
Book Section:
Marsden, Emma Josephine (2020) Methodological transparency in applied linguistics and its consequences for the quality and scope of research. In: McKinley, Jim and Rose, Heath, (eds.) Routledge Handbook of Research Methods in Applied Linguistics. Routledge, pp. 15-28.

Part Five: Current and future considerations for methods applied linguistics research

METHODOLOGICAL TRANSPARENCY AND ITS CONSEQUENCES FOR THE QUALITY AND SCOPE OF RESEARCH

Emma Marsden
University of York (UK)

How to cite:

Marsden, E. J. (2020). Methodological transparency in applied linguistics and its consequences for the quality and scope of research. In J. McKinley, & H. Rose (Eds.), *Routledge Handbook of Research Methods in Applied Linguistics* (pp. 15-28). Routledge.

Methodological transparency and its consequences for the quality and scope of research

1.0 Introduction

Methodological transparency can involve all aspects of the research process, from initial design, through peer review, to dissemination of findings. It means making the research process fully transparent so that reviewers and readers can understand exactly what the researchers did to elicit, analyse and understand their data; that is, how they moved from their research aims to data to findings to interpretation. As a key component of the Open Science movement, transparent methods are being (and have the potential to be) adopted to different extents across the subdomains within the broad field of applied linguistics. In this chapter, I describe the different practices that are required to make our research methodologically transparent and what the driving forces are behind these practices. I examine the extent to which methodological transparency is already established in the field and sub-fields of applied linguistics. I illustrate some of the costs of limited transparency by drawing on sobering examples from research into second language learning and teaching. I then describe some of the key benefits of working towards increased methodological openness and clarity. The chapter goes on to highlight some of the key developments and initiatives that can help us to improve the transparency of our methods to improve the quantity, quality, scope and usefulness of applied linguistics research, whilst also acknowledging the challenges that this involves.

It is important to distinguish between at least two interpretations of 'transparency'. One - the 'soft version' - involves making our research process fully available for other researchers, including reviewers, editors, and readers in the academic community who have access to journal articles and books that are usually behind paywalls. A second – 'strong version' - involves making our research process fully open to everyone. That is, transparent to those who are *beyond* the paywalls of the academe, by adopting Open Science practices and providing materials, data, and publications that are available at no cost at the point of access. In this chapter, I will attempt to identify which of these interpretations I refer to at different points, but the two are closely intertwined and inevitably merge into one another. (For an account of the broader notion of 'Open Science', please see Marsden, 2019).

In terms of its scope – its relevance to different subdomains and methods within applied linguistics – the need for methodological transparency can apply to almost all of them. However, the current chapter largely relates to research into multilingual language learning and education (broadly defined), largely because those areas have been, to date, the focus of meta-science into methodological transparency. This focus is also partly due to the fact that concerns about, and changes in, practices have tended to be driven by researchers working in more quantitative or hypothesis-driven research, often a characteristic of those subdomains. Nevertheless, most transparent practices can arguably be adopted and adapted within any subdomain, from ethnographic approaches through to laboratory experiments. However, different designs, methods, and epistemologies (philosophies about the nature of knowledge) do incur different sets of purposes and challenges for methodological transparency, issues that are touched upon at various points in this chapter.

2.0 What is methodological transparency?

The life cycle of methodological transparency runs in tandem with the life cycle of a research project. It affects decisions right from the start of the research process (conceptualisation and

design) through to its end (reporting the research findings). At the start of the research process, funding requests and planning must consider what resources and steps will be required to make it possible for others to understand and use materials, procedures, or data. For example, to make data available for full scrutiny, human participants must agree before data collection to have their data made available to more than just the researcher(s), and this must be approved via institutional ethics boards; data management plans must include details of where and how the data will be stored and the level of anonymity that is possible and necessary. To attain fully open research methods, participants are asked to agree to having their (anonymised) data made available in a public repository and held there indefinitely, as is increasingly being made a condition of funding by bodies across the world (see Marsden, Trofimovich & Ellis, 2019 and Trofimovich & Ellis, 2015 for some more information on these.)

At the next stage of the research process, during the design of materials (such as protocols, stimuli, schedules, and tests) researchers must bear in mind that *all* of their materials will be scrutinised by reviewers and readers. Knowing this beforehand may affect decisions about the design process itself, akin to a 'backwash' effect of transparency on the research process. For example, at this point, researchers might ask themselves, how well should I pilot this material (questionnaire, interview protocol, oral production test)? What do I already know about the reliability of this instrument, in terms of either the data that it elicits (internal reliability) or the extent to which it can be coded or scored reliably by more than one person? Do I need to get others to check language used (for accuracy and appropriacy)? If I use images or videos, do I need the permission of others (e.g. authors, publishers, participants) so I can share these materials later? Do the materials alone allow others to carry out a similar study with the confidence that they could compare their findings to ours, or do I need to provide an additional document laying out a protocol or explaining our decisions (e.g. specific words said to participants, the layout of a room, operation of equipment, the way in which access to a context or participants was obtained, or the order in which events must flow)? What data do I need to collect so that I will be able to describe my participants fully, with demographic information such as age, proficiency, language background, and context? In short, can others evaluate the relevance of my research to their own context and, where appropriate, could my methods be replicated? The knowledge that actual research materials will be made fully available to reviewers and other researchers (or, in the case of Open Science, to all), probably forces us to consider our design more carefully than if our materials were only to be described or a small sample of them provided. Indeed, this backwash effect already has some empirical support in some disciplines (described briefly below).

