The Point of Replication: Investigating Social Psychologists' Beliefs on Direct and Conceptual Replication

Colm O'Fuarthain

S4018958

Department of Psychology, University of Groningen

PSB3E-BT15: Bachelor Thesis

Date of Completion: February 9, 2022

2122_1a_21

Supervisor: J.M. Hoek, N. Schwarzbach

Second evaluator: Dr. I. Wessels

In collaboration with: Maria Bompa, Kaiti Evgeniou, Rafael Funke, Larissa Hoß, Robert van Ark

Month 02, 2022

Abstract

Many believe social psychology to be in crisis due to poor replicability. As a result, a "Reform Movement" has developed that consists of researchers pushing for improved methodological practice. Regarding reform, it can be considered that a spectrum of perspectives exists ranging from "reformer" (desiring regulatory reform) to "challenger" (challenging the reform movement). However, a general understanding of social psychologists' perspectives on replication and its functions is not well established. This knowledge gap poses a problem for reformers, as the uptake of reform proposals is mostly self-determined by social science researchers. It is also an opportunity for social psychologists to provide feedback that has largely gone unsolicited. In this pilot study, I used quantitative and qualitative means to investigate 77 social psychologists' beliefs about replication. Here, I distinguished two typologies that often dominate replication discussion in social psychology; direct replication and conceptual replication. The descriptive results illustrated strong support for both direct and conceptual replication in social psychology, with conceptual replication receiving higher agreement using 1-100 analogue scales. Furthermore, the contradicting ideas that 1) social psychology is highly context-sensitive and subjective and that 2) objectivity and universalism are fundamental cornerstones of research may both underlie beliefs concerning replication. Our results suggest the need for replication studies to be locally suited to social psychology. Our respondents' nuanced understanding of replication should encourage reformers to include social psychologists when developing new replication reforms. Future research should build on these exploratory results and investigate other purposes of replication more robustly.

Keywords: replication crisis, direct replication, conceptual replication, social psychology

Replication Crisis

Recently, it has been declared that social psychology, amongst other scientific fields, faces a replication crisis (Open Science Collaboration, 2015; Shrout & Rodgers, 2018). The arguments of Ioannidis (2005) that most published scientific findings are false were later supported by a multi-lab replication effort (Open Science Collaboration, 2015), which estimated a replication rate of merely 25% for social psychology. Increasing recognition of poor reproducibility drew attention to a host of methodological and cultural issues in psychological science besides replication (Munafò et al., 2017; Wiggins & Christopherson, 2019). For this reason, some observers instead now refer to the current period as a "crisis of confidence" or "the credibility revolution" (Barrett, 2019; Earp & Trafimow, 2015). Regardless of the term used, this scientific *zeitgeist* has sparked debate concerning improved research practices in psychology.

In response to the crisis narrative, several solutions for psychology's low replication rate have been proposed. Broadly speaking, these solutions and their sponsors, often described under the umbrella term "Reform Movement", attempt to realign scientific practice with long-held scientific values, such as transparency, objectivity, and prospect of falsification (Watson, 2015). These proposals include, but are not limited to, preregistration of methods and analysis plans, open research access, open peer review, and open sharing of data to facilitate replication (Nosek & Bar-Anan, 2012; Nosek et al., 2012).

However, in addition to disagreement concerning the proposed solutions' ability to remedy psychology's replication "problems", it is disputed whether problems indeed exist (Peterson & Panofsky, 2020). On this subject, Marowski (2021) considers a dichotomy of views to exist; a party of "reformers", designating those pushing for reform, and "challengers", the

label used for critics of the reform movement. Although in reality the positions individuals may take likely represent a continuum rather than a strict dichotomy, this framework is still helpful in exploring the epistemological and ontological underpinnings of perspectives on replication. Therefore, using this framework, I will explore how reformers and challengers perceive replicability, the different forms of replication, and its relevance for social psychology, a field in which disputes over the value of replication tend to occur (Earp and Trafimow, 2015). This framework will assist me in investigating the purposes social psychologists assign to replication in their field, and the beliefs underlying their positions.

Theoretical Framework

Reformers and Challengers

According to Morawski (2021), "reformers" represent a perspective that posits scientific effects to be real, stable, and consistently observable despite changing conditions. This view often champions a Popperian view that holds falsification, and thus reproducibility, as a cornerstone of science (Derksen, 2019). Their faith in these traditional values of science leads to a perspective that psychology's failures are not a result of faulty epistemology or ontology, but rather an issue of incentivized and subjective researchers impeding objective interpretation (Flis, 2019; Morawski, 2020). As such, reformers emphasize regulating psychologists' scientific conduct through methodological reform as the only way to align practice with long-held scientific ideals, such as transparency and objectivity (Morawski, 2021). In other words, a main goal of those adhering to a reformist perspective is to develop and encourage a broad set of standards for how scientific research should be conducted, regardless of discipline.

The challengers' perspective, meanwhile, "take[s] psychology's objects to be complex, fluid, and context-sensitive" (Morawski, 2021, p. 2). This view claims psychological science to be distinctly different from the natural sciences that objectivists attempt to model, as the human mind is not reproducible from one situation to another (Iso-Ahola, 2017). Challengers also emphasize the need for researcher creativity and multiple investigative techniques (Morawski, 2021). Adherents to this perspective criticize reformers' apparent lack of appreciation for idiosyncratic differences between scientific fields (Penders et al. 2019; Guttinger, 2020). For example, some claim that imposing the methodological regulations of one singular epistemic culture reduces disciplinary pluralism that makes science indeed science (Freese and Peterson, 2017). Other complaints are simpler, focusing on the quixotic nature of reformers' regulations. As one sceptic at the Metascience Symposium recently stated, "These are really smart people, … but, Jesus, you got to be in the trench to see how it really works and how slippery the truth really is" (Peterson & Panofsky, 2020, p. 24). This quote speaks to a recurring theme; a potential impasse between what reformers believe *should* be done and what actual psychological researchers believe *can* be done.

As replication is still considered by some as the greatest focus of reformers (Derksen, 2019), I will now describe how reformers and challengers view replicability.

