Scientometrics (2016) 108:1243–1265
DOI 10.1007/s11192-016-2012-4

An instrument to measure individuals' research agenda setting: the multi-dimensional research agendas inventory

Hugo Horta[1] · **João M. Santos**[2,3]

Received: 13 January 2016 / Published online: 25 June 2016

Abstract This study developed the Multi-Dimensional Research Agendas Inventory to measure the key factors associated with the process of research agenda setting. Research agendas reflect the preferences, strategies, influences and goals that guide researchers' decisions to investigate specific topics. The results of exploratory and confirmatory factor analyses indicated that the instrument has eight distinct dimensions: Scientific Ambition, Convergence, Divergence, Discovery, Conservative, Tolerance for Low Funding, Mentor Influence and Collaboration. The model underlying the instrument exhibited a very good fit [$X^2/df = 1.710$; CFI $= 0.961$; PCFI $= 0.791$; RMSEA $= 0.035$; $P(\text{rmsea} \leq 0.05) < 0.001$], and the instrument itself was found to have excellent measuring properties (in terms of validity, reliability and sensitivity). Potential interpretations of the instrument and its implications for research and practice are also discussed in this article.

Keywords Research agendas · Survey instrument · Researchers' research agenda setting · Instrument validation · Confirmatory factor analysis

Introduction

Given the contribution of research to knowledge creation and accumulation, it is important in our increasingly global, fast-paced, multifaceted risk societies to understand the process by which individual researchers set research agendas (Pump 2011). Research agenda

✉ Hugo Horta
horta@hku.hk

[1] Division of Policy, Administration and Social Sciences Education, Faculty of Education, The University of Hong Kong, Pokfulam road, Hong Kong, SAR, China

[2] Centro de Investigação e Intervenção Social (CIS-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Lisbon, Portugal

[3] Center for Innovation, Technology and Policy Research, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

setting by governments, field experts, communities and organisations has received substantial attention in the literature (e.g. Andrews and Johnson 2016). The type of research agenda setting practised by organisations such as research-funding agencies, which is generally associated with government priorities for national development, is relatively straightforward and usually less complex than the process followed by researchers, whose careers increasingly follow non-linear paths (Cantwell 2011). The process by which researchers set research agendas—which drive research—is largely unexplored in the literature (Harris 2001). Although the idea of a research agenda is ubiquitous in contexts and artefacts related to scientific research (e.g. articles and conference proceedings), the term is generally understood intuitively. The concept of a research agenda as instituted by individual researchers has been formally defined in only one recent publication, in which it is envisaged as both a problem-solving framework and a set of actions taken to pursue goals (Ertmer and Glazewski 2014). Based on this definition, a research agenda can be interpreted as a high-level plan implemented via a subset of low-level actions. Although this definition is helpful, a fundamental question remains unanswered: what factors drive researchers' decisions about research agendas, which so powerfully shape the knowledge they produce and ultimately their careers?

This article offers new insights into how researchers set research agendas. An instrument—the Multi-Dimensional Research Agendas Inventory (MDRAI)—was developed to measure the endogenous dimensions of research agenda setting. A questionnaire survey was constructed based on the literature and qualitative data obtained from researchers in the field of higher-education. The data were subjected to exploratory factor analysis (EFA) and structural-equation modelling (SEM), specifically confirmatory factor analysis (CFA). In the next section, a brief overview of the literature providing the rationale for the questionnaire is presented. In the third section, the method is explained. This is followed by a comprehensive analysis of the methodology and results. The article concludes with an overview of the contributions made by the authors and recommendations for future research.

Literature review

Research careers no longer follow linear paths (Cantwell 2011). As researchers' career trajectories become more and more uncertain and complex, the process by which they set research agendas is increasingly determined by their motivation (Bourdieu 1999), incentives (Cole and Cole 1973), collaborations (Mamun and Rahman 2015), career paths (Horta and Yonezawa 2013), influence of their mentors (Pinheiro et al. 2014) and other factors likely to influence research choices.

When considering the relevance of research agendas to a researcher's career, it is vital to acknowledge that the more a researcher has invested in a given field (through learning, researching and publishing), the less likely he or she is to move into other fields (Bourdieu 1999). Specialising in a single field is a well-known predictor of research productivity, as moving into another field is likely to incur hidden transaction costs that may outweigh the benefits of the change (Leahey 2007). Nevertheless, recent studies have shown that approaches to research change over the course of academic careers: researchers usually focus on a single subject while studying for a Ph.D., diverge from this subject to address unrelated issues during their post-doctoral years and then converge to a single research focus later in their careers (Horlings and Gurney 2013). Expertise in multiple fields is particularly desirable given the increasing complexity of problems tackled by researchers today, many of which require multi-disciplinary approaches (Martimianakis and Muzzin 2015; Schut et al. 2014). The changing trajectory of the typical research career suggests that convergence and divergence are two possible—competing, yet concomitant—dimensions of the process of research agenda setting.

As acquiring a position of authority in a field of knowledge takes a long time, convergence is usually attained at a late stage in a researcher's career (in line with the cumulative advantage in scientific fields; see Allison et al. 1982). Most researchers are driven by a certain amount of scientific ambition, and by the desire to be recognised by their peers as authoritative in their respective fields (Merton 1968). Although such recognition increases researchers' status and prestige (Bourdieu 1999; Latour and Woolgar 2013), it may also affect their research-agenda choices. Researchers' social positioning within a field of knowledge may constrain their research agenda setting, as their research may not be completely autonomous—in some cases, it may be significantly influenced by others (such as Ph.D. mentors; see Levitt 2010). Therefore, both scientific ambition and mentor influence must be considered when analysing how individuals set research agendas. In this context, it is also worth noting that the extent of the autonomy that researchers have to set individual research agendas varies across disciplines and their communities. Research agendas tend to be set through interactions with peer communities that are socially and cognitively informed by specific beliefs, traditions, sets of rules, norms and taken-for-granted behaviour (Whitley 2000). To a large extent, the influence of these communities on the research agenda setting of individual researchers is exerted through a consensus over what are the significant research challenges to be addressed (Becher and Trowler 2001); achieving such a consensus is more common in the pure sciences and related fields of knowledge than in the social sciences and the humanities (Becher 1994).[1]

Other relatively endogenous features may also influence researchers' decisions about research agendas. Collaboration has been shown to affect access to resources and ideas (Ebadi and Schiffauerova 2015), publication output and citation outcomes (Horta and Santos 2016), and behavioural and career considerations (Hoffman et al. 2014). It is thus important to determine whether and how collaboration with other researchers affects the process of research agenda setting. Collaboration is considered desirable when tackling multi-disciplinary subjects (Katz and Martin 1997). Additionally, studies of network effects in scientific-collaboration networks have shown that individuals who more frequently collaborate attain a greater visibility and thus are more likely to be invited to participate in future collaborations (Uddin et al. 2013). However, as not all individuals are 'team players' (Barrick and Mount 1991), the collaboration dimension of research agenda setting must be assessed in terms of both willingness and opportunity to collaborate.

