1: Descriptive data for journal articles (n=812) published by authors from the Faculty of Health Sciences at the University of Cape Town (UCT) in 2011 that were approved for subsidy by the Department of Higher Education and Training (DHET)

	DHET subsidy units	Citation count (2011–2015)	Field-weighted citation impact (2011–2015)	Number of non-UCT authors	Proportion of non-UCT authors
Mean±s.d.	0.53±0.35	19±32	2.05±4.29	4±7	0.46±0.35
Range	0.02–1.00	0–356	0–61.71	0–99	0–0.98

Materials and methods

We analysed journal articles published in 2011 by authors from the University of Cape Town's (UCT's) Faculty of Health Sciences that were audited and approved for subsidy by DHET. For each article, we identified the subsidy units assigned by DHET to UCT and the proportion of non-UCT authors. For each article, we extracted (on 15 December 2015) the citation count (total number of citations received) and the field-weighted citation impact for the period 2011–2015 using data drawn from Scopus[11] accessed via SciVal[12]. Articles not listed on the Scopus database were excluded from the analysis.

The relationship between subsidy units assigned to articles and their citation count, as well as between subsidy units and the articles' field-weighted citation impact, were examined using a Pearson correlation, with a two-tailed p-value and a 95% confidence interval.

Results

Following exclusion of 38 articles not listed on the Scopus database, 812 articles were included in the analysis, with a mean subsidy unit assignment of 0.53±0.35 units (Table 1). Of these, 589 (72.5%) articles had non-UCT co-authorship.

Both citation count and field-weighted citation impact were negatively correlated with subsidy units (Figure 1). While the shared variance was small, the correlation was significant in both cases.

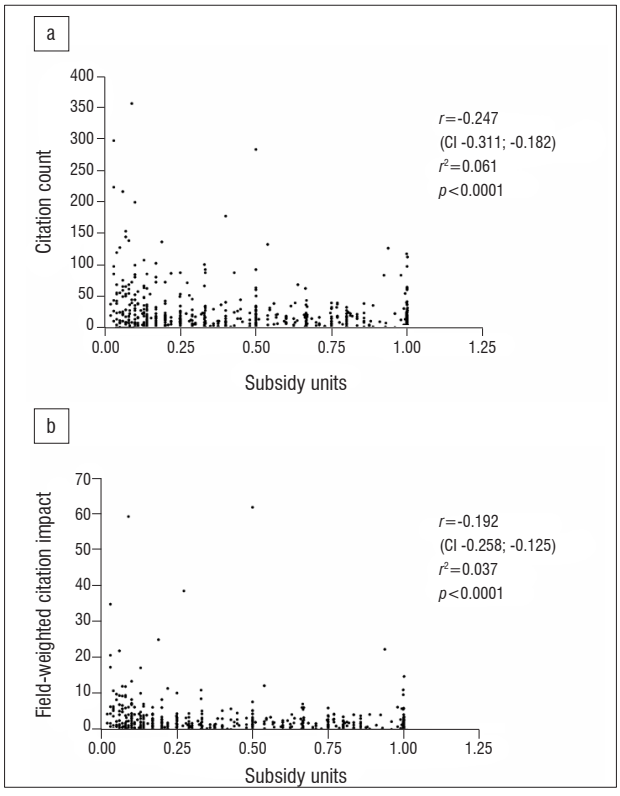


Figure 1: Correlation of the subsidy units assigned to the University of Cape Town for journal articles published by authors from the Faculty of Health Sciences in 2011 with (a) the citation count and (b) the field-weighted citation impact, for the period 2011–2015 (accessed 15 December 2015).

Discussion

DHET subsidy units assigned to UCT for journal articles published by the Faculty of Health Sciences in 2011 had a significant inverse relationship with both citation count and field-weighted citation impact. This finding implies that subsidy allocation is smaller for articles receiving a greater number of citations than it is for those receiving a lower number of citations, whether in absolute terms (citation count) or when compared with the average number of citations received by similar publications[10] (i.e. with the same publication year, publication type and discipline).