Conceptual and Direct Replication

Replication is generally considered as an attempt to repeat an original study. However, under this umbrella term, many different types have been specified that vary in terms of purposes and approach (Leonelli, 2018; Zwaan et al., 2018). In psychology, two replication types dominate: direct replication and conceptual replication (Guttinger, 2020). Direct replication is

often considered as an attempt to follow the procedure and conditions of an original study as closely as possible (Crandall & Sherman, 2016). Therefore, its results may act as a test of the reliability of the original finding and the robustness of the underlying theory.

Reformers believe that direct replications are necessary for self-correction in science and hold them as the only valid way to guard against false-positives (Freese & Peterson, 2017). According to a reformist perspective, without self-correction, science just becomes a series of uninteresting, unfalsifiable results, especially in today's age where bad incentive structures exist (Simons, 2015). Such efforts may be considered crucial for a field like social psychology, which has historically strained itself to be considered as "real" science (Elms, 1975).

Challengers object to these claims in many ways (e.g. Feest, 2019). The most common and consequential objection is that in highly context-sensitive fields like social psychology, direct replication is not capable of serving the purposes reformers assign to it (Guttinger, 2020). Even if the replication perfectly adheres to the rest of the original study's methodology, there are a number of conditions, such as time period or population of participants, that will likely be different. In studies where the independent and dependent variables may be highly culturally or socially dependent (as is oftentimes the case in social psychology), these conditions may have resulted in the production of the effect (Stroebe & Strack, 2014). Furthermore, Peterson and Panofsky (2021) argue that tacit researcher knowledge, which goes unacknowledged in the methodologies of studies, may also be required in effectively testing the underlying theory and producing an effect. Taken together, direct replications appear impossible to conduct, rendering its success or failure not only non-corrective, but also uninteresting.

Considering this, some challengers advocate the utility of conceptual replication studies over direct replications. Here, conceptual replication is the attempt to test the underlying theory behind an original study by deliberately altering the study design, population, and operationalization of the study variables (Crandall & Sherman, 2016). According to a challenger perspective, the advantage of conceptual replication is that testing our ideas in a novel manner not only increases or decreases our confidence in the theory but may also expand the theory to new contexts (Crandall & Sherman, 2016). In this way, conceptual replication not only serves the purposes reformers claim direct replication does, but also facilitates a deeper understanding of the extent to which our theories truly reflect phenomena (Stroebe & Strack, 2014).

Naturally, adherents to the reform perspective respond in critical manners. While acknowledging that conceptual replication has a role to play in psychological science and that direct replications can never be perfect recreations of original studies, they reject that conceptual replication can be a sober process of verification (Simons, 2014). Here, reformers point out that due to the customary disregard for nonsignificant results, failed conceptual replications become difficult to publish (Freese & Peterson, 2017; Hales et al., 2015). What then occurs is a file-drawer problem in which “supposed” failed tests of the theories of contentious studies go unnoticed, leading to unchallengeable theories. This same file drawer problem does not afflict direct replication studies, as both significant and non-significant results can be informative to the reliability of the original study, and therefore publishable (Hales et al., 2015). Regardless, if a failed conceptual replication attempt is indeed published, reformers often contend that the theory is revised to accommodate the new results (Freese & Peterson, 2017). Reformers may argue that rather than developing theories, the accommodations introduced by conceptual replications are

simply ad-hoc explanations for studies and theories with validity issues, thereby weakening a discipline's credibility.

Motivation and Research Aims

Metascientific scholars often compare the value of conceptual and direct replication. Therefore, a spectrum of perspectives held by those assessing science's practices, ranging from reformer to challenger, is well established. However, a broad understanding of social psychologists' views on replication is not known. Indeed, few studies, to our knowledge, have surveyed psychologists' beliefs on reform and replication specifically. Recently, Agnoli et al. (2021) surveyed Italian and Australian psychologists' about the role of replication in their research. Majorities of the participants indicated that replication studies are very important, that more should be conducted, and that more resources should be dedicated to replications. However, the work failed to delineate direct and conceptual replication. Furthermore, epistemological beliefs underlying psychologists' views on replication were not sought.

Currently, the implementation of the Reform Movement's proposals is mostly self-determined by social scientists (Washburn et al., 2018). Therefore, an incomplete understanding of social psychologists beliefs concerning direct and conceptual replication represents a significant knowledge gap and poses a problem for reformers. A lack of understanding of how social psychologists view the role of replication in their field makes bridging any epistemological and methodological impasse much more difficult. If positions of social psychologists on the different forms of reproducibility were known, the reform movement may be able to improve their current conceptualizations for replication and in turn, improve uptake of replication in social psychology (Peterson and Panofsky, 2020). Social psychologists

may also wish to reveal their views on replication; at a recent Metascience Symposium, some disciplinary researchers expressed their desire for reformists to hear their feedback (Peterson and Panofsky, 2020).

Considering this, the current research investigates where social psychologists lie on the reformer-challenger spectrum by asking two questions; 1) what purposes (i.e. with respect to conceptual and direct replication) do social psychologists believe replication should serve in their field, and 2) why they are committed to these purposes. Concerning the latter, I am seeking both practical and epistemological grounds behind their beliefs.

The current work is a pilot study for future studies which will assess a wider population of social psychologists. Previous undergraduate theses assessed psychologists' views on the replication crisis by qualitative, explorative means. Here, we attempt to build on such work by developing and testing a novel survey that can be used to evaluate social psychologists' opinions on the reform movement on a larger scale. Accordingly, we aim to produce exploratory, descriptive results which may further inform possible improvements for the survey and future studies.

Methods

Researcher Description and Reflexivity

As the study is exploratory and descriptive in nature, researcher reflexivity is valuable for my analysis. I am a third-year international psychology student at the University of Groningen, who has previously completed another bachelor's degree in engineering. I find myself possessing conflicting views on the perceived replication crisis. While I am convinced that the reform movement has been insensitive to disciplinary differences, I harbour reservations against some

defenses offered by challengers. I am concerned that opposition to direct replication provides avenues for researchers to make ad-hoc adjustments for flawed studies and theories. Despite my skepticism, it should be noted that I am not well acquainted with the creative component required in research. My past education in a STEM-field and interest in data science as a career direction leads me to be data- and methodology-oriented.

Ethical Considerations

Prior to sample collection, the project was approved by the BSS-Psychology Ethics Committee at the University of Groningen. The code for approval is PSY-2122-S-0016. Participation was voluntary and could be ended at any time during the survey. Participants also provided informed consent prior to data collection. Email addresses required for survey dissemination were available publicly. Data was processed anonymously. We did not collect personal data such as name or email address during the survey or metadata such as IP addresses; therefore, a participant's answers could not be linked back to them.