Other, more abstract factors reported in the literature are relevant to the development of an instrument measuring researchers' research agenda setting. For example, research agendas in emerging or relatively new fields of knowledge incur greater risk and uncertainty than agendas in well-established fields, due to the greater probability of both dead ends, with no compensation for time and other resources consumed, or substantial rewards for persistence. Both the probability of failure and the probability of success affect research decisions and behaviour (Cummings and Kiesler 2005). Individuals' responses to risk vary

[1] Decision-making processes related to research focus also tend to be collective rather than individual in some fields of knowledge such as biomedicine (Verbree et al. 2015), and are substantially centralised in some fields of knowledge such as physics, particularly in the context of large experimental laboratories (Boisot 2011).

widely, from risk seeking to risk aversion (Hillson and Murray-Webster 2007). Some researchers are less willing than others to pursue or persist with research in high-risk fields. For example, researchers in biomedicine have been shown to pursue conservative research strategies, which become more conservative over time and are considered to be a safer choices for careers, even if these strategies do not significantly advance the field (Rzhetsky et al. 2015). Perceptions of risk have also been shown to vary based on the nature of the risk (Slovic et al. 1982), the resources available (mostly financial; see Ebadi and Schiffauerova 2015) and the amount of information provided, all of which influence a researcher's decision to implement a more or less risky research agenda (assuming a bounded-rationality approach; see Simon 1992). Therefore, it is important to consider research area (e.g. emerging or mature) and limited funding as separate yet equally integral determinants of researchers' risk propensity and thus their choice of research agendas. Such a measure, at an individual level, would complement measures of conformity associated with researchers' belief systems, that is, the tendency for research practices and outputs to reinforce existing knowledge rather than develop innovative findings (Klavans et al. 2013).

In the following sections, the methodology and operationalisation of the study's constructs are discussed.

Methods

Structural equations modelling

This study was largely conducted using Structural Equations Modelling (SEM), specifically, confirmatory factor analysis (CFA), using the statistical software package AMOS 22. In this section, a brief overview of the procedure is provided to assist readers unfamiliar with the procedure to better understand the following sections.

SEM is a modelling technique used to test hypothetical causal relations between variables. It extends traditional generalised linear modelling and exploratory factor analysis (EFA) techniques by combining the strengths of both methods. (For an in-depth analysis of the applications of SEM techniques, see Arbuckle 2007; Bollen 2014; Jöreskog and Sörbom 1989; Kline 2011; Marôco 2010).[2] SEM has two main advantages over traditional methods: (1) the capacity to specify latent variables, which are variables that cannot be directly observed but can be estimated using other variables, similar to disturbance terms (Bentler and Weeks 1980), and (2) its incorporation of multiple traditional linear modelling techniques, such as analysis of variance, analysis of covariance and linear regression, into a single analytical method (Marôco 2010). In addition, SEM provides a vast number of fit indicators that can be used to evaluate and further refine a model. SEM also mitigates the over-inflation of disturbance terms by allowing researchers to consider systemic relations between variables that are difficult to detect or specify using traditional linear modelling techniques (Bollen 2014; Maroco 2007; Marôco 2010).

CFA is a specific case of SEM, in which the model can be written as follows (Bollen 2014; Marôco 2010):

[2] David Kenny (whose work on linear modeling is seminal) has maintained very comprehensive and up-to-date guidelines for SEM on his personal webpage, http://davidakenny.net/cm/causalm.htm, which may be useful to readers interested in learning how to operate SEM software.

$$X = \Lambda_x \xi + \varepsilon$$

where X is the vector for the manifest variables; Λ_x is the matrix for the factorial weights of ξ in x; ξ is the vector for the latent variables; and ε is the disturbance term.

CFA is typically (but not necessarily) used as a follow-up to a more traditional modelling method, EFA. The critical difference between the two types of factor analysis is that whereas EFA allows variables to be loaded freely into any of the extracted factors (hence 'exploratory'; the procedure is used to extract an otherwise unknown structure) (Maroco 2003), CFA requires the factorial structure to be specified a priori, creating constraints on the variables' factorial loading (Brown 2015). The specification of the model is typically based on insights obtained previously using EFA (and may also be used to test a structure identified via EFA) or findings reported in previous studies. Either way, CFA can be used to confirm the specified factorial structure.

Maximum Likelihood (ML) estimation was used in this study because it is robust to deviations from multivariate normality, which makes it safe for use in most analytical contexts (Marôco 2010). Details of the implementation of ML estimation using SEM software and the calculations underlying this method can be found in numerous related books and articles (see, for example, Arbuckle 2007; Jöreskog and Sörbom 1989).

Model estimation is generally followed by fit evaluation. A vast number of fit indicators are available, typically categorised by the functions they serve or the dimensions of fit they evaluate. There is no widely accepted set of 'best' indicators; researchers simply tend to choose those with which they are most familiar, selecting one indicator from each category to ensure a more comprehensive fit evaluation (Bentler 1990). For the purposes of this study, the most commonly used indicators are reported: the X^2/df indicator (Arbuckle 2007; Barrett 2007; Bentler 2007; Marôco 2010); the comparative-fit index (CFI) (Bentler 1990); the parsimony CFI (PCFI) (Marôco 2010); the root mean square error of approximation (RMSEA) (Steiger et al. 1985); the Akaike information criterion (AIC); and the Browne-Cudeck criterion (BCC) (Anderson, Burnham, and White 1998; Marôco 2010).

If the model fit is poor, the model can be re-specified to improve its fit with little effort. The first and most conservative strategy for improving model fit involves eliminating non-significant trajectories or trajectories with low loadings (Marôco 2010), followed by re-specification typically based on modification indexes (MI). MIs are used to estimate the relative change in the X^2 statistic when parametric or trajectory adjustments are made to the model. It is primarily an optimisation procedure, but cannot be performed automatically, as adjustments that benefit the model statistically may be theoretically implausible. Therefore, the researcher must carefully consider which adjustments to the model make sense (Arbuckle 2007). In AMOS 22, MIs are implemented using the Lagrange multipliers method, as described by Bollen (2014). MI adjustments are conducted iteratively. In the first pass, only adjustments with the highest MI value are performed, followed by a re-estimation of the model and a re-evaluation of the fit and MIs. This process is repeated until optimal fit is attained. Typically, the first pass involves adjustments with an MI value higher than 11, which corresponds to a type I error probability of 0.001; the second pass involves adjustments with an MI value higher than 4, representing a type I error probability of 0.05 (Marôco 2010).

Pilot study

Prior to the main study, a small pilot study was conducted. In a first step, a large pool of questions was drafted based on the literature. At this stage, the questions were discussed informally with researchers and academics to obtain their feedback, which was used to refine the structure and content of the questions (see Kassam et al. 2012). Next, preliminary validation exercises were conducted with the goal of reducing the initial number of questions and obtaining preliminary insights at a structural level. In this section, the procedures and results of this process are briefly described.

Question drafting

The initial set of questions was based on themes that emerged from the literature review. Further discussion of these themes with colleagues led to an initial draft of the questionnaire. This initial version contained 84 Likert-style questions (answers ranging from 1 to 7, with 'don't know' options) with a mixture of true-scored and reverse-scored items, divided into seven blocks according to the following themes: scientific ambition; convergence; divergence; risk propensity (field); risk propensity (funding); mentor influence; and collaboration.

The Scientific Ambition block contained questions measuring the participant's desire to excel in the field and gain recognition for his or her scientific endeavours (e.g. 'I aim to be recognised by my peers'). The Convergence block contained questions regarding specialisation in a single field of science (e.g. 'I have mastered a single scientific area'). Conversely, the Divergence block comprised questions on diversification (e.g. 'I would be interested in pursuing research in other fields'). The Risk Propensity (Field) block dealt with the participant's attitude toward fields of science whose outcomes are considered risky or uncertain (e.g. 'I find "cutting-edge" scientific areas more appealing than well-established ones'). The questions in the Risk Propensity (Funding) block also addressed risk perception, but dealt with fields with limited funding (e.g. 'Limited funding does not constrain my choice of field'). The Mentor Influence block contained questions regarding the degree to which the participant's Ph.D. mentor continues to influence his or her decision making (e.g. 'My Ph.D. mentor's opinion carries much weight in my research choices'). The questions in the Collaboration block dealt with the participant's willingness and opportunity to engage in collaborative work (e.g. 'I often seek peers with whom I can collaborate on scientific articles').