Once we have designed our instruments and collected our data, another set of decisions that affect transparency face us, about data preparation and data description. We will need to report if and how we cleaned our data (e.g., the level of anonymisation, removal of outliers, normalisation techniques), how we coded or scored our data, and how we analysed it. For quantitative research, certain details of reporting are necessary and increasingly expected by journals, such as descriptive statistics (means, sample sizes before and after data cleaning, standard deviations, confidence intervals), effect sizes and their confidence intervals, and instrument and rater reliability information.

A closely related aspect of transparency involves making any further analysis process as visible as possible – how exactly were the raw data reduced to the format in which they are presented as 'results.' For quantitative research, this means describing how, for example, percentages were calculated; which specific measures or indices were used (e.g. which eye movement measurements, of the many available; which regions/samples/extracts of language; which measures of complexity, accuracy or fluency were adopted); which criteria

were used to select the analysis procedures; which assumptions underpinning certain statistical procedures were checked; which criteria were used to select factors for inclusion or interpretation (e.g. in exploratory factor analysis, structural equation modelling, latent growth curve analyses); which modelling procedure was adopted (e.g. for regression-based analyses). Increasingly, quantitative researchers can use open source software and provide their actual analysis code (see Larson-Hall & Mizumoto, this volume), allowing others to replicate the analysis exactly.

At this point, the reader might be thinking that all these issues simply constitute good practice: 'our research methods training and our peer review and editorial processes should take care of all this!'. The next section lays out how, in fact, the field of applied linguistics has not yet achieved transparency in many of the ways laid out above. The following section then illustrates how this situation affects our ability to understand, evaluate, and replicate research.

3.0 What is the state of play of methodological transparency in applied linguistics?

The extent of methodological transparency in our field, in its entirety, has not yet been systematically evaluated, but a burgeoning meta-science is systematically examining the field's methodological and reporting practices. As Byrnes (2013) notes: "it appears that at this point in the development of applied linguistics, [methodological issues] demand a kind of professional scrutiny that goes directly to the core of what we do and what we know and what we can tell our publics that we know—and not only how we do it" (p. 825). This meta-science includes a growing number of systematic syntheses of particular methodological practices (such as research design and data elicitation or analysis techniques). It has served to highlight a severe lack of methodological transparency. In this section, I bring together some of this research to provide a short narrative account of the extent and nature of methodological transparency in terms of materials, data and analysis. The picture about to be described is sobering, with few bright points. However, it is very important to note that there is no intention to criticise or 'blame' researchers, reviewers or editors: changing expectations and standards are entirely inevitable as our research aims, cultures, and capacities shift in concert with evolving societal views and technological innovation. The aim here is simply to describe the situation to date and indicate the direction of change.

3.1 Materials transparency to date. As noted above, authors can make their materials (stimuli, procedure, analysis protocols, code) and data (raw and/or coded or reduced at some level) available. Early steps in this direction were taken by individual researchers, in the absence of larger, more sustainable infrastructure and incentives: For example, vocabulary research materials have been available on Paul Meara's Lognostics site (http://www.lognostics.co.uk/) for many years, as have materials for research into language attrition on Monika Schmidt's Language Attrition website (https://languageattrition.org/). More examples regularly emerge, such as Atsushi Mizumoto's resources for analysis and natural language processing. As committed as these individuals are, repositories that are sustainable and community supported (such as IRIS www.iris-database.org or the Open Science Framework (OSF)) are now available and offer perhaps greater hope of sustainability, visibility and reach across broader domains of research. For these reasons, such repositories (rather than individual or institutional platforms) are endorsed by the *Centre for Open Science,* a large international philanthropic initiative established in 2011 to promote and facilitate open science practices across disciplines.

To facilitate methodological transparency in the domain of language learning, use and education, an initiative began in 2012 known as IRIS[1] - Instruments for Research in Second Languages (Marsden, Mackey, & Plonsky, 2016; Marsden, Thompson, & Plonsky, 2017). This repository now holds over 4,300 files of materials and analysis protocols, including 77 files of second language learning data. Among these materials are numerous examples of the methods covered by this Handbook. IRIS offers a discipline-specific, highly searchable platform, hosting only materials and data from peer reviewed publications (including PhD theses). This is in contrast to the OSF which also holds works-in-progress and non-peer reviewed work. Enthusiasm for and engagement with IRIS can be seen in the approximately 36 journals that encourage their accepted authors to upload their materials to IRIS. In 2017, this practice was endorsed in the *American Association for Applied Linguistics*' Publication Guidelines, and by the *British Association of Applied Linguistics*.

The IRIS resource offers great potential and has received very active support from some quarters. But what *proportion* of articles have materials that are actually available on IRIS? An ongoing study is investigating this by examining how many data collection instruments *could be* available from the ten journals that have the most materials on IRIS, ranging from 2013, just after IRIS became live, to the end of 2018 (Marsden, Thompson, & LaFlair, in preparation). They have found that, in total over those five years, approximately just 13.6% of the articles that used data collection instruments have made some materials available on IRIS. The trajectory increased in the first few years, but seems to have plateaued at about 15% annually. This low proportion is despite the fact that 36 journals in applied linguistics report they routinely invite their authors to upload their materials upon acceptance of a manuscript.