Participants

Our target population consists of social psychologists. Since the main aim of our pilot study was to receive and integrate the feedback on our survey before future distribution, we aimed for a relatively small minimum sample size. Furthermore, as an important part of our study was the thematic analysis of the open questions, we regarded a minimum of 20 participants as sufficient. We deemed a response rate of 10% as realistic and in turn contacted 246 psychologists. Using a convenience sample, we approached researchers from University of Groningen (UG) (102), VU Amsterdam (27), University of Amsterdam (47), Tilburg University

(34), Radboud University (36). The universities mentioned above were selected because they clearly separated social psychology from other departments such as organizational psychology and because email addresses were easily extractable from researcher profiles on department webpages. The selection sequence began with the present author's own university (UG) and thereafter continued by decreasing city population size within the Netherlands. After extracting all email addresses from the fifth university (RU), we had obtained 246 email addresses and stopped the sampling procedure.

Procedure

We sent a Qualtrics link and the informed consent form out to our target population via email. This email explained that we were looking for social psychologists to share their perceptions on the crisis debate, the reform movement and their methodological proposals, as seen in Appendix A. Respondents were also informed that the resulting data will be used for several bachelor theses and may eventually contribute to publication in a scientific journal. The survey ran for three weeks and reminders to answer the survey were sent one and two weeks after our initial invitation.

Limitations of the Sampling Procedure

As we worked with a convenience sample, certain types of responses may be under- or overrepresented. Moreover, it is quite likely that researchers who participated in our study are different from those who chose not to fill out the survey. One possibility is that those with stronger opinions on the open science movement and its practices are more likely to answer. In addition, a heated public Twitter debate regarding the reform movement, escalating when a

“reformer” publicly criticized a newly published social psychology study, occurred days prior to survey distribution (Brown, 2021). Due to the perceived denunciative and personal nature of the criticism, and the surrounding tone debate (e.g. Derksen & Field, 2021) underlying reform discussions, polarised opinions might have been exacerbated by the debate. The possibility of bias is taken into account in the discussion of the results.

Survey Design

Qualitative work from former bachelor and master theses investigating psychologists' perspectives on the replication crisis and open science practices contributed to item generation for this novel survey (Futjes, 2021; Hershler, 2021; Nicolai, 2021; Pool, 2021; Sales, 2021; Schmidt, 2021; Schwarzbach, 2021). We also consulted survey designs that assessed the role of replication in ecology (Fraser et al., 2020) and psychology (Agnoli et al., 2021) and literature considering epistemological and ontological differences between reformers and challengers (Derksen, 2019; Flis, 2019; Morawski, 2019).

The survey consists of four core sections and it was anticipated that it would take 15 minutes for participants to complete it. First, the participants were asked about their epistemological and ontological views regarding (psychological) science. The second section, most important for the current research purposes, broadly investigated the participants’ views on (1) the purposes of new replication studies (generalizability, falsification, and/or confirmation of established results), (2) the importance of conducting direct and conceptual replication, and (3) the extent to which direct and conceptual replication are indicative of research quality in social psychology. Then, with an open-ended question, we asked why participants believed that either

direct or conceptual replication, when successful, is or is not indicative of research quality. Further optional open-ended questions asked if there were other important quality indicators outside replication in their field, and allowed respondents to provide broad thoughts on replication or the replication items. The third block of questions gathered information on open science concepts, practices and applications. Lastly, participants evaluated critical reflections on the reform movement and obstacles to proposed reforms. The complete survey can be found in Appendix B.

Data Analysis

Quantitative Analysis

Of the 246 invitations, we collected 94 responses (a response rate of approximately 38%). We excluded participants who indicated another field of expertise (17 people) outside social psychology, along with respondents who indicated a lack of effort (0 people) or lack of honesty (1 non-social psychologist), yielding a response sample of 77 social psychologists. Of these remaining responses, 15 were incomplete.

Before data collection and analysis, we produced a quantitative analysis plan and later conducted the analysis in accordance. To address the question of what purposes social psychologists believe replication should serve, seven quantitative items from the survey were descriptively analysed: Q24 (generalization), Q25 (falsification), Q26 (confirmation), Q28 (importance of direct replications), Q29 (importance of conceptual replications), Q30 (direct replication as indication of research quality), and Q31 (conceptual replication as indication of research quality). The quantitative items yielded 0-100 scores (100 corresponding to complete

agreement and 0 to no agreement at all), or a missing value when a “not applicable to my field” choice was selected. In such cases, these responses were excluded from analysis. Medians and interquartile range (IQR) were calculated from the 0-100 scores. We used boxplots to visualize the scores of the seven replication items, including all data points and highlighted outliers. To conduct median and spread calculations, we considered all numerical responses as valuable and as such, we did not remove partial responses.

We also considered descriptive correlations to be potentially informative. For example, a strong negative correlation ($r < -.6$) between items Q25 (falsification) and Q26 (confirmation) could show a pattern wherein people believe that new replication studies should either falsify *or* confirm existing results, whereas a strong positive correlation ($r > .6$) could show a pattern wherein people believe that new replication studies should both falsify *and* confirm existing results. No predictions were made as to the outcomes of the correlation computations. The data violated the assumptions of linearity, normality, and absence of outliers for Pearson's correlation (see Figure C.1, Figure C.2, and Table C.1). Therefore, Spearman’s rank-order correlation was computed for a single pair for items Q30|Q31 (extent to which each replication type is indicative of research quality), which met its assumption of monotonicity. A scatterplot was used to visualize this relationship. Although desired, correlations were not computed for each pair in the Q24-Q26 cluster (purposes underlying replication) and for Q28|Q29 (extent to which each replication type is important), as they did not meet the assumption of monotonicity (see Figure C.1).

We analyzed the data using R, an open-source statistical software programme (R Development Team, 2018). The code used for analysis is included in Appendix D.

Qualitative Analysis

Thematic analysis was used to address the *why* behind social psychologists' beliefs about replication's purposes. As thematic analysis by nature encourages deviations, we have explicated the actual procedure with which the qualitative analysis was carried out below.