Preliminary test

A key issue that emerged during the pilot study was that an 84-question survey was too long to be of practical use. Reducing the number of items used in the final survey was necessary. As long questionnaires can have poor response rates, one of the primary tasks at this stage was to reduce the number of questions per theme (for similar work see, for example, Rammstedt and John 2007). A preliminary test was conducted in May 2015 to obtain initial feedback on the questionnaire, gain insights into the factorial structure and reduce the number of questions. The original 84 questions were given to a limited sample of 43 researchers in a range of fields, who were affiliated with various institutions worldwide. The respondents were asked to provide qualitative feedback on the questionnaire in addition to their survey responses. The 84 questions were presented in random order for each participant.

The data obtained in this preliminary deployment were analysed by EFA using Varimax rotation and subsequently CFA. As the small sample size did not allow factor analysis to be conducted on all 84 questions simultaneously, analysis was performed separately on each of the seven blocks, with each block containing 12 questions. We had two goals at the EFA stage: (1) to perform a first pass of question elimination, and (2) to obtain insights into the underlying lower order factorial structure (see Bentler and Weeks 1980). Anti-image matrices were produced and questions that had a measure of sampling adequacy (MSA) value smaller than 0.50, indicating poor fit (Maroco 2003), were eliminated. After removing some questions and performing EFA again with Varimax rotation (Ebrahimy and Osareh 2014), the optimal numbers of factors and corresponding questions were determined based on three criteria: (a) the Kaiser criteria; (b) scree-plot analysis; and (c) factor and total extracted variance. The extracted factors were labelled according to the content and themes of their highest-loaded constituent questions.[3] Subsequently, a model was specified and estimated using the extracted structure, and subjected to preliminary CFA. At this stage, questions with factor loadings under 0.50 were removed. Additionally, items with MIs that suggested implausible correlations were eliminated (Marôco 2010). Finally, once all of the problematic questions had been excluded, the items with the lowest factor loadings were removed until only six items remained, all containing the same number of lower-order factors. The goal was to maintain a balanced number of items per factor, facilitating the calculation of composite scores (DiStefanoet al. 2009). The choice of six items was thus determined by the factor with the smallest number of non-problematic items (the scientific ambition factor). Forty-two questions were removed, leading to a final pool of 42 questions. Finally, Cronbach's alpha was computed to measure the questionnaire's reliability. The findings of these preliminary tests are summarised in Table 1.

Several issues emerged from the preliminary analysis. First, certain reverse-scored items were loaded into entirely separate factors. This occurred in both Risk Propensity scales, in the Mentor Influence scale and in the Collaboration scale. This effect has been documented in the literature, and occurs when the reverse wording used in a question is perceived as the exact opposite of the true wording (Spector et al. 1997). In the Risk Propensity scales, for example, risk-seeking behaviour (true-scored questions) and risk-averse behaviour (reverse-scored questions) were probably perceived by the participants as incompatible. As treating the two sets of questions as a single factor worsened the fit, we opted against merging them into a single factor at this stage. During CFA, the reverse-coded factors that emerged in the mentor influence and collaboration scales were found to be non-significant and detrimental to model fit, and were thus removed entirely from the analysis. Finally, EFA indicated an additional factor in the Risk Propensity (Funding) scale, which we labelled Competition. This factor comprised questions on the participants' perceptions of competition for funding. However, CFA revealed that this factor was non-significant and detrimental to fit, so it was also removed from the analysis.

Main study

Following the pilot study, a tentative final questionnaire was distributed to a much larger sample to conduct a full validation exercise, including EFA, CFA, validity, reliability and sensibility evaluation. In this section, both the procedures and the results of this exercise are described in detail.

Procedures

Before distributing the tentative questionnaire, a search was conducted on the Scopus database in May 2015 to identify the corresponding authors of all articles published

[3] More in-depth information on these lower-order factors is provided in later sections of this article.

Table 1 Preliminary results

Factor	Average loading	Cronbach's alpha	Number of items
Scientific ambition	**0.788**	**0.906**	**6**
Prestige	0.866	0.897	3
Scientific recognition	0.868	0.893	3
Collaboration	**0.718**	**0.873**	**6**
Willingness to collaborate	0.866	0.900	3
Invited to collaborate	0.756	0.799	3
Convergence	**0.776**	**0.905**	**6**
Mastery	0.903	0.928	3
Stability	0.803	0.796	3
Divergence	**0.765**	**0.911**	**6**
Branching out	0.875	0.866	2
Multi-disciplinarity	0.955	0.953	2
Flexibility	0.915	0.910	2
Mentor influence	**0.786**	**0.906**	**6**
Risk propensity (funding)	**0.630**	**0.822**	**6**
Risk seeking	0.863	0.899	3
Risk aversion	0.751	0.810	3
Risk propensity (field)	**0.785**	**0.906**	**6**
Risk seeking	0.900	0.951	3
Risk aversion	0.876	0.906	3

Note Values in bold indicate single-factor parameters

between 2004 and 2014 with 'tertiary education' or 'higher education' in the journal title. There were two methodological reasons to restrict the respondents to higher education researchers. First, as the responses were highly likely to differ considerably between fields of knowledge, we restricted the participants to a single field to prevent inter-field variability.[4] Second, the authors have published extensively on higher education research (Horta and Jung 2014; Jung and Horta 2013, 2015), and their knowledge and expertise in this field makes it an ideal choice for an exploratory study. The Scopus search yielded 5985 authors. The sampling process was conducted in a non-probabilistic manner through availability sampling, as all of the matching authors were invited to participate.

The MDRAI was implemented via an online surveying platform. Invitations to participate in the study were sent by e-mail in a series of waves between June and November 2015; each e-mail provided a description of the survey's purpose and a link to the platform. An opt-out link was also provided for recipients who did not wish to be contacted again. The recipients who followed the link to the platform were directed to a page containing an informed consent letter; they were required to provide their informed consent before proceeding to the survey itself.

A minimum of 500 subjects was considered necessary to adequately conduct the analysis (as recommended by MacCallum et al. 1999). Of the 1348 researchers who agreed

[4] However, higher education research is to some extent multi-disciplinary, with contributions from most of the social science fields (e.g. economics, political science, sociology and psychology). In future research, using a new set of data, the authors will carry out further validation exercises with different cohorts (in this case, academics from other fields) to maximise the robustness of the instrument.

to participate, 416 did not complete the MDRAI questions, and were thus excluded from further consideration, leading to a sample size of 932, meeting the proposed threshold. Females comprised 495 (53.1 %) of the participants, and the remaining 437 (46.9 %) were male. The participants ranged from 24 to 84 years old ($M = 51.01$, $SD = 11.23$). The majority of the participants were affiliated with institutions in the United States (231; 24.8 %), followed by those with affiliations in Australia (143; 15.3 %) and the United Kingdom (127; 13.6 %). Scholars in these countries have produced the majority of publications on higher education research worldwide (Kosmützky and Krücken 2014). The remaining 431 (46.3 %) participants were affiliated with institutions in 65 other countries.