The IRIS repository also has two special collections of materials that have been reviewed in published methodological syntheses: 62 self-paced reading tests (Marsden, Thompson, & Plonsky, 2018) and 110 acceptability judgement tests (Plonsky, Marsden, Crowther, Gass, Spinner, in (2019). Although this may seem impressive, in fact these special collections were only gathered after an intensive effort from the IRIS team to contact all the authors (where possible), to seek their instruments. *Before* this effort, Marsden et al. (2018) found only 4% of self-paced reading studies had openly available materials and 77% had only a brief example of stimuli available in their articles (not the full instrument). Similarly sobering is that for judgement tests, the 110 materials that are now available still only represent just over one third of the total 385 JTs that were found by the authors for inclusion in their synthesis. Another indication of low levels of materials transparency was found by a synthesis of replication studies in which Marsden, Morgan-Short, Thompson & Abugaber (2018) found very low levels of availability of materials in the initial studies that had been claimed to be replicated: 17% of the 67 self-labelled replication studies claimed to replicate initial studies that had not provided any materials at all; 41% of the initial studies provided only partial examples in the article; 37% provided at least one full instrument that had been used, but not *all* of the instruments used to collect the data. Only three of the studies that had been replicated provided all of their materials.

3.2 Data and analysis transparency. Early adopters of data and coding transparency in applied linguistics include several corpus linguistics projects, ahead of the game, that made their data, often richly tagged, available to the community, such as: the French Language Learning Oral Corpus (e.g., Marsden, Myles, Rule, & Mitchell, 2003); the Spanish Language

[1] The IRIS project was initially funded by the *Economic and Social Research Council* UK (RES-062-23-2946) and has long-term funding from *The British Academy* (AN110002).

Learning Oral Corpus (Mitchell, Dominguez, Arche, Myles, & Marsden, 2008); the Corpus of Contemporary American English (Davies 2008); MacWhinney's TalkBank; the Multimedia Adult ESL Learner Corpus (Reder et al. 2003). Ongoing open data projects continue in this vein, such as Bell, Collins & Marsden (2018). See Tracey-Ventura & Huensch (2018) for a discussion of the importance and usefulness of open L2 corpora.

Perhaps testament to a changing culture, in the two years in which IRIS has welcomed data (2016-2018), 77 datasets have been made available. Individual researchers seem to be increasingly making their data and materials openly available (e.g. van den Broek, Takashima, Segers, & Verhoeven, 2018), and some even explicitly inviting other researchers, in the title of the study, to use their materials and data to replicate their study (Arroyo & Yilmaz, 2018).

Other signs of transparent methods include a growing number of researchers producing and using open-source software, such as R data analysis scripts (see Larson-Hall & Mizumoto, this volume; Mizumoto & Plonsky, 2015; Norouzian, de Miranda & Plonsky, 2018) and calls for large-scale data collection available to all via online platforms (MacWhinney, 2017). However, neither of these practices are common across the sub-domains of applied linguistics as yet.

Overall, a somewhat bleak picture of methodological transparency emerges which has a range of negative consequences, to which we now turn.

4.0 Some negative consequences of poor methodological transparency

There are a multitude of negative consequences for not making our methods more transparent during peer-review and after publication. Specific examples of these consequences are discussed by Gudmestad & Edmonds (2018) and Marsden & Plonsky (2018). Here I draw on a few meta-scientific and methodological syntheses of applied linguistics research into second language learning to illustrate the importance of methodological transparency for research quality, rigour, reliability and validity.

4.1 Consequences of poor transparency on replicability[2], reproducibility[3], and syntheses. For more than a decade, the trustworthiness of research findings has been called into question in a number of disciplines, most recently in the psychological sciences by a series of large-scale multi-site replication studies that have found that the results of important studies could not be reproduced to within satisfactory levels (for a review see Marsden, Morgan-Short, Thompson, & Abugaber, 2018). The problem is thought to have gone undetected due to an historic lack of replication research (Munafò et al., 2017) and has been dubbed by many as constituting a "replication crisis". It has been a driving force behind huge methodological introspection and infrastructure creation in the psychological and clinical sciences.

Reflecting this trend, in the domain of second language learning research, Marsden et al. (2018) carried out a synthesis of replication studies and found that fewer than one in every 400 journal articles has been a self-labelled replication study, since the first such study was published. They argue that one cause of this paucity may be the poor availability of materials

[2] The term 'replicability' is used to refer to the extent to which a study's methods are such that they can be replicated i.e., run in the same or similar way again by others

[3] The term 'reproducibility' is used here to refer to the extent to which findings are reproduced reliably across different iterations of a study.

in the first place. Also, the lack of published replication research probably reduces the need to make methods accessible, as there is a low chance of anyone replicating it. Thus, low rates of published replication research probably both partially account for and lead to inadequate methodological transparency.