Using Braun and Clarke (2006)'s description of thematic analysis as a guide, we recursively carried out six phases: 1) familiarization with the data, 2) code generation, 3) searching for themes, 4) reviewing themes, 5) defining and naming themes, and 6) writing the report. The conducted thematic analysis is inductive. We described what is explicitly mentioned in the data and did not go beyond what has been said by participants by examining underlying ideas and assumptions.

We first familiarized ourselves with all open question responses to the survey. All contributors (Kate Evgeniou, Colm O'Fuarthain, and Robert van Ark) individually highlighted informative responses about replication and generated a non-exhaustive list of initial codes that described the highlighted responses, without attaching them to specific text extracts.

Thereafter, all contributors collaborated to generate one collective codebook. Similar codes identified by multiple collaborators were combined to form one code. Codes that were unique to a collaborator were also included. We ensured that codes ending up in the codebook needed to be specific and distinct by creating definitions for all codes. This preliminary codebook consisted of 25 codes.

We then used the codebook to perform another round of coding. Codes were only applied if a consensus was met by all collaborators that the code fit that specific data extract. Once more,

an emphasis was placed on avoiding interpretation of the data that would suit our particular research interests. During the coding process, eight new codes were added to describe relevant data that existing codes did not yet describe. Four codes that were not assigned to any pieces of data (i.e. a frequency of 0) were removed from the codebook.

We then began the theme generation process. Here, all collaborators first independently sorted similar codes into themes (i.e. groups of codes) and thought of initial theme names. Following this, we compared our independent themes and discussed the appropriate groupings and theme names. The process of grouping and naming was done recursively. Simultaneously with these discussions, codes were again assessed on their distinctness and relevance. Consequently, we removed four more codes and reassigned the relevant text extracts. As a result of this process, we identified six themes from 25 codes. The final codebook can be found in Appendix E.

Results

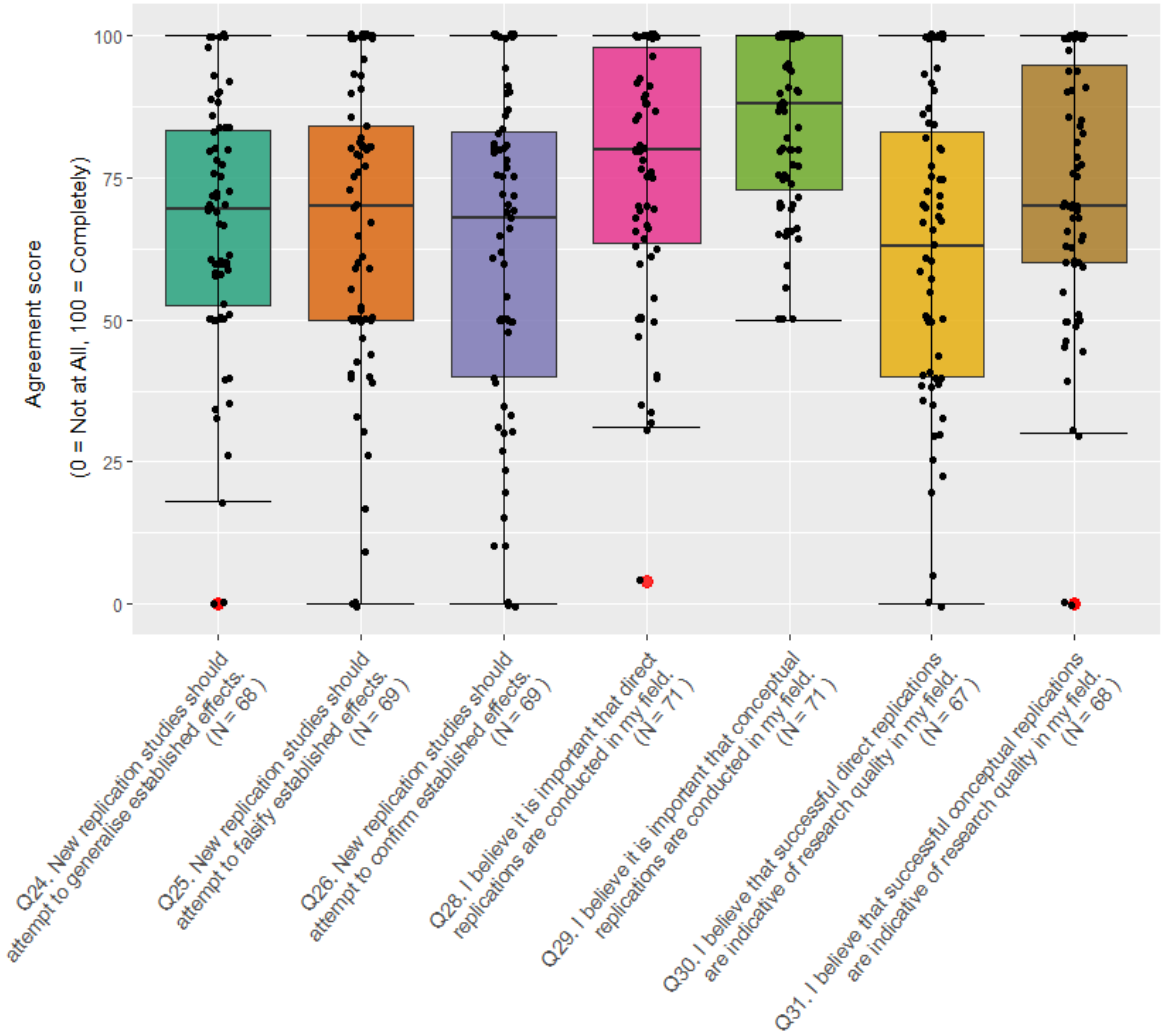
Of our sample, 73 participants worked in the Netherlands and four in various other countries. The median number of years that participants had been working in academia was six years, however, the spread was quite large (IQR: 7 [4, 11]). Considering current job positions, 35 participants were PhD students, 17 assistant professors, eight associate professors, six full professors, four post-doctorates, and seven respondents identified working in other positions.

Quantitative Results

Figure 1 shows the responses on the seven items asking about the value and aims of replication. Each item is discussed in detail below.