For the purpose of cross-validation, the sample was randomly split into two sub-samples, in line with similar studies (e.g. Johnson and Stevens 2001). The original sample was further randomly divided into a training sample, which contained approximately 40 % of the individuals ($N = 342$) and was used for the exploratory factor analysis, and a holdout sample, which contained the remaining 60 % of the participants ($N = 590$) and was used for the confirmatory factor analysis.

Imputation

Missing values were handled using a series of data-imputation techniques. A Markov Chain Monte Carlo (MCMC) multiple imputation procedure was conducted to obtain five complete EFA datasets. The datasets were subjected to further analysis, and the pooled estimates and parameters for all five were used for decision-making and reporting purposes. For the structural-equation models, imputation was conducted via the Full Information Maximum Likelihood (FIML) estimation, which is considered superior to other imputation techniques (Enders and Bandalos 2001). However, it was not possible to use FIML estimation to calculate the MIs due to computational limitations; one of the MCMC datasets was used instead.

EFA

Prior to CFA, an EFA was conducted on the training sample, using a method similar to that used in the preliminary tests, to obtain a tentative structure for specification by CFA. The procedure followed here was the same as that previously described.[5] As most of the potential problems with fit had been dealt with in the preliminary test, no problematic items were identified at this stage (based on the MSA criteria). However, the primary loadings of some items shifted toward other factors, and some factors collapsed altogether. The most notable differences were as follows. The loading of one item shifted from the Prestige factor toward the Scientific Recognition factor. As the two remaining items measuring the prestige factor were related to publishing scientific articles, we renamed this factor Drive to Publish. One item from the Invited to Collaborate factor shifted toward the Willingness to Collaborate factor. The lower-order constructs on the Convergence factor collapsed into a single factor, with predictably lower factorial loadings and reliability. Two-factor extraction was used to replicate the structure identified in the preliminary tests, and yielded a greater percentage of explained variance at the cost of lower factorial loadings for some of the items, along with unacceptably low levels of reliability. As a result, we opted to carry out CFA with a single-factor structure for these items. Our findings are summarised in the following table (Table 2).

[5] Although the sample size at this stage allowed EFA to be conducted on all of the items simultaneously, we opted to perform EFA with separate question blocks, as in the preliminary test, to ensure consistency.

Table 2 Summary of EFA results

Factor	Average loading	Cronbach's alpha	Number of items	% Variance explained
Scientific ambition	**0.758**	**0.846**	**6**	**57.311** %
Scientific recognition	0.840	0.854	3	57.311 %
Drive to publish	0.670	0.719	3	18.419 %
Collaboration	**0.840**	**0.915**	**6**	**70.472** %
Willingness to collaborate	0.808	0.900	4	70.472 %
Invited to collaborate	0.892	0.915	2	11.913 %
Convergence	**0.647**	**0.714**	**6**	**42.683** %
Divergence	**0.682**	**0.755**	**6**	**47.433** %
Branching out	0.822	0.650	2	14.066 %
Multi-disciplinarity	0.898	0.839	2	21.725 %
Flexibility	0.923	0.868	2	47.433 %
Mentor influence	**0.818**	**0.903**	**6**	**67.069** %
Risk propensity (funding)	**0.637**	**0.705**	**6**	**40.203** %
Risk seeking	0.817	0.759	3	40.203 %
Risk aversion	0.808	0.743	3	27.175 %
Risk propensity (field)	**0.770**	**0.859**	**6**	**59.688** %
Risk seeking	0.782	0.775	3	13.673 %
Risk aversion	0.819	0.844	3	59.688 %

Note Values in bold indicate single-factor parameters

Model specification

Using the holdout sample, the first implemented model-specification strategy was to replicate the factorial structure previously identified by EFA. An initial estimation of the model with this structure was conducted to identify specification problems. However, an inadmissible solution was obtained due to an estimation problem affecting two factors: the variance in the disturbance terms in the Risk Propensity (Field) factor and the Risk Propensity (Funding) factor was negative. This situation, known as a Heywood case, is typically caused by either model misspecification or a small sample size (Kolenikov and Bollen 2012). As our sample size was adequate for the analysis conducted, the structure extracted during the EFA was probably not fully confirmed at the CFA stage. This possibility has been acknowledged in the literature (Marôco 2010). To mitigate the estimation problem, EFA was conducted again using all 12 items belonging to both of the Risk Propensity factors. A three-factor solution rather than the expected four-factor solution was extracted, because three items ('If a research area has little available funding, then I will not consider it'; 'I am afraid of engaging in research areas with no funding'; and 'I only research topics for which research funding is available') loaded into two separate factors. Therefore, these items were eliminated from further analysis. In addition, the three remaining Risk Propensity factors were separated to give three independent factors without a second-order structure. They were renamed to more accurately convey their new content, as follows. Risk Seeking as a low-level construct of Risk Propensity (Field) was renamed Discovery (e.g. 'I find "cutting-edge" scientific areas more appealing than well-established ones'). Risk Aversion as a low-level construct of Risk Propensity (Field) was renamed Conservative (e.g. 'I prefer "safe" or "stable" fields of study'). Finally, Risk Seeking as a low-level construct of Risk Propensity (Funding) was renamed Tolerance for Low Funding (e.g. 'Highly limited funding does not constrain my choice of field'). This re-specification of the model corrected the observed problems.

As a follow-up strategy, we scanned the factorial items and the low-level constructs to identify poor loadings, and conducted an initial analysis of validity and reliability. We identified a potential problem with the low-level construct of Flexibility, which had a low loading into the second-order construct of Divergence ($\lambda = 0.43$). As a result, the validity of the Divergence factor was low, with an average variance extracted (AVE) of 0.492, which was below the threshold of 0.5 suggested by Hair et al. (2007).[6] Further analysis of the content of the corresponding items revealed that they both contained the expression 'jack of all trades' ('In terms of research-field preferences, I like to think of myself as a "jack of all trades"' and 'I am a "jack of all trades" when it comes to research preferences'). These items probably loaded into the same factor due to their shared use of this expression, not because they represented the concept of Divergence. This would explain why they loaded strongly into a low-level construct that in turn loaded poorly into its second-order construct. As a result, these two items were eliminated from further analysis to ensure the validity of the Divergence construct, which improved significantly after this change. All of the items had factorial loadings above the 0.50 threshold, with two exceptions: one item in the Conservative factor, which had a loading of 0.46 ('I find emerging fields of science less preferable than well-established ones') and one item in the Mastery factor, a low-order construct of the Scientific Ambition factor, with a loading of 0.40 ('I have mastered a single scientific area'). These items were excluded, because loadings under 0.5 indicate poor construct validity (Marôco 2010).

The third and final strategy was to specify the covariance between selected disturbance terms based on the MI criteria, as previously described in the section on methodology. A single pass was conducted at a threshold of 11, which corresponds to a type I error probability of 0.001, and only the disturbance terms for items in the same factor were considered. Based on information theory fit indexes, the re-specified model had a much better fit (AIC = 1179.351; BCC = 1198.360) than the original model (AIC = 2254.491; BCC = 2288.065). After implementing these three strategies, the final model was analysed. The findings are reported below.

CFA

The final model was estimated using FIML. The overall model and all of the individual trajectories were found to be significant ($P < 0.001$). Evaluation of the various fit indicators with commonly used thresholds (Barrett 2007; Hair et al. 2007; Hooper et al. 2008; Marôco 2010) revealed that the model had a very good fit [$X^2/df = 1.710$; CFI = 0.961; PCFI = 0.791; RMSEA = 0.035; $P(\text{RMSEA} \leq 0.05) < 0.001$]. Table 3 provides the factorial loadings for all of the items, and the full model is represented in Fig. 1.