Not just the quantity, but also the *quality* of replication research (or indeed of any research that claims comparability to a previous study) is threatened by poor materials and data transparency. Poor availability of materials means that researchers need to create their own materials by extrapolating from the often-cursory descriptions and examples provided in published articles. The resulting (often unacknowledged) heterogeneity between studies adversely affects the inter-connectedness within research agendas. Indeed, Marsden et al. (2018) found a positive association between the extent of materials availability and the extent to which the results of replication studies supported the results of the initial studies. Also, poor materials availability means that collaboration with the author(s) of the initial study becomes necessary to access materials. This situation is not ideal if we aspire for objective, independent validation of previous findings. Illustrating concerns about such non-independent replication, Marsden et al. (2018) found a positive, statistically significant association between the likelihood of a replication study supporting the initial study's findings when there was some 'author overlap' between the two studies (90% of replication studies) compared to when there was no author overlap (59% of replication studies).

Also useful for improving the quality of research, more transparent materials would help to improve the field's ability to check the reliability of instruments when administered across different contexts and participant demographics (see Plonsky & Derrick, 2016 for a range of concerns about instrument and rater reliability).

The systematicity and scope of synthetic work in our field would also be greatly enriched by more transparent data. Non-availability of data reduces our capacity to compare results across studies (such as meta-analyses) and to re-analyse previously collected data. For example, only aapproximately one third of the requests made by Plonsky (2011) and Lee et al. (2014) resulted in authors providing missing data. Out of the 255 studies that were initially identified by Plonsky, Egbert & Laflair (2015) for inclusion in their re-analysis study, only 37 data sets (14.5 per cent) were received from authors and only one of the studies had reported the raw data in the initial publication. Plonsky et al. then had to eliminate 11 studies because the initial reported findings did not align with the data that had been sent or because the data sent were indecipherable/uninterpretable. This situation not only makes such activities difficult, decreases the power of re-analysis, and reduces our ability to detect trends over multiple studies, it is clearly a huge waste of resources. For discussions of similar challenges in linguistics, see Berez-Kroeker et al. (2017), and in psychology, see Fecher, Friesike, and Hebing (2015) and Lindsay (2017).

4.2 Non-transparency permitting Questionable Research Practices. The non-transparency of complete sets of materials and data that were involved in a study can result in yet more practices that are known to affect the trustworthiness of research findings. These are known as 'Questionable Research Practices' (National Science Foundation, 2015) and include: reporting only a subset of the data (e.g., those that might be more likely to get published); adding or removing participants or trying more analyses until a statistically significant finding emerges (known as *p*-hacking or data dredging); and claiming or implying that theorising occurred *before* a study (i.e., presenting confirmatory research, important for hypothesis testing) when in fact theorising emerged *after* data analysis (i.e., exploratory research, important for hypothesis generation). This latter practice, known as HARKing—hypothesizing after the results are known (Chambers, 2017; Kerr, 1998; Lindsay, 2015) –

increases the apparent plausibility of the research and makes it difficult to establish reliable means of testing theories and track the evidence base for and against theories.

4.3 Reporting practices cannot fully compensate for an absence of materials and data. Reports in journal articles and books can only supply a limited amount of information about materials, thus having full materials available may be the only way that readers can access enough information to evaluate a study's quality and relevance. Relying on reporting alone, important details about materials are missed due to an inevitable lack of standardisation across different journals and the organic nature of reporting standards caused by the sheer rate of methodological innovation. For example, Marsden, Thompson & Plonsky (2018) observed very spotty reporting of key features of self-paced reading tasks: sentence length was reported in only half the studies (30/64) and out of the 50 studies that used questions to check comprehension during reading, only half provided an example of the questions and only a small handful provided sufficient examples to enable an evaluation of the quality of those questions. Similarly, but spanning a wide variety of domains in L2 research, the reporting practices surrounding 925 data collection instruments found in 385 articles were reviewed by Derrick (2016) and indicated very poor transparency. For example, only 17% reported piloting of instruments and 28% reported reliability coefficients.

Omissions in reporting, without full materials and data, can affect our understanding of the validity and significance of research findings, as a lack of clarity about the nature of the data can lead to unintended, or unknown, data heterogeneity between studies. For example, Plonsky et al. (2019), in examining studies using judgement tests, observed low levels of reporting about whether grammatical (acceptable) or ungrammatical (unacceptable) sentences were analysed together or separately. They also found modality (oral vs written) and timing (timed vs untimed) of the tests were frequently not reported; yet, where these details *were* reported, Plonsky et al. demonstrated through meta-analysis that these methodological characteristics systematically and reliably influenced findings, with scores on oral tests being lower than written, and on timed tests lower than untimed. Similarly, Marsden et al. (2018) found that in self-paced reading studies, the number of items in different conditions was not always reported, yet the extent to which conditions are balanced is important as too many items per condition can affect performance (fatigue participants, desensitize them to ungrammaticality, or raise awareness about the manipulations), whereas too few items per condition can fail to provide a trustworthy result.

We have now seen the importance of methodological transparency given that methodological choices demonstrably affect the quality and usefulness of findings. I now highlight a number of initiatives that offer researchers increased opportunities to enhance replicability, to reduce concerns about the integrity of the research process, and to scrutinise the validity of data and analyses.

5.0 Ongoing and future developments for methodological transparency: Research infrastructure, policy, and practice.

Many studies, across a wide range of disciplines, have empirically demonstrated benefits of methodological transparency (McKiernan, 2016). For example, holding open *data* has been associated with better research and reporting quality (Wicherts, Bakker & Molenaar, 2011) and with increased citations (Piwowar & Vision, 2013); publishing open *reports* can increase citations (Wagner, 2010); and transparent *statistical procedures* have been positively associated with rates of journal citations (Al-Hoorie & Vitta, 2018).