Figure 1

Boxplots for the seven relevant items (with jitter function; outliers are indicated by red dots)



Generalizability, Falsification, and Confirmation

Participants were asked to assess their agreement with purposes that may predicate replication. Results indicate similar patterns of responses on the statements that new studies should attempt to generalize (Mdn = 69.5, N = 68), falsify (Mdn = 70, N = 69), and confirm

(Mdn = 68, N = 69) established effects. However, the spread of agreement among respondents was less similar across the three items (generalizability: IQR 30.8 [52.5, 83.3]; falsification: IQR 34 [50, 84]; confirmation: IQR 43 [40, 83]).

Importance of Direct and Conceptual Replication

There was high agreement that the conduction of direct replication (Mdn = 80, N = 71) as well as that the conduction of conceptual replication (Mdn = 88, N = 71) is important in social psychology. Moreover, the spread of agreement among participants was smaller for conceptual replication (IQR: 27 [73, 100]) than for direct replication (IQR: 34.5 [63.5, 98]). Notably, the lowest agreement scores concerning the importance of conceptual replication and direct replication were 50 and 4, respectively. This was not observed in any of the other item scales, which all recorded responses at the lowest possible value: zero.

Successful Replication as an Indication of Research Quality

When we asked participants if successful direct/conceptual replication was indicative of research quality, the median agreement for direct replication (Mdn = 63, N = 67) was lower than that of conceptual replication (Mdn = 70, N = 68). Again, the spread of agreement among participants was found to be smaller for conceptual replication (IQR: 34.8 [60, 94.8]) than for direct replication (IQR: 43 [40, 83]).

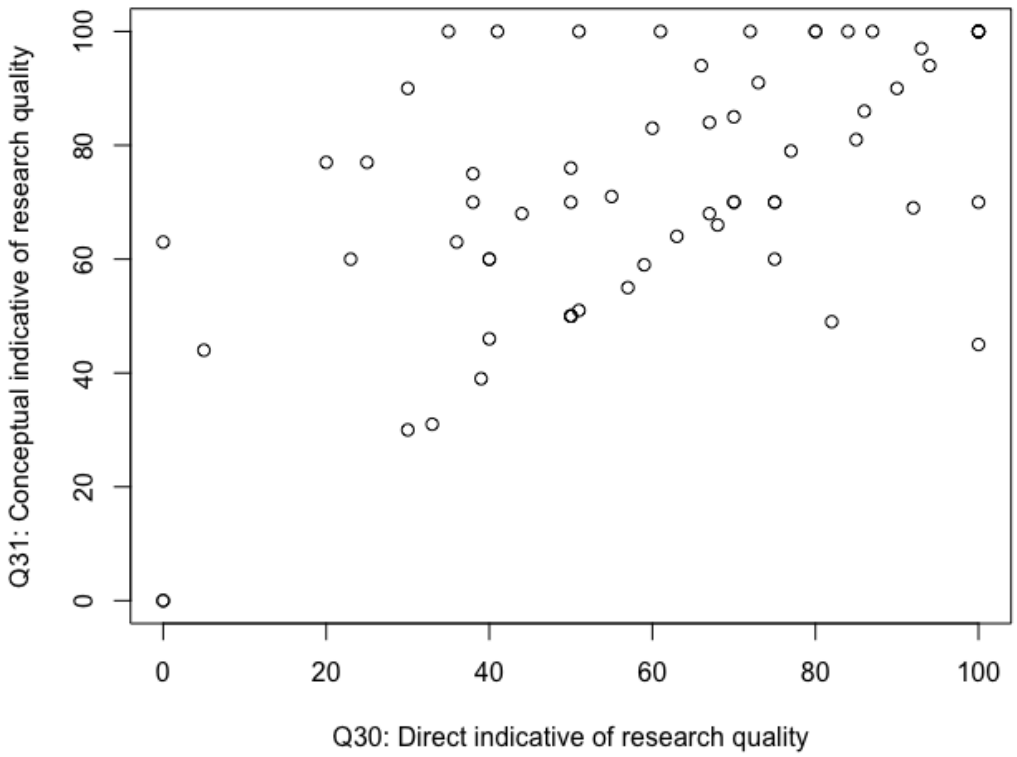
Correlation Pattern

Spearman's correlation coefficient was found to be moderately strong for agreement scores on the statements that successful direct and conceptual replication are indicative of research quality (pair Q30-31, r_S = .55). As shown in Figure 2, these item pairs share a

monotonically increasing relationship, where high agreement on one statement (i.e. direct replication) often is paired with high agreement on the other (i.e. conceptual replication).

Figure 2

Scatter plot for the relationship between responses on agreement that successful direct and conceptual replication are indicative of research quality



Qualitative Results

Using thematic analysis, 25 codes were identified, resulting in eight themes (see Table 1). The number of unique responses (i.e. the number of free-text responses in which a code was identified) is indicated in brackets for each code. Appendix E contains the complete codebook containing all codes and response examples.

Table 1

Identified themes with pertaining codes (number of instances in brackets)

Themes	Assigned Codes
Process and Conclusions of Replication	*Successful single replication is not conclusive (7)* *Replication is a learning and quality process (4)* *Incentives and bias for failed replications (2)* *Replication should not have purposes (1)*
Direct Replication: Functionality and Drawbacks	*Direct replication for reliability (9)* *Direct replication is uninformative (5)* *Direct replication reinforces original bias/mistakes (3)* *Direct replication for robustness (3)* *Direct replication indicative of quality of methodology (3)* *Direct replication not applicable in social psychology (2)*
Conceptual replication: functionality	*Conceptual replication for generalizability (12)* *Conceptual replication for theory development (7)* *Conceptual replication for validity (4)* *Conceptual replication and context-sensitivity (3)* *Conceptual replication overcomes methodological limitations and bias (3)* *Conceptual replication for theory boundary conditions (2)*
Broad judgements regarding both replication types	*Conceptual over direct replication (13)* *Both replication types are uninformative (3)* *Both replication types are similarly important (2)* *Direct over conceptual replication (2)* *Nature of study determines which type of replication (1)*
Epistemology & Ontology	*Social psychology and context-sensitivity (10)* *Objectivity and truth as foundations for science (7)* *Research is subjective by nature (6)* *Universal and stable effects exist in science (2)*

Process and Conclusions of Replication

Participants had differing views regarding how replications should be performed and what can be concluded from them. One prominent perspective was that less attention should be paid to the outcomes of single studies, which in and of themselves are inconclusive or biased.