Greater collaboration is associated with greater citation rates and, author affiliation aside, citation count tends to increase with the number of authors (with self-citation likely to play only a minor role).[8,9,13,14] It is therefore not unexpected that lower-subsidy-earning publications, which will have been more collaborative, are more highly cited. The analysis has confirmed our hypothesis. By directly relating subsidy to the proportional contribution of authors from a university, and therefore penalising universities for collaborative research, the annual subsidy awarded by DHET to South African universities for research publication output may also be inadvertently penalising high-citation, 'high-impact' publication.

Given the financial benefit of subsidy unit assignment, the existing model discourages inter-institutional collaboration. This situation seems particularly punitive in the case of international collaboration. Universities only receive subsidy for the proportional contribution of authors from that university, whether the external authors are based at other South African institutions or international ones. The annual research publication output subsidy distributed by DHET is currently valued at about ZAR1.6 billion.[15,16] Such large-scale funding should align with the strategic goals of the government, i.e. should incentivise collaborative research that may be associated with high-impact science.

Collaborative publishing has increased in recent years in African institutions, and Pouris and Ho[17] suggest that the large proportion of inter-institutional articles from South African universities indicates that factors encouraging collaboration outweigh the adverse impact of the funding model. Indeed, we found 72.5% of the publications in our analysis had non-UCT co-authorship. That noted, South African universities have different methods of internally allocating subsidy received, and this internal funding distribution could influence their researchers' publishing behaviour. We also note that the nature of research output in the university sector is differentiated with respect to volume, journals, level of collaboration, citation rates and scientific field. Our findings are representative of health sciences articles from a research-intensive university that does not directly allocate publication subsidy to researchers.

Our findings support the recommendation of the 2014 Report from the Review of the Funding of Universities[5] that the funding framework be revised. The Report proposed that subsidy units be divided only among South African authors of articles, so the model no longer actively discourages international collaboration. This revision would better align the DHET model with the DST's Innovation Plan[4], the SAMRC's Strategic Plan[3] and the NRF Strategic Plan[1], which encourage collaboration.

DHET research output subsidy is a means of distributing government funding in a way that factors productivity. Research output has been steadily increasing in South Africa over the last decade and the subsidy system is thought to have contributed to this increase.[5,18] However, quantity should not be emphasised at the expense of quality.[19] It is perhaps notable that while South Africa's medical publication output increased during 1996–2010, the number of citations per document declined.[20] The 2014

Report from the Review of the Funding of Universities[5] also recommends that quality and scientific impact of publications be directly factored in the model; preferential weighting of journals with higher impact factors was suggested. The notion of a journal's impact factor being a measure of the quality of papers published in it has been contested, and deficiencies of this measure as a tool for research assessment have been highlighted.[21] The impact factor may have some utility in the funding framework as an indicator of the quality of the journals in which South African universities publish, especially given research output is aggregated at an institutional level. This is in line with the original intention of the impact factor, namely as a measure of journal (rather than article) quality.[22] Revising the subsidy model to include a weighting for research impact would potentially better align it with the NRF Strategic Plan[1] and SAMRC Plan[3], which advocate for targeting impact. As bibliometric measures of quality and impact evolve in an attempt to minimise inherent flaws (e.g. through the addition of weighting by field), their utility in directing funding allocation should be evaluated on a regular basis.

In summary, the annual subsidy awarded by DHET to South African universities for research publication output may be inadvertently penalising high-citation publication. Revision of the funding model to address this effect would better align DHET funding allocation with government strategic plans and may better support publication of greater impact research.

Authors' contributions

Y.X.H. was the project leader; Y.X.H., E.H., C.H. and T.D. were responsible for experimental design; E.H., C.H. and Y.X.H. collected and compiled the data; T.D. made conceptual contributions; Y.X.H. performed the analysis; Y.X.H. wrote the manuscript; and E.H., C.H. and T.D. gave input on the draft manuscript.