Validity, reliability and sensitivity

Validity is commonly assessed in three dimensions: factorial validity, convergent validity and discriminant validity (Hair et al. 2007; Marôco 2010). *Factorial validity* is confirmed

[6] This indicator is described in detail in a later section of the article.

Table 3 Factorial loadings for the MDRAI

Code	Item	Loading
A1	I aim to one day be one of the most respected experts in my field	0.898
A2	Being a highly regarded expert is one of my career goals	0.817
A3	I aim to be recognized by my peers	0.784
A4	Standing out from the rest of my peers is one of my goals	0.706
A5	I feel the need to constantly publish new and interesting papers	0.808
A6	I am constantly striving to publish new papers	0.858
C1	My expertise is focused on a single scientific area	0.683
C2	I believe that specialization in one area is preferable to diversification	0.666
C3	Shifting towards another field of science is not a part of my plans	0.517
C4	Studying subjects outside of my main field of work is pointless	0.618
C5	I have invested far too much in my current field to consider branching out into another	0.628
DI1	I find "cutting-edge" scientific areas more appealing than well-established ones	0.781
DI2	I would rather conduct revolutionary research with little chance of success than replicate research with a high chance of success	0.523
DI3	I prefer "cutting-edge" research to "safe" research, even when the odds of success are much lower	0.839
CN1	I prefer "safe" or "stable" fields of study	0.872
CN2	I prefer fields of study that are considered "safe" or "stable	0.799
TL1	Limited funding does not constrain my choice of field	0.846
TL2	Highly limited funding does not constrain my choice of field	0.822
TL3	The availability of research funding for a certain topic does not influence me doing research on that topic	0.676
CO1	I enjoy collaborating with other authors in my scientific articles	0.870
CO2	My scientific articles are enhanced by collaboration with other authors	0.831
CO3	I see myself as a team player when it comes to research collaboration	0.762
CO4	I often seek peers with whom I can collaborate on scientific articles	0.731
CO5	My peers often seek my collaboration in their scientific articles	0.855
CO6	I am often invited to do collaborative work with my peers	0.856
M1	My Ph.D. mentor's opinion carries much weight in my research choices	0.827
M2	A part of my work is largely due to my Ph.D. mentor	0.651
M3	My research choices are highly influenced by my Ph.D. mentor's opinion	0.865
M4	My Ph.D. mentor is responsible for a large part of my work	0.771
M5	My Ph.D. mentor still often works alongside me	0.742
M6	My Ph.D. mentor largely determines my venues of research	0.820
D1	I look forward to diversifying into other areas	0.816
D2	I would be interested in pursuing research in other fields	0.742
D3	I enjoy multi-disciplinary research more than single-discipline research	0.864
D4	For me, multi-disciplinary research is more interesting than single-discipline research	0.921

when all of the individual items have standardised loadings above 0.50 (Marôco 2010). As previously described, items with loadings below this threshold were removed, so the model had full factorial validity. *Convergent validity* is confirmed when the manifest items for a latent factor load heavily into that factor. The average variance extracted (AVE) of a given

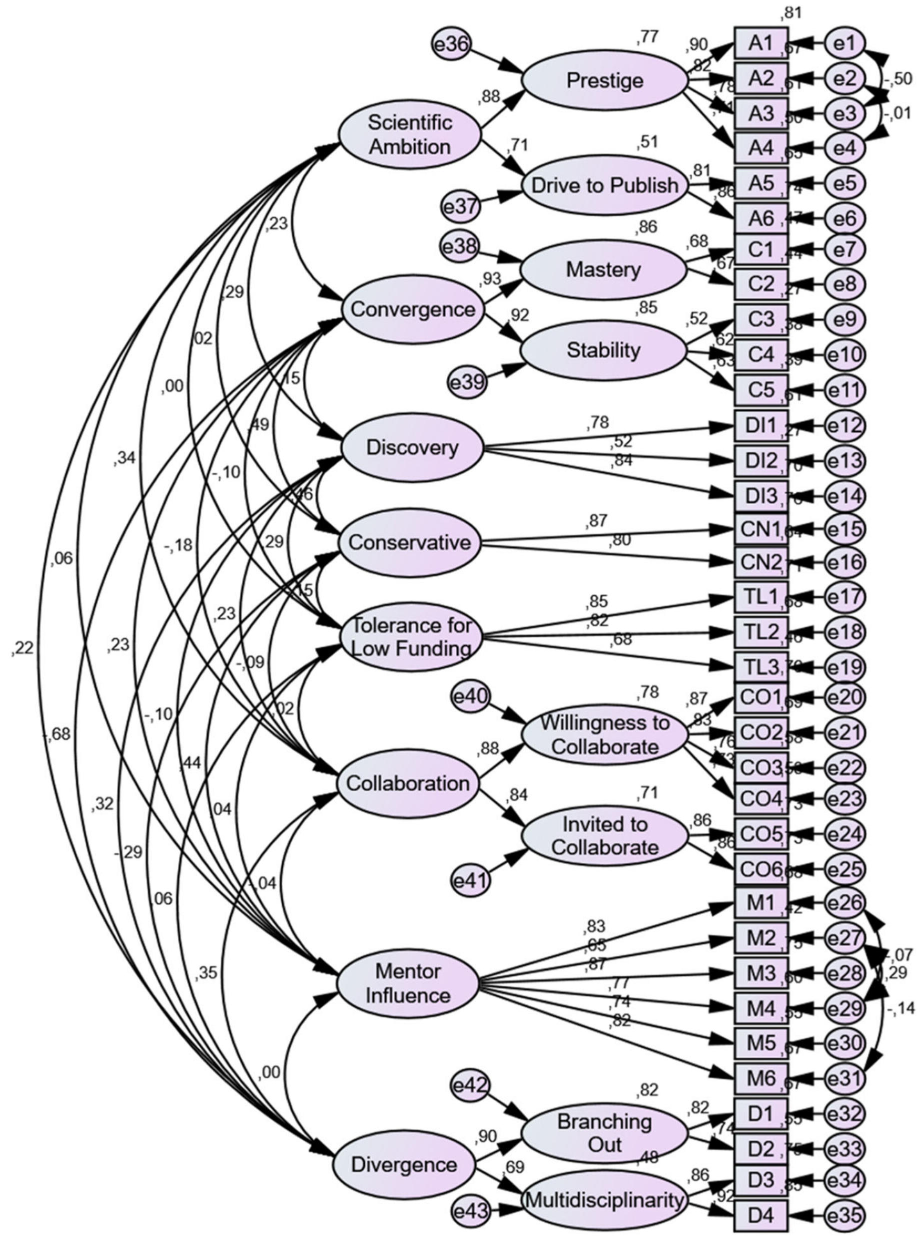

Fig. 1 Measurement model for the MDRAI with standardized regression weights (loadings). *Note ellipses* indicate latent variables, and *squares* indicate manifest variables. *Disturbance terms* are indicated by the latent variables labeled "e."

factor has been proposed as a useful index of convergent validity (Fornell and Larcker 1981). AVE is calculated as follows:

$$\widehat{\mathrm{AVE}}_J = \frac{\sum_{i=1}^{k} \lambda_{ij}^2}{\sum_{i=1}^{k} \lambda_{ij}^2 + \sum_{i=1}^{k} \varepsilon_{ij}}$$

where λ_{ij}^2 are the squared standardised factorial loadings for each item and ε_{ij} is the disturbance terms for those items.