Such benefits are now facilitated by several top down and grassroots initiatives. One such initiative is 'pre-registration', whereby the design, full materials, and (likely) analyses are reported before data collection begins. There are two main routes for this process: via a public site (a pre-registration) or a journal publication route (a Registered Report).

5.1 Pre-registration. One route to providing transparency before and after data collection is to pre-register a study on a public site, such as the Open Science Framework – a free platform provided by the *Centre for Open Science.* The plans are 'date-stamped' and 'locked' so that changes cannot be made, and open to allow public visibility to the intended research approach. Once the data are collected and the study submitted for publication, reviewers and editors then check whether the study adhered to the pre-registration. There are several advantages to pre-registering in an open forum. One is that it assures readers that the researchers have not indulged in Questionable Research Practices. Another benefit is that other researchers at different sites can join the endeavour, using the same protocol, materials and analyses, as done by Morgan-Short, Marsden, Heil, et al. (2018). Such open, multi-site collaboration reduces 'researcher degrees of freedom' (the variability in how different researchers operationalise the same study), increases sample size, and permits the investigation of contextual effects. It can also allow researchers with different theoretical views to work together in open 'adversarial collaboration' (e.g. Mellers, Hertwig, & Kahneman, 2001). In sum, pre-registration has the potential to address concerns with the *research* process. However, addressing these concerns plus additional ones relating to *peer-review and publication* processes requires another mechanism: Registered Reports.

5.2 Registered Reports: Methodological transparency during the peer-review process. This article type, first introduced in the journal *Cortex* (Chambers, 2013), has been adopted by about 70 journals as a permanent article type (https://cos.io/rr), and by one, *Language Learning*, in the field of applied linguistics (see Marsden, Morgan-Short, Trofimovich & Ellis, 2018). Authors first submit a manuscript prior to data collection, containing the justification, methods and proposed analyses. This 'stage 1 manuscript' is sent for review as per normal journal procedures to evaluate the proposed study's worth, design, and methods. Reviewers suggest amendments, leading to revisions and potentially further reviews. Once approved, the editor issues In Principle Acceptance (IPA). This gives a green light to the author(s) to collect and analyse data, safe in the knowledge that reviewers cannot now reject the manuscript on the grounds of its justification, design, or methods, nor can reviewers suggest changes to those elements. So long as the authors then adhere to the registered stage 1 manuscript, this IPA protects research from several potential problems: publication bias (as surprising, controversial, or non-statistically significant results have to be published); reviewer bias (as reviewers cannot reject findings that do not align with their theoretical or methodological persuasions); retrospective criticism of methods after the data collection has finished. Of most relevance to the current chapter is that methods are transparent - the reviewers and editors of Registered Reports see *all* materials that will be used to collect the data as well as the coding and analysis procedures. This addresses the problem that under standard publication routes, most reviewers seek further information about instrumentation, data cleaning, coding and analysis (see DeKeyser & Schoonen, 2007). Such requests slow the publication process and can sometimes block it once reviewers see materials when it is too late to change them.

In addition to this transparency-at-review, an increasing number of journals also ask authors to archive their stage one manuscript on an open repository (such as the OSF), with the release embargoed to suit the preferences of those involved. This additional, 'strong

version' of transparency provides the opportunity for anyone to scrutinise the extent to which the pre-registered protocol was adhered (Hardwicke & Ioannidis, 2018).

5.3 Peer Reviewers' Openness Initiative. This is another grass-roots movement that enhances methodological transparency (Morey et al., 2016), whereby researchers refuse to review manuscripts unless full materials and data are available for peer-review and made open after acceptance or the authors justify why this is not appropriate. This initiative has hardly impacted applied linguistics to date. Although somewhat extreme, it may serve in the best interests of both the author (as concerns about methodology are aired earlier in the review process) and science (as peer review under this model always includes an evaluation of full materials, not of author-picked examples).

5.4 Incentivisation by journal editors to promote methodological transparency. One initiative that has been adopted to some extent in our field are the *Transparency and Openness Promotion Guidelines* (Nosek et al., 2015), which encourage and recognize journals that facilitate and promote the various stages towards full methodological transparency. Another, related initiative is found in the Open Science Badges, which flag articles that have open materials, data, and/or pre-registration (see Figure 1).

Figure 1. Badges indicating open science practices in participating journals (created 2013). From Blohowiak, B. B., Cohoon, J., de-Wit, L., Eich, E., Farach, F. J., Hasselman, F., … DeHaven, A. C. (2018, November 14). Badges to Acknowledge Open Practices. Retrieved November 27, 2018, from osf.io/tvyxz.

Kidwell et al. (2016) and Giofrè Cumming, Fresc, Boedker, and Tressoldi (2017) provide quantitative evidence for the positive effect this initiative has on the sustainable availability of materials and data. In applied linguistics, this method of valuing materials and data transparency was adopted by *Language Learning* (Trofimovich and Ellis, 2015) and also currently by *Studies in Second Language Acquisition* and *The Modern Language Journal*. *Language Learning* also values pre-registration via the relevant open science badge and accepts Registered Reports (Marsden, Morgan-Short, Trofimovich & Ellis, 2018).