References

1. South African National Research Foundation (NRF). NRF vision 2015: Strategic plan of the National Research Foundation. Pretoria: NRF; 2008.
2. South African National Research Foundation (NRF). Evaluation and rating 2016: Guidelines and instructions for completing the application for evaluation and rating via the NRF Online Submission System. Pretoria: NRF; 2015. Available from: https://www.cput.ac.za/storage/research/research_directorate/funding/apply_nrf_rated_researcher_18022016.pdf
3. The South African Medical Research Council (MRC) Strategic Plan for the fiscal years 2014/15–2018/19. South African Medical Research Council; 2014.
4. South African Department of Science and Technology (DST). Innovation towards a knowledge-based economy: Ten-Year Plan for South Africa (2008–2018). Pretoria: DST; 2007.
5. Report of the Ministerial Committee for the Review of the Funding of Universities. Pretoria: Department of Higher Education and Training; 2013.
6. Macleod C. The incentive system for research is bad for scholarship. University World News. 2010 June 06. Available from: http://www.universityworldnews.com/article.php?story=20100605063654105
7. Research outputs policy. Pretoria: Department of Higher Education and Training; 2015.
8. Khor KA, Yu LG. Influence of international collaboration on the research impact of young universities. In: Salah AA, Tonta Y, Akdag Salah AA, Sugimoto C, Al U, editors. Proceedings of the 15th International Society of Scientometrics and Informetrics Conference; 2015 June 29 – 04 July; Istanbul, Turkey. Istanbul: Bogaziçi University; 2015. p. 278–279. Available from: http://www.issi2015.org/files/downloads/all-papers/0278.pdf
9. Sooryamoorthy R. Do types of collaboration change citation? Collaboration and citation patterns of South African science publications. Scientometrics. 2009;81(1):177–193. http://dx.doi.org/10.1007/s11192-009-2126-z
10. Colledge L. Snowball Metrics recipe book. 2nd ed. Amsterdam: Snowball Metrics program partners; 2014. Available from: http://www.snowballmetrics.com/wp-content/uploads/snowball-recipe-book_HR.pdf
11. Scopus [database on the Internet]. No date [cited 2015 Dec 15]. Available from: http://www.scopus.com/.
12. SciVal [database on the Internet]. c2015 [cited 2015 Dec 15]. Available from: http://scival.com/.
13. Aksnes DW. A macro study of self-citation. Scientometrics. 2003;56(2):235–246. http://dx.doi.org/10.1023/A:1021919228368
14. Van Raan AFJ. The influence of international collaboration on the impact of research results – Some simple mathematical considerations concerning the role of self-citations. Scientometrics. 1998;42(3):423–428. http://dx.doi.org/10.1007/BF02458380
15. Ministerial statement on university funding: 2015/16 and 2016/17. Pretoria: Department of Higher Education and Training; 2014.
16. University performance within block grants: Public report. Pretoria: Department of Higher Education and Training; 2015.
17. Pouris A, Ho YS. Research emphasis and collaboration in Africa. Scientometrics. 2014;98(3):2169–2184. http://dx.doi.org/10.1007/s11192-013-1156-8
18. Pouris A. Scientometric research in South Africa and successful policy instruments. Scientometrics. 2012;91(2):317–325. http://dx.doi.org/10.1007/s11192-011-0581-9
19. Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet. 2014;383(9912):166–175. http://dx.doi.org/10.1016/S0140-6736(13)62227-8
20. Pillay TS. Subject and discipline-specific publication trends in South African medical research, 1996–2011. S Afr J Sci. 2013;109(9/10), Art. #2012-0054, 4 pages. http://dx.doi.org/10.1590/sajs.2013/20120054
21. San Francisco Declaration on Research Assessment (DORA) [document on the Internet]. c2012 [cited 2015 Dec 15]. Available from: http://www.ascb.org/files/SFDeclarationFINAL.pdf
22. Alberts B. Impact factor distortions. Science. 2013;340(6134):787. http://dx.doi.org/10.1126/science.1240319