Convergent validity is confirmed when AVE is higher than a 0.5 threshold (Hair et al. 2007). In our model, all of the factors were above this threshold, confirming the convergent validity of the MDRAI. Finally, *discriminant validity* describes the extent to which the items for a given factor are correlated with those for other factors. Discriminant validity can be confirmed by determining whether the AVE for factors i and j is equal to or greater than the squared correlation between the two factors (Fornell and Larcker 1981). Additionally, the AVE must be equal to or greater than both the maximum shared variance (MSV) and the average shared variance (ASV). Again, all of the factors in our model exceeded these thresholds, confirming that the model exhibited adequate discriminant validity.

We also evaluated *reliability*, defined as measurement consistency and replicability (Marôco 2010), by calculating the commonly used composite reliability (CR) indicator (Fornell and Larcker 1981). For a factor j with k items, CR is obtained as follows:

$$\widehat{\mathrm{CR}}_J = \frac{\left(\sum_{i=1}^{k} \lambda_{ij}\right)^2}{\left(\sum_{i=1}^{k} \lambda_{ij}\right)^2 + \sum_{i=1}^{k} \varepsilon_{ij}}$$

where λ_{ij} are the standardised factorial loadings for each item and ε_{ij} is the disturbance terms for those items.

A CR value above 0.7 is considered to confirm the reliability of a measure (Hair et al. 2007). As all of our factors received CR values above this threshold, we concluded that the MDRAI is a reliable instrument overall. Table 4 summarises the results of our validity and reliability assessment.

Finally, sensitivity is defined as the ability of an instrument to differentiate between individuals. Sensitivity is confirmed if the items have a normal distribution (Marôco 2010). The skewness and kurtosis of each of the items were analysed to identify any deviation

Table 4 Validity and reliability

Factor	Composite reliability	Average variance extracted	Maximum shared variance	Average shared variance
Scientific ambition	0.778	0.639	0.112	0.043
Convergence	0.922	0.855	0.465	0.126
Discovery	0.765	0.529	0.215	0.082
Conservative	0.823	0.699	0.241	0.109
Tolerance for low funding	0.827	0.616	0.082	0.017
Collaboration	0.855	0.747	0.125	0.047
Mentor influence	0.904	0.612	0.189	0.037
Divergence	0.784	0.649	0.465	0.119

from normality. The items were considered to exhibit an acceptably normal distribution if their skewness and kurtosis were each lower than an absolute value of 3 (Kline 2011). Using these criteria, no issues with normality were detected, indicating that the MDRAI is a sensitive instrument. Table 5 presents a summary of the descriptive statistics (mean and standard deviation) for each of the higher-order latent factors identified in the present study. For simplicity of presentation, the totals for the latent variables were computed using the mean for their respective manifest variables.

Discussion

In this section, some issues regarding the interpretation and scoring of the MDRAI are discussed.

Scientific Ambition was found to be a key variable in researchers' research agenda setting, consistent with the literature. Gaining recognition for one's research from academic peers and thereby moving up the scientific community hierarchy are important incentives for engaging in research (Bourdieu 1999; Latour and Woolgar 2013; Merton 1968), and thus have a considerable influence on research agendas. Researchers with high scores for this factor can be said to be research-community driven, in that they aim to become prominent in their respective fields (with the corresponding career benefits). This factor is subdivided into Prestige, the desire to acquire recognition per se, and Drive to Publish, the desire to produce scientific articles (which may be related to the accumulative-advantage hypothesis; see Allison et al. 1982 and/or the current 'publish or perish' paradigm; see Jung 2014).

The second factor, Convergence, relates to the intention to specialise in a single field of knowledge, which is a traditional professional strategy (Leahey 2007) with a significant influence on research agenda setting. A researcher scoring high for this factor is likely to create a research agenda characterised by much time and effort devoted to a single field of knowledge. This factor has two dimensions: Mastery, representing the goal of becoming an expert in a specific topic; and Stability, which represents the time investment made in a topic. The results for both dimensions were aligned with previous findings (Bourdieu 1999). The third factor, Divergence, reflects the desire to branch out into other fields of knowledge, a useful approach to the complex problems of modern science (Horlings and Gurney 2013). Researchers with high scores for this factor are likely to create research agendas with particular emphasis on establishing themselves in (or pursuing research interests in) many fields of knowledge (or researching inherently multi-disciplinary topics that encompass or relate to many fields of knowledge). Divergence is subdivided into Branching Out, the desire to expand one's research work to address other (potentially many) research topics, and Multi-disciplinarity, which indicates a preference for multi-disciplinary research ventures.

The next factor, Discovery, indicates a propensity for risky fields of knowledge. Researchers with high scores for this factor usually create research agendas in emerging and largely unexplored fields with greater risk—and greater potential reward—than more established fields of knowledge. Conservative represents the opposite: a preference for setting research agendas in established and thus safer fields in which outcomes are more predictable. The Tolerance for Low Funding factor represents the extent to which the availability of funding affects a researcher's choice of research topic. Researchers with high scores for this factor do not place particular emphasis on funding when setting

Table 5 Descriptive statistics

	Gender				Age								Degree					
	Female		Male		≤ 40		41–50		51–60		>60		Bachelor		MSc. MBA		Ph.D.	
	M	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
Ambition	4.92	1.04	5.00	1.06	5.20	0.98	4.94	1.09	4.89	1.05	4.85	1.04	5.02	0.90	4.68	1.07	5.05	1.05
Convergence	3.51	0.87	3.45	0.98	3.52	0.85	3.45	0.97	3.50	0.92	3.46	0.95	3.42	0.92	3.17	0.76	3.50	0.94
Divergence	4.99	0.93	4.98	1.06	5.05	0.93	5.05	1.02	5.01	0.98	4.81	1.02	5.05	1.06	5.10	0.91	4.97	0.97
Discovery	4.37	1.07	4.63	1.11	4.46	1.07	4.47	1.16	4.51	1.12	4.54	1.01	4.59	1.08	4.63	1.11	4.46	1.12
Conservative	3.04	1.03	2.94	1.04	3.24	0.96	3.01	1.09	2.94	1.03	2.82	1.00	2.98	1.07	2.75	0.98	3.00	1.07
Tolerance for low funding	4.43	1.28	4.72	1.26	4.49	1.20	4.49	1.36	4.58	1.30	4.73	1.20	4.64	1.19	4.60	1.20	4.62	1.32
Mentor influence	2.65	1.27	2.63	1.29	3.19	1.27	2.59	1.27	2.48	1.24	2.35	1.19	2.74	1.29	2.53	1.47	2.62	1.26
Collaboration	5.43	0.88	5.20	1.02	5.21	0.89	5.39	1.00	5.37	0.91	5.27	1.00	5.22	0.96	5.26	0.82	5.39	0.96

research agendas, whereas researchers with low scores create research agendas based largely on the availability of funding.

The next factor, Collaboration, represents a researcher's engagement (as reflected in his or her research agenda) in collaborative research endeavours, a critical strategy in science (Katz and Martin 1997; Uddin et al. 2013). A researcher with a high score for this factor is both willing and able to collaborate to produce research, as reflected in the lower-order factors Willingness to Collaborate (which measures a researcher's intrinsic inclination to set up research agendas in collaboration with others) and Invited to Collaborate (which indicates the frequency with which the researcher is actively invited to partake in research ventures initiated by others, as reflected in his or her research-agenda setting).

Finally, Mentor Influence indicates the extent to which a researcher's research agenda setting is influenced by his or her Ph.D. mentor. Individuals with high scores for Mentor Influence are likely to perform research alongside their Ph.D. mentors (perhaps even long after completing their doctoral work), and their research agendas and corresponding work are significantly shaped by this relationship (Pinheiro et al. 2014). Researchers with low scores for this factor are likely to set research agendas without considering the opinions of their Ph.D. mentors, perhaps because the mentor-researcher relationship has weakened over the years.