Methodological transparency also involves making the process of research more accessible to those *outside* academia. Although several initiatives review or summarise research for, say, education practitioners, very few make the *methods* of the research transparent, even though it is methodological information that is necessary for readers to gauge the relevance of the research (see Marsden & Kasprowicz, 2017, for discussion of these issues). Readers need to know about the design, participants and the measures used (e.g. how was 'learning' operationalized in this study?). This problem is being systematically addressed by a cross-journal, sustainable initiative in applied linguistics known as OASIS (Open Accessible Summaries in Language Studies) (Marsden, Trofimovich, & Ellis, 2019). Encouragingly, thirteen journals have pledged to ask authors to write non-technical summaries, and six of these are already routinely doing so.

In sum, there are now many initiatives and opportunities to serve methodological transparency: Digital platforms have eliminated any barrier caused by insufficient space in

journals; open infrastructure is available at no cost to researchers; top-down policies (from governments and granting agencies) increasingly enforce researchers to deposit the products of research (e.g., DARPA, National Science Foundation, UKRI, and see OECD Principles of Guidelines for Access to Research Data from Public Funding); and professional associations and journals promote methodological transparency (e.g., the American Psychological Association, the American Educational Research Association, Centre for Open Science). Yet, as we saw in section 4, progress towards methodologically transparency seems slow.

6.0 Challenges and concerns

A range of issues might prevent researchers from being more transparent about methods and data. Some relate to the limited influence of government pressure through funder regulation. For example, although the funder requirements alluded to above promote open *products* of research (data and reports), very few focus on the *process* of research through open archiving of materials. Perhaps more importantly, however, funder requirements by no means affect all applied linguistics research, as our research is not always funded by external agencies or is charity- (rather than government-) funded. Thus, full methodological transparency currently remains largely voluntary and is vulnerable to a range of complex barriers and concerns (Marsden, Thompson, & LaFlair, in preparation). As we do not yet have a very robust understanding of what, precisely, prevents improved methodological transparency, what is laid out here is somewhat speculative but gleaned from almost a decade of working with the research community on open science initiatives, including with students, early career researchers, established academics, journal editors and presidents of associations.

Some of the barriers are mundane, such as a simple lack of time to make one's methods and data transparent, or lack of physical access to previous materials and data. Others, however, involve an array of psychological, cultural, personal, social and philosophical issues.

Some concerns about making one's methods fully transparent relate to a series of fears. One is that of being 'scooped': that others will use your materials (whether behind a paywall or held openly) and pre-empt what you were planning to do. This fear can be allayed to some extent by embargoing the release of materials and data until after publication. However, even once published, researchers can still fear that others may scoop their ideas before they have exhausted their own use of their materials and data. Such concerns tend to conflate transparency with 'losing' intellectual property. Another, related fear is that others will find fault with one's materials, analysis or data. The personality characteristics that seem to lead researchers to succumb to or resist such fears are described in rich detail by Laine (2017). Also countering these concerns, proponents of open science argue that no study is complete where materials and data are not transparent and that intellectual property is not weakened by transparency but instead its quality is enhanced, as more information is available to help interpret, validate, or replicate the knowledge. In addition, Creative Commons Licensing means that downloading open materials signals a commitment to cite them appropriately and it continues to be incumbent upon the community to ensure that use or adaptation of transparent materials respects authorship.

A broader concern, for methodological rigour in general, is that the mere fact of being transparent might increase the chances of a particular instrument being used more widely (rather than its validity or reliability, for example). In addition to existing quality assurance mechanisms (dissertation panels and peer review), the most powerful way to counter any such 'transparency bias' is for the whole field to move together – once transparency is the norm, methodological choices cannot be influenced by transparency *per se.*

A practical concern worth airing relates to the perceived 'extra' time and resource costs involved in ensuring materials and data are documented such that they are interpretable by others. However, such costs are arguably inherent in our cumulative and collaborative research endeavour, rather than constituting an extra burden. Relatedly, the methodological review that happens before data collection for Registered Reports provides time savings later in the research timeline due to a less burdensome final review and publication process.

Finally, an important challenge for some aspects of methodological transparency is that certain mechanisms (such as pre-registration and Registered Reports) are likely more suited to particular epistemologies, where ontologies, coding, counting, and hypotheses are determined and can be documented, at least to some extent, prior to data collection (e.g., quantitative and experimental approaches). Note, however, that both pre-registration and Registered Reports *can* accommodate qualitative research designs, so long as there is some element of the methods that can be usefully pre-determined. Similarly, exploratory (unplanned) analyses and serendipitous findings can certainly be documented in Registered Reports, separated from the analyses that were pre-registered (see Marsden, Morgan-Short, Trofimovich & Ellis, 2018 for discussion). Of course, it is fully acknowledged that mechanisms such as these may not be appropriate for some approaches to research, such as Grounded Theory (see Hadley, this volume), for which the purposes and constraints of pre-registration are not relevant. Another challenge related to applicability to different subdomains of research is that qualitative data is often so enriched by contextual information that anonymisation (necessary for the data to be made openly available) remove the data's utility and/or that secondary analysis by others would not be rigorous due to the separation between the analyst and the data collection itself (a pivotal relationship in some qualitative approaches such as ethnography). However, work is ongoing that is attempting to address such challenges (Irwin & Winterton, 2012; Piñeiro & Rosenblatt, 2016).

The concerns and challenges raised above suggest there is a strong need for unified directives and incentives from professional associations, promotion systems, funders, and journals.