For both direct and conceptual replication, a single outcome can not tell us definitively whether a finding is "*true*", but "*can reveal more nuances to our understanding of the effect*". Instead, greater attention should be paid to "*patterns*" that may arise across multiple studies. As a whole, replication should be viewed as a process that allows researchers to learn and improve the field's research quality: "*It's not the results of the replication that matters. What matters is that we do them and learn from them.*"

Direct Replication: Functionality and Drawbacks

Many responses associated direct replications with strengthening the reliability and robustness within their field. Due to the unchanging circumstances between the original study and replication study, a successful direct replication could increase our confidence in not only the original effect, but also "*establish the quality of the research methods, the paradigms, and design".* However, some responses argued against this claim, suggesting that direct replication reinforces biased designs, thereby failing to say anything about research quality. Instead, successfully conducting a study in the same manner *"is not very meaningful because the results of the field are so contextual",* rendering direct replication rather uninformative.

Conceptual Replication: Functionality

Notably, respondents did not highlight drawbacks about conceptual replication, instead drawing attention to various positive functions of conceptual replication. Overwhelmingly, conceptual replication was seen as a tool to generalize findings, informing us if "*the way we study things can be applied to other contexts or samples or methods".* Through generalizing results, conceptual replication was also considered to assist theory development. As one

respondent said, by *"indicating in what context something is and is not present … we can build on [it] theoretically"*. In addition, conceptual replication can "*overcome the methodological limitations or unique methodological features of the previous studies"* and offer *"convergent validity"*. Lastly, as conceptual replication accounts for changes in context, and "*... social psychological theories can be quite time, culture and context dependent*", conceptual replication was considered to function well in social psychology.

Epistemology & Ontology

Participants' epistemological and ontological convictions were considered to be informative with regard to the purposes they assigned replication. For example, many responses agreed with the extract above that social psychology and its objects are highly context sensitive. Linking it to replication, one suggested it "*might be only possible in more stable situations of complex systems, and hence is not a good concept to study more turbulent stages in social systems*". Similarly, respondents stressed that the subjectivity of the researcher must be acknowledged: *"the background of the researcher can influence the kinds of research questions that they ask and the way the research is designed"*. However, in possible incongruence, Popperian ideas of objectivity and truth-seeking were also considered to be worthwhile: *"just like Popper, I believe we can get closer to the truth via conjucture [sic] and refutation"*. Overall, recognizing the complexity and epistemological nuance of conducting science in social psychology may be required. As one scientist claimed, "*I think there are phenomena in reality that have stable characteristics, but that can never be described in any words. Multiple theories are needed to approach the quality and complexity of these real phenomena. The phenomena*

themselves will never be described in their essence but they can be described and predicted with the clumsy tools we call theories."

Broad Judgements on Direct and Conceptual Replication

When comparing direct and conceptual replication, more responses indicated conceptual replication to be distinctly more valuable than direct replication, highlighting conceptual replication's functions. For example, if "*you can replicate a finding conceptually it contributes more to possibly finding theories applicable across different contexts than just copying the exact context in which a study has been first conducted"*. Relatively fewer responses identified direct replication to be more valuable, criticizing the extent of "*interpretation and degrees of freedom"* afforded to conceptual replications. Some participants highlighted the utility of conducting both types, while others believed neither replication type is necessarily informative: "*any form of 'successful' replication should not be considered a sign of quality. Finding support for the existence of an effect is only as much informative as finding support for the absence of it.*"

Discussion

Summary

Our respondents strongly agreed that both conceptual and direct replication were of value to their field. Conceptual replication items had lower spreads and higher minimum scores, indicating that respondents were more willing to at least moderately endorse conceptual replication. This was supported by free-text responses which indicated that conceptual replication had more functions and was a better fit for social psychology than direct replication.

Respondents indicated similarly strong agreement with all three purposes of replication (generalizing, falsifying, or confirming).

Interpretation and Relation to Existing Literature

Replication and its Purposes

Earlier, I introduced Morawski's (2021) reformer-challenger framework as a spectrum of views in which challengers were considered more likely to endorse conceptual replication, while reformers typically advocate for direct replication. Our participants, however, endorsed both types of replication to a high degree. This supports the findings of a recent survey of Italian and Australian psychologists, where replication studies were considered crucial and respondents wished more were conducted (Agnoli, 2021). Furthermore, descriptive correlation in our results showed that respondents who agreed with the conceptual replication being indicative of research quality were more likely to agree that direct replication also indicated quality.

This would appear to be in contrast to conventional notions that endorsement of direct and conceptual replication is competing (Feest, 2019; Lynch Jr et al., 2015; Morawski, 2021). It could be that other factors besides the kind of replication used may underlie perceived utility. Penders et al. (2020) suggest that the utility of replication may rely on the type of question researchers ask in the original study. Researchers may not seek factual "X leads to Y" statements, but rather intend to generate a diversity of arguments. Thus, the authors suggest that fields such as social psychology encourage a plurality of conclusions as part of normal science. As one participant noted, both direct and conceptual replication are important and their usage may depend on the nature of the study. Providing an example, they said "*observing behavior in*

experiments … direct replications have value. If we're talking about field experiments, conceptual replication might be more fruitful." Leonelli (2018) argues this case, distinguishing between more replicable research setups, such as standardized experiments, to less replicable, like participant observation. Leonelli (2018) argues that in the context of social psychology's typically semi-standardized experimental setup, both types of replication have their potential utility. This assertion appears to be reflected in our results.

In line with respondents endorsing both conceptual and direct replication, we observed equally strong agreement scores for generalizing results, conventionally a function of conceptual replication (Stroebe & Strack, 2014), and confirming and falsifying results, functions of direct replication (Simon, 2014). However, the higher endorsement of explicit conceptual replication items did not materialize in the implicit purposes items. This contrasting result may suggest that we failed to identify and quantitatively assess other purposes underlying conceptual and direct replication. Indeed, free-text responses also identified theory development, overcoming methodological limitations in the original study, and testing construct validity as key functions for conceptual replication, as well as replication being a learning process, regardless of the results or type used.

Our respondents' functions appear to be unidentified in the existing literature. Schmidt (2009) suggests that replication can also be performed to control for a lack of internal validity or address researcher fraud. But as Machery (2020) notes, beyond Schmidt (2009) there is surprisingly little discussion of what functions replication should serve, with scholars instead defining procedures and superiority of different types of replication.