Concerning scoring, there is a vast range of computational methods (DiStefano et al. 2009) to calculate composite scores, but it is important to avoid computing scores by simple summation alone. As the factorial dimensions contain different numbers of items (despite our efforts to balance item numbers in the preliminary tests), simple summation would yield composite scores with a different range of values for each dimension, requiring the scores to be standardised further to enable unbiased comparison. The simplest way of calculating composite scores is to compute equally weighted averages for the items in each dimension. In our case, this results in standardised continuous scores ranging between 1 and 7, which were considered adequate. In cases in which imputation is impossible or undesirable, averages have the added benefit of mitigating score deflation resulting from missing values. Alternatively, weighted averages can be used to provide additional robustness. The factor loadings presented in Table 3 can be used as weights. Finally, in future research based on the MDRAI, the use of lower-order factors should be considered optional, depending on the purposes of the research undertaken.

Conclusion

This article describes the first instrument capable of evaluating the endogenous aspects of researchers' research agenda setting. The Multi-Dimensional Research Agendas Inventory enables the examination of a broad range of factors critical to researchers' decision making, and has robust measuring properties in terms of validity, reliability and sensitivity. In addition, the model underlying the instrument has a very good fit. The development of this instrument has important implications for both researchers and policy makers. It will add value to studies of scientometric research, higher-education research and scientific policy research by providing a tool for investigating researchers' individual agendas. This has the potential to open up new directions for research. The instrument will also be of interest to policy makers, especially funding managers and university managers, as it offers a new method of prediction and evaluation. For convenience, the full instrument in its final form is provided in an appendix. It is recommended that the items be randomised prior to use.

Nevertheless, it is important to highlight some limitations of the instrument developed in our study. First and foremost, the instrument is perception-based, and thus incurs all of the risk inherent in subjective measures. Therefore, it is necessary to consider and prepare for the possibility of individual bias in the responses before implementing the survey (in part by obtaining a sample large enough to mitigate such bias). Second, the validation exercise was carried out in a single field (higher education), and should thus be tested in other fields of knowledge. It is expected that the current instrument is more applicable to individual researchers in the social sciences and humanities than in the pure sciences. This is due to the greater autonomy that individual researchers in these fields of knowledge have in setting the research focuses that guide their research practice. (This is particularly evident in studies of research topics contributed by researchers from different disciplinary backgrounds within the social sciences; see Morley 2003.) The authors plan to carry out additional validation exercises in the near future; the aim is to eventually identify the dimensions that make the instrument more applicable to a wider range of fields of knowledge. As this is the first instrument of its kind, some important dimensions of research agenda setting may not have been considered. We hope that as the MDRAI is used and discussed by members of the scientific community, further dimensions will be identified and included in future revised versions of the instrument. An example that arose from the review process is that a dimension relating to items inquiring about the role and influence of collective research agenda consensus by research communities on individual research agenda setting is required to broaden the applicability of the instrument to all fields of knowledge. It was also recognised that the instrument should be revised to include additional questions measuring the factors that have only two items in the current version. Although the use of two items per factor is not entirely unheard of (see, for example, Rammstedt and John 2007), it may lead to problems if the respondent skips one or both of the questions.

Acknowledgments This study was supported by doctoral Grant PD/BD/113999/2015 from the Fundação para a Ciência e Tecnologia (FCT), co-funded by the European Social Fund (ESF) and the Portuguese Ministry of Science and Education. The study was also supported by the Seed Funding Programme for Basic Research of The University of Hong Kong.

Appendix: Multi-dimensional research agendas inventory (MDRAI)

You will be asked a series of questions regarding your motivations and goals as an academic. To respond to this questionnaire, read each statement carefully and decide how much do you agree with each of them. For each statement, check one of the 7 boxes next to the corresponding item. If you don't know or a particular sentence does not apply to you, check the N/A box.

There are no right or wrong answers. Please read each statement and check the box which best applies to you.

How much do you agree with the following statements?

		Completely disagree	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Completely agree	N/A
A1	I aim to one day be one of the most respected experts in my field.								
A2	Being a highly regarded expert is one of my career goals.								
A3	I aim to be recognized by my peers.								
A4	Standing out from the rest of my peers is one of my goals.								
A5	I feel the need to constantly publish new and interesting papers.								
A6	I am constantly striving to publish new papers.								
C1	My expertise is focused on a single scientific area.								
C2	I believe that specialization in one area is preferable to diversification.								
C3	Shifting towards another field of science is not a part of my plans.								
C4	Studying subjects outside of my main field of work is pointless.								
C5	I have invested far too much in my current field to consider branching out into another.								
DI1	I find "cutting-edge" scientific areas more appealing than well-established ones.								
DI2	I would rather conduct revolutionary research with little chance of success than replicate research with a high chance of success.								
DI3	I prefer "cutting-edge" research to "safe" research, even when the odds of success are much lower.								
CN1	I prefer "safe" or "stable" fields of study.								
CN2	I prefer fields of study that are considered "safe" or "stable."								
TL1	Limited funding does not constrain my choice of field.								
TL2	Highly limited funding does not constrain my choice of field.								

		Completely disagree	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree	Completely agree	N/A
TL3	The availability of research funding for a certain topic does not influence me doing research on that topic.								
CO1	I enjoy collaborating with other authors in my scientific articles.								
CO2	My scientific articles are enhanced by collaboration with other authors.								
CO3	I see myself as a team player when it comes to research collaboration.								
CO4	I often seek peers with whom I can collaborate on scientific articles.								
CO5	My peers often seek my collaboration in their scientific articles.								
CO6	I am often invited to do collaborative work with my peers.								
M1	My Ph.D. mentor's opinion carries much weight in my research choices.								
M2	A part of my work is largely due to my Ph.D. mentor.								
M3	My research choices are highly influenced by my Ph.D. mentor's opinion.								
M4	My Ph.D. mentor is responsible for a large part of my work.								
M5	My Ph.D. mentor still often works alongside me.								
M6	My Ph.D. mentor largely determines my venues of research.								
D1	I look forward to diversifying into other areas.								
D2	I would be interested in pursuing research in other fields.								
D3	I enjoy multi-disciplinary research more than single-discipline research.								
D4	For me, multi-disciplinary research is more interesting than single-discipline research.								

References

Allison, P. D., Long, J. S., & Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review, 47*, 615–625.

Anderson, D., Burnham, K., & White, G. (1998). Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture–recapture studies. *Journal of Applied Statistics, 25*(2), 263–282.

Andrews, R. N., & Johnson, E. (2016). Energy use, behavioral change, and business organizations: Reviewing recent findings and proposing a future research agenda. *Energy Research and Social Science, 11*, 195–208.

Arbuckle, J. (2007). *AMOS 16.0 user's guide*. Chicago, IL: SPSS.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences, 42*(5), 815–824.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*(1), 1–26.

Becher, T. (1994). The significance of disciplinary differences. *Studies in Higher Education, 19*(2), 151–161.

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories: Intellectual inquiry and the culture of disciplines*. Buckingham: Open University Press/SRHE.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238.

Bentler, P. M. (2007). On tests and indices for evaluating structural models. *Personality and Individual Differences, 42*(5), 825–829.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika, 45*(3), 289–308.

Boisot, M. (2011). *Collisions and collaboration: The organization of learning in the ATLAS experiment at the LHC*. Oxford: Oxford University Press.