Conclusion

Benefits of methodological transparency include improving the accessibility, visibility, rigour, scrutiny, reproducibility, replicability and systematicity of research. Indeed, methodological transparency is increasingly regarded as an indicator of study quality. Individual researchers can take several steps to move toward more open practices including: seeking consent from participants to hold data on open repositories; making materials and data open and easily searchable on sustainable repositories; using pre-registration and Registered Reports that make the research process fully open to reviewers and, eventually, to all. Collective action means that individual researchers are not disadvantaged by some of the challenges outlined above, so it is important that Editors and professional associations facilitate and require these practices. Many research communities beyond applied linguistics, including some of its sister disciplines such as social and cognitive psychology, education, neuroscience, and linguistics, have been experiencing similar cultural changes. For example, nearly a decade ago, in 2010, the *Linguistic Society of America* and the *American Psychological Association* made policy statements prescribing good practice in data sharing. Researchers in applied linguistics are increasingly feeling pressures and incentives and now have a range of infrastructures in place. In some respects, the field is at the forefront of methodological transparency and there are high expectations that it will improve the quality, quantity and reach of our research, but a more concerted effort is required.

Perhaps most importantly, a set of philosophical, moral, social, and economic arguments make the case for a fully transparent research process. One is an 'economic efficiency' argument, that greater transparency could ease the cost of (re)creating materials and allow multiple uses of data. Another powerful argument is that what researchers do with tax-payers' investment should be fully and freely available for public and scientific use and scrutiny. Such arguments are, surely, further underpinned by many researchers' sense of epistemic responsibility: an individual and collective duty to share the knowledge that we gain, which includes our methodological know-how.

References

Al-Hoorie, A. H. & Vitta, J. (2018) The seven sins of L2 research: A review of 30 journals' statistical quality and their CiteScore, SJR, SNIP, JCR Impact Factors. *Language Teaching Research* https://doi.org/10.1177/1362168818767191

Arroyo, D. C. & Yilmaz, Y. (2018). An open for replication study: The role of feedback timing in synchronous computer-mediated communication. *Language Learning* 68.4, 1–31.

Bell, P., Collins, L., & Marsden, E. (2018). Building an oral and written learner corpus of a school programme: Methodological issues. In B LeBruyn & M. Paquot, Learner Corpus Research and Second Language Acquisition. Amsterdam: John Benjamins.

Berez-Kroeker, A., Gawne, L., Kung, S., Kelly, B., Heston, T., Holton, G., et al. (2017). Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics*, 56, 1–18. https://https://doi.org/10.1515/ling-2017-0032

Byrnes, H. (2013). Notes from the editor. *Modern Language Journal, 97*, 825-827

Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex*, *49*, 609-610.

DeKeyser, R., & Schoonen, R. (2007). Editors' announcement. *Language Learning,* 57, ix–x.

Derrick, D. J. (2016). Instrument reporting practices in second language research. TESOL Quarterly, 50, 132–153.

Fecher, B., Friesike, S., & Hebing, M. (2015). What drives academic data sharing? PLOS ONE, 10, e0118053. https://doi.org/10.1371/journal.pone.0118053

Giofrè, D., Cumming, G., Fresc, L., Boedker, I., & Tressoldi, P. (2017). The influence of journal submission guidelines on authors' reporting of statistics and use of open research practices. PLOS ONE, 12, e0175583. https://doi.org/10.1371/journal.pone.0175583

Hardwicke, T. E., & Ioannidis, J. (2018). Mapping the universe of Registered Reports. https://doi.org/10.31222/osf.io/fzpcy

Irwin, S. & Winterton, M. (2012) Qualitative secondary analysis and social explanation *Sociological Research Online,* 17 (2) 4 <http://www.socresonline.org.uk/17/2/4.html> 10.5153/sro.2626

Kidwell, M. C., Lazarevi´c, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S.,

Falkenberg, L. S., et al. (2016). Badges to acknowledge open practices: A simple, low-cost, effective method for increasing transparency. *PLOS Biology*, 14, e1002456. https://doi.org/10.1371/journal.pbio.1002456

Laine, H. (2017). Afraid of Scooping – Case study on researcher strategies against fear of scooping in the context of open science. *Data Science Journal*, 16, p.29 http://doi.org/10.5334/dsj-2017-029

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65 (Supp. 1), 127–159.

Leal, T. (2018) *Data analysis and sampling: Methodological issues concerning proficiency in SLA research.* In A. Gudmestad, & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition.* Philadelphia, PA: John Benjamins. pp. 63-68

MacWhinney, B. (2017). A shared platform for studying second language acquisition. *Language Learning*, 67, 254–275.

Marsden, E, (2019). Open Science and transparency in Applied Linguistics. In C. Chapelle (Ed.), *The Concise Encyclopedia of Applied Linguistics* Oxford: Wiley-Blackwell.

Marsden, E., Mackey, A., & Plonsky, L. (2016). The IRIS repository: Advancing research practice and methodology. In A. Mackey & E. Marsden (Eds.), Advancing methodology and practice: The IRIS repository of instruments for research into second languages (pp. 1–21). New York, NY: Routledge.