In short, multiple purposes of replication may exist besides the ones quantitatively assessed in the current study, representing a significant limitation. A proper discussion of replication and its functions, as proposed by Machery (2020), is required.

Beliefs Underlying Replications Purposes

Respondents highlighted epistemological and ontological commitments that may inform why they strongly agreed to the purposes of generalizing, falsifying and confirming results. Some respondents aligned themselves with prototypical challenger perspectives (Morawski, 2021) by suggesting that social psychology is highly context-sensitive and subjective. In such dynamic disciplines, conceptual replication may be a better epistemic tool (Crandall & Sherman, 2016). Responses also stressed the inevitable bias introduced by researchers, and that social psychologists should accept such subjectivity rather than try to eliminate it. This claim is supported by Wiggins and Christopherson (2019), who argue that psychology's commitment to a sort of objectivism simply hides bias under the cloak of objectivity. Upholding objectivity and universalism instead of subjectivity and contextualism may lead to future blind spots, where research must simply look "objective" to avoid scrutiny, concerns some respondents appear to share. This perhaps underscores the strong scores we observed for generalization as a function of replication, as this naturally lends itself towards an appreciation for context and uncertainty.

However, in contrast, other responses drew attention to objectivity and universalism as valuable and realistic cornerstones of scientific practice. Many retained Popperian ideas concerning falsification and took psychology's objects to be true, which Morawski (2021) considers to be affiliated with a reformist perspective. Such perspectives are considered to

sympathize with the aims of direct replication, that being to falsify or increase confidence in existing results (Derksen, 2019).

That these competing epistemologies were both highlighted in the response pool may simply mean that separate respondents align themselves differently along the reformer-challenger spectrum; while some criticize direct replication, others may hold falsification as a key tenet in social psychology. However, it may also mean that participants themselves may hold competing epistemological views, giving rise to their belief in both direct and conceptual replication. Although perhaps the least likely of the scenarios, to conceive of such contradictions is not hard considering the complexity of the situation. Take reformers, for example. As Flis (2019) points out, the reform movement is committed to Popperian scientific ideals and methodologies while often framing the problems of the replication crisis in heavily social, Kuhnian terms, taking the researcher as an inherently biased, irrational actor. Something similar may have been the case with our respondents, where an awareness of the context-sensitivity of their field may collide with a falsificationist research standard (Mulkay & Gilbert, 1981). Future work should investigate the degree to which social psychologists hold different epistemological positions along the reformer-challenger spectrum.

Limitations

Our data, in particular the rich free-text responses, can provide future directions for researchers to explore the role of replication social psychology. However, at the risk of being repetitive, it's important to once again acknowledge the descriptive and non-representative nature of the current study, as this has implications on what the data can and cannot tell us. As

participation was voluntary, I expect that social psychologists who are more interested in research practices or replication debates were more likely to respond. Bias may have further been amplified due to a highly public Twitter debate concerning the reform movement days prior to data collection. This event was mentioned in many of the open response questions, in which respondents criticized reformers' behaviours and tones. Taken together, it is clear that the data can only be interpreted with regard to our sample.

It should also be noted that our results would have benefited from more quantitative items assessing replication. We only utilized seven items; two for each replication type, and one for each of our explicit purposes. As mentioned earlier, considering the thematic analysis, this study may have quantitatively assessed too few purposes of replication. However, more items framing the utility of conceptual and direct replication in different manners could have helped illuminate the nuance that our respondents attribute to this subject. Many participants suggested that the items were hard to answer, with one even suggesting that they were extremely leading. A multitude of items may help eliminate measurement bias due to the framing of our questions.

Implications

This work has a number of implications for the field of social psychology and the reform movement. First, the support for conceptual replication and contextualism supports the idea that replication's value in social psychology should be considered locally (Guttinger, 2020). Overall, the reform movement would do well to ask *what works for social psychology?* when designing interventions aimed to improve the validity and reliability of findings.

That said, there was also broad agreement among participants about the positive utility of direct replication. Reformers should consider formulating replication in a system that would satisfy the idiosyncratic needs of social psychology while also providing an avenue for the falsification of results. For example, requiring a combination of successful direct and conceptual rather than their separate parts may strengthen the research quality of social psychology. In another approach, Leonelli (2018) suggests employing “scoping replications” in which experiments are re-run to identify the most relevant sources of variation, which may indirectly assist falsification efforts.

Our sample consisted mostly of new and young researchers, most being PhD students. Still, there was considerable support for ideas that would commonly align with a challenger perspective. However, this would seem to be in contradiction with the conventional image of the young researcher. Everett and Earp (2015) believe early-career researchers to be eager to conduct direct replications, but instead are beholden to a “publish or perish " culture which encourages novel and exciting studies. Indeed, Morawski (2021) even pits reformers and challengers against each other as “generation 2.0 versus generation 1.0” (p. 2). Instead, our results may suggest that younger researchers take a variety of positions along the reformer-challenger continuum that may be different from commonly held assumptions. This perhaps has practical relevance for both social psychologists and reformers, as it shows that the next generation may not be overwhelmingly committed to either end of the reformer-challenger spectrum.

Overall, it should be noted that many participants exhibited a great degree of nuance and familiarity when discussing direct replication and conceptual replication. This could illustrate that those in the reform movement should trust the knowledge that social psychology researchers

may have concerning methodological reform, and therefore make a greater attempt to include them in the conversation. Future efforts should be made to assess if a representative pool of social psychologists similarly reflects a nuanced understanding of replication and metascientific reform.

Concluding Remarks

In conclusion, we found support among social psychologists for both direct and conceptual replication. In terms of replication's goals, respondents strongly agreed that replication studies should attempt to generalize, falsify, and confirm existing results. We discovered beliefs highlighting the context-sensitivity and subjectivity of social psychology, and contradictingly, commitments to objectivity and universalism to be important considerations with regard to why they hold their beliefs about replication. Taken together, partakers in the crisis discussion must acknowledge that the reformer-challenger debate is a spectrum rather than a dichotomy. Concerning social psychologists, it is likely they hold various positions along this spectrum that are informed by the peculiarity of their field. To progress social psychology forward, discussions about reform likely require cooperation from those outside and within the reform movement.