Bollen, K. A. (2014). *Structural equations with latent variables*. London: Wiley.

Bourdieu, P. (1999). The specificity of the scientific field. In M. Biagoli (Ed.), *The science studies reader* (pp. 31–50). New York, NY: Routledge.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Publications.

Cantwell, B. (2011). Transnational mobility and international academic employment: Gatekeeping in an academic competition arena. *Minerva, 49*(4), 425–445.

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago, IL: University of Chicago Press.

Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science, 35*(5), 703–722.

DiStefano, C., Zhu, M., & Mindrila, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation, 14*(20), 1–11.

Ebadi, A., & Schiffauerova, A. (2015). How to receive more funding for your research? Get connected to the right people! *PLoS One, 10*(7), e0133061.

Ebrahimy, S., & Osareh, F. (2014). Design, validation, and reliability determination a citing conformity instrument at three levels: Normative, informational, and identification. *Scientometrics, 99*(2), 581–597.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling, 8*(3), 430–457.

Ertmer, P. A., & Glazewski, K. D. (2014). Developing a research agenda: Contributing new knowledge via intent and focus. *Journal of Computing in Higher Education, 26*(1), 54–68.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2007). *Multivariate data analysis*. Ahmedabad: Bookman.

Harris, M. (2001). Centering in on professional choices. *College Composition and Communication, 52*(3), 429–440.

Hillson, D., & Murray-Webster, R. (2007). *Understanding and managing risk attitude*. Aldershot, Hants, UK: Gower Publishing Ltd.

Hoffman, D. M., Blasi, B., Ćulum, B., Dragšić, Ž., Ewen, A., Horta, H., et al. (2014). The methodological illumination of a blind spot: Information and communication technology and international research team dynamics in a higher education research program. *Higher Education, 67*(4), 473–495.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6*(1), 53–60.

Horlings, E., & Gurney, T. (2013). Search strategies along the academic lifecycle. *Scientometrics, 94*(3), 1137–1160.

Horta, H., & Jung, J. (2014). Higher education research in Asia: An archipelago, two continents or merely atomization? *Higher Education, 68*(1), 117–134.

Horta, H., & Santos, J. M. (2016). The impact of publishing during the Ph.D. on career research publications, visibility, and collaborations. *Research in Higher Education, 57*(1), 28–50.

Horta, H., & Yonezawa, A. (2013). Going places: Exploring the impact of intra-sectoral mobility on research productivity and communication behaviors in Japanese academia. *Asia Pacific Education Review, 14*(4), 537–547.

Johnson, B., & Stevens, J. J. (2001). Exploratory and confirmatory factor analysis of the School Level Environment Questionnaire (SLEQ). *Learning Environments Research, 4*(3), 325–344.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications*. Chicago, IL: SPSS.

Jung, J. (2014). Research productivity by career stage among Korean academics. *Tertiary Education and Management, 20*(2), 85–105.

Jung, J., & Horta, H. (2013). Higher education research in Asia: A publication and co-publication analysis. *Higher Education Quarterly, 67*(4), 398–419.

Jung, J., & Horta, H. (2015). The contribution of East Asian countries to internationally published Asian higher education research: The role of system development and internationalization. *Higher Education Policy, 28*(4), 419–439.

Kassam, A., Papish, A., Modgill, G., & Patten, S. (2012). The development and psychometric properties of a new scale to measure mental illness related stigma by health care providers: The Opening Minds Scale for Health Care Providers (OMS-HC). *BMC Psychiatry, 12*(1), 1.

Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy, 26*(1), 1–18.

Klavans, R., Boyack, K. W., Sorensen, A. A., & Chen, C. (2013). Towards the development of an indicator of conformity. In 14th international society of scientometrics and informetrics conference. ISSI 2013.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances: Is a Heywood case a symptom of misspecification? *Sociological Methods and Research, 41*(1), 124–167.

Kosmützky, A., & Krücken, G. (2014). Growth or steady state? A bibliometric focus on international comparative higher education research. *Higher Education, 67*(4), 457–472.

Latour, B., & Woolgar, S. (2013). *Laboratory life: The construction of scientific facts*. Princeton, NJ: Princeton University Press.

Leahey, E. (2007). Not by productivity alone: How visibility and specialization contribute to academic earnings. *American Sociological Review, 72*(4), 533–561.

Levitt, D. G. (2010). Careers of an elite cohort of US basic life science postdoctoral fellows and the influence of their mentor's citation record. *BMC Medical Education, 10*(1), 80.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84.

Mamun, S. A. K., & Rahman, M. M. (2015). Is there any feedback effect between academic research publication and research collaboration? Evidence from an Australian university. *Scientometrics, 105*(3), 2179–2196.

Maroco, J. (2003). *Análise estatística: com utilização do SPSS*. Lisbon, Portugal: Edições Sílabo.

Maroco, J. (2007). Consistency and efficiency of ordinary least squares, maximum likelihood, and three type II linear regression models. *Methodology, 3*(2), 81–88.

Marôco, J. (2010). *Análise de equações estruturais: Fundamentos teóricos, software & aplicações*. Pêro Pinheiro: Report Number.

Martimianakis, M. A., & Muzzin, L. (2015). Discourses of interdisciplinarity and the shifting topography of academic work: Generational perspectives on facilitating and resisting neoliberalism. *Studies in Higher Education, 40*(8), 1454–1470.

Merton, R. K. (1968). The Matthew effect in science. *Science, 159*(3810), 56–63.

Morley, L. (2003). *Quality and power in higher education*. Berkshire: Society for Research into Higher Education, Open University.

Pinheiro, D., Melkers, J., & Youtie, J. (2014). Learning to play the game: Student publishing as an indicator of future scholarly success. *Technological Forecasting and Social Change, 81*, 56–66.

Pump, B. (2011). Beyond metaphors: New research on agendas in the policy process. *Policy Studies Journal, 39*(s1), 1–12.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*(1), 203–212.
Rzhetsky, A., Foster, J. G., Foster, I. T., & Evans, J. A. (2015). Choosing experiments to accelerate collective discovery. *Proceedings of the National Academy of Sciences, 112*(47), 14569–14574.
Schut, M., van Paassen, A., Leeuwis, C., & Klerkx, L. (2014). Towards dynamic research configurations: A framework for reflection on the contribution of research to policy and innovation processes. *Science and Public Policy, 41*(2), 207–218.
Simon, H. A. (1992). *Economics, bounded rationality, and the cognitive revolution*. Aldershot: Elgar.
Slovic, P., Fischhoff, B., & Lichtenstein, S. (1982). Why study risk perception? *Risk Analysis, 2*(2), 83–93.
Spector, P. E., Van Katwyk, P. T., Brannick, M. T., & Chen, P. Y. (1997). When two factors don't reflect two constructs: How item characteristics can produce artifactual factors. *Journal of Management, 23*(5), 659–677.
Steiger, J. H., Shapiro, A., & Browne, M. W. (1985). On the multivariate asymptotic distribution of sequential Chi square statistics. *Psychometrika, 50*(3), 253–263.
Uddin, S., Hossain, L., & Rasmussen, K. (2013). Network effects on scientific collaborations. *PLoS One, 8*(2), e57546.
Verbree, M., Horlings, E., Groenewegen, P., Van der Weijden, I., & Van den Besselaar, P. (2015). Organizational factors influencing scholarly performance: A multivariate study of biomedical research groups. *Scientometrics, 102*(1), 25–49.
Whitley, R. (2000). *The intellectual and social organization of the sciences*. Oxford: Oxford University Press.