Marsden, E. J., & Kasprowicz, R. E. (2017). Foreign Language Educators' Exposure to Research: Reported Experiences, Exposure Via Citations, and a Proposal for Action. *The Modern Language Journal*, *101*(4), 613–642. 10.1111/modl.12426

Marsden, E. J., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews, and recommendations for the field. *Language Learning*, *68*(2), 321-391. DOI: 10.1111/lang.12286

Marsden, E. J., Morgan-Short, K., Trofimovich, P., & Ellis, N. (2018). Introducing Registered Reports at *Language Learning*: Promoting transparency, replication, and a synthetic ethic in the language sciences. *Language Learning*, *68*(2), 309-320. DOI: 10.1111/lang.12284

Marsden, E., Myles, F., Rule, S., & Mitchell, R. (2003). Using CHILDES tools for researching second language acquisition. In S. Sarangi & T. van Leeuwen (Eds.), Applied linguistics and communities of practice (Vol. 18, pp. 98–113). London: BAAL/Continuum.

Marsden, E. J., & Plonsky, L. (2018). Data, open science, and methodological reform in second language acquisition research. In A. Gudmestad, & A. Edmonds (Eds.), *Critical reflections on data in second language acquisition.* Philadelphia, PA: John Benjamins.

Marsden, E. J., Thompson, S., & Plonsky, L. (2017). Open science in second language acquisition research: The IRIS repository of research materials and data. In C. Granget, M-A. Dat, D. Guedat-Bittighoffer, & C. Cuet (Eds.), *Connaissances et Usages en L2 / Knowledge and Usage in L2: SHS Web of Conferences* (Vol. 38). DOI: 10.1051/shsconf/20173800013

Marsden, E., Thompson, S., & Plonsky, L. (2018). A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics 39*(5), 861–904.

Marsden, E., Trofimovich, P., & Ellis, N. (2019). Extending the reach of research:

Introducing open accessible summaries at Language Learning [Editorial]. *Language Learning,* 69, v - vii.

McKiernan et al. (2016) Point of view: How Open Science helps researchers succeed. *eLife*;5:e16800. DOI: 10.7554/eLife.16800

Mellers, B. A., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, 12, 269–275. https://doi.org/10.1111/1467-9280.00350

Mitchell, R., Domínguez, L., Arche, M., Myles, F., & Marsden, E. (2008). SPLLOC: A new database for Spanish second language acquisition research. *EuroSLA Yearbook*, 8, 287–304.

Mizumoto, A. & Plonsky, L. (2016) R as Lingua Franca: Advantages of using R for quantitative research in applied linguistics. *Applied Linguistics* 37(2), 284–291

Morgan-Short, K. Marsden, E., Heil, J., Issa. B., Leow, R., Mikhaylova, A.,… Szudarski, P. (2018). Effects of attending to form whilst comprehending: A multi-site replication study. *Language Learning,* 68 (2), 392-437. DOI: 10.1111/lang.12292

National Science Foundation (2015). Social, behavioral, and economic sciences perspectives on robust and reliable science. *Report of the subcommittee on replicability in science.* Retrieved from http://www.nsf.gov/sbe/AC_Materials/SBE_Robust_and_Reliable_Research_Report.pdf (1 November, 2017).

Morey, R. D. et al. (2016) The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review. *R. Soc. open sci*.3: 150547. http://dx.doi.org/10.1098/rsos.150547

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., du Sert, N. P., ... & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*, 0021.

Norouzian, R., de Miranda, M. A., & Plonsky, L. (2018). The Bayesian revolution in second language research: An applied approach. *Language Learning. 68,* 1032-1075.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, D., Breckler, S. J., et al. (2015). TOP guidelines. Retrieved November 1, 2017, from https://cos.io/top

Piñeiro, R., & Rosenblatt, F. (2016). Pre-analysis plans for qualitative research/Planes Pre-Analisis para la investigacion cualitativa. *Revista de Ciencia Politica*, *36*(3), 785-797.

Plonsky, L., Egbert, J., & LaFlair, G. T. (2015). Bootstrapping in applied linguistics: Assessing its potential using shared data. *Applied Linguistics*, *36*, 591-610.

Piwowar H.A. & Vision T.J. (2013). Data reuse and the open data citation advantage. PeerJ 1:e175 https://dx.doi.org/10.7717/peerj.175

Plonsky, L., Marsden, E., Crowther, D., Gass, S., & Spinner, P. (2019). A methodological synthesis and meta-analysis of acceptability judgment tests. *Second Language Research.* https://doi.org/10.1177/0267658319828413.

Tracy-Ventura, N. & Huensch A. (2018) The potential of publicly shared longitudinal learner corpora in SLA research. In A. Gudmestad & A. Edmonds (Eds.) *Critical Reflections on Data in Second Language Acquisition*. 149-170

Trofimovich, P., & Ellis, N. (2015). Open science badges [Editorial]. *Language Learning*, 65, v–vi.

van den Broek G, S. E., Takashima, A. Segers, E., & Verhoeven, L. (2018). Contextual richness and word learning: Context enhances comprehension but retrieval enhances retention. *Language Learning* 68, 546–585.

Wagner, A. B. (2010). Open access citation advantage: An annotated bibliography. Issues in Science, Technology & Librarianship. 60: Winter. Retrieved 29 January 2019 from http://www.istl.org/10-winter/article2.html

Wicherts, J., Bakker, M., & Molenaar, D. (2011). Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. *PLoS One*, 6, e26828