References

Barrett, L. F. (2019). Looking at psychology through the lens of metascience. APS Observer, 32(9).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Brown, N. [@sTeamTraen]. (2021, November 10). *It's 2021 and social psychologists are still publishing papers like this: Women wearing not much clothing on a night out don't get cold because they believe that they "look hot" (geddit?)* [Tweet]. Twitter. https://twitter.com/sTeamTraen/status/1458217774677954560

Crandall, C. S., Sherman, J. W. (2016). On the scientific superiority of conceptual replications for scientific progress. *Journal of Experimental Social Psychology, 66*, 93–99.

Derksen, M. (2019). Putting Popper to work. *Theory & Psychology*, *29*(4), 449-465.

Derksen, M., & Field, S. M. (2021). The tone debate: knowledge, self, and social order. *Review of General Psychology, 00(0),* 1-12.

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, 6, 621.

Elms, A. C. (1975). The crisis of confidence in social psychology. *American psychologist, 30*(10), 967.

Everett, J. A. C., & Earp, B. D. (2015). A tragedy of the (academic) commons: Interpreting the replication crisis in psychology as a social dilemma for early-career researchers. *Frontiers in Psychology, 6*, 1152.

Feest, U. (2019). Why replication is overrated. *Philosophy of Science, 86*(5), 895-905.

Flis, I. (2019). Psychologists psychologizing scientific psychology: An epistemological reading of the replication crisis. *Theory & Psychology*, *29*(2), 158-181.

Furtjes, J. (2021). *Where to go from here? Researcher perspectives on the replication crisis and the state of the academic world* [Unpublished undergraduate thesis]. University of Groningen.

Freese, J., & Peterson, D. (2017). Replication in social science. *Annual Review Sociology, 43*, 147–165.

Guttinger, S. (2020). The limits of replicability. *European Journal for Philosophy of Science*, *10*(2), 1-17.

Hales, A. H., Wesselmann, E. D., & Hilgard, J. (2019). Improving psychological science through transparency and openness: An overview. *Perspectives on Behavior Science, 42*(1), 13-31.

Hershler, A. (2021). *In crisis do we unite? Debates on the nature and impact of the credibility revolution.* [Unpublished undergraduate thesis]. University of Groningen.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS medicine*, *2*(8), e124.

Iso-Ahola, S. E. (2017). Reproducibility in psychological science: When do psychological phenomena exist? *Frontiers in Psychology, 8,* 879-916.

Leonelli, S. (2018, October). Rethinking reproducibility as a criterion for research quality. In *Including a symposium on Mary Morgan: curiosity, imagination, and surprise.* Emerald Publishing Limited.

Lynch Jr, J. G., Bradlow, E. T., Huber, J. C., & Lehmann, D. R. (2015). Reflections on the replication corner: In praise of conceptual replications. *International Journal of Research in Marketing, 32*(4), 333-342.

Morawski, J. (2019). The replication crisis: How might philosophy and theory of psychology be of use? *Journal of Theoretical and Philosophical Psychology*, *39*(4), 218-238.

Morawski, J. (2020). Psychologists' psychologies of psychologists in a time of crisis. *History of Psychology, 23*(2), 176-198.

Morawski, J. (2021). How to true psychology's objects. *Review of General Psychology. 0*(0), 1-15.

Mulkay, M., & Gilbert, G. N. (1981). Putting philosophy to work: Karl Popper's influence on scientific practice. *Philosophy of the Social Sciences, 11*(3), 389–407.

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*(1), 1-9.

Nicolai, R. (2021). *Replication crisis: The current view of early career researchers on the psychological research climate.* [Unpublished undergraduate thesis]. University of Groningen.

Nosek, B. A., & Bar-Anan, Y. (2012). “Scientific Utopia: I. Opening Scientific Communication.” *Psychological Inquiry*, *23*(3), 217–243.

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). “Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability.” *Perspectives on Psychological Science 7*(6), 615–631.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251).

Penders, B., Holbrook, J. B., & De Rijcke, S. (2019). Rinse and repeat: Understanding the value of replication across different ways of knowing. *Publications*, 7(3), 52.

Peterson, D., & Panofsky, A. (2020, August 4). Metascience as a scientific social movement. *SocArXiv.* DOI: 10.31235/osf.io/4dsqa

Pool, A. (2021). *You, Science and The Replication Crisis – A Thematic Analysis Attempting To Understand What Replication Means To Researchers.* [Unpublished undergraduate thesis]. University of Groningen.

Sales, S. B. H. (2021). *Researchers’ experiences of the replication debate and their well-being.* [Unpublished undergraduate thesis]. University of Groningen.

Schmidt, H. (2021). *ECRs stuck between two worlds.* [Unpublished undergraduate thesis]. University of Groningen.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*, 90-100.

Schwarzbach, N. R. (2021). *Motivations for and reservations about the use of open science practices: An interview study* [Unpublished master's thesis]. University of Groningen.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual review of psychology*, 69, 487-510.

Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science, 9*(1), 59-71.

Washburn, A. N., Hanson, B. E., Motyl, M., Skitka, L. J., Yantis, C., Wong, K. M., Sun, J., Prims, J. P., Mueller, A. B., Melton, Z. J., & Carsel, T. S. (2018). Why do some psychology researchers resist adopting proposed reforms to research practices? A description of researchers' rationales. *Advances in Methods and Practices in Psychological Science*, *1*(2), 166-173.

Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, *39*(4), 202.

Appendix A - Invitation/Reminder E-mails, Study Information Form, and the Informed Consent Form

Initial Email

Dear [title+ name],

We are contacting you, because we are doing a pilot study for a large-scale study about perceptions of the replication/credibility crisis and the 'reform movement'. In this context, social psychology is a field that is often **talked about**, but in our opinion, not **talked to** enough. We are curious how you, as a social psychologist, have experienced the crisis debate, the reform movement and the proposed changes. The results of this survey will facilitate a critical evaluation of the aims and accomplishments of the reform movement. Because this is a **pilot survey**, we are especially interested in your feedback about questions and repeatedly ask for that.

We hope that you would like to take your time to participate in our survey. Participation will take approximately 15 minutes. You can access the survey via this link:

https://rug.eu.qualtrics.com/jfe/form/SV_8quywigev6mhQa2

Your contribution would be greatly appreciated!

In the attachment of this email, you can find more information about the study. If you would like to be kept up to date about this research and its results, please send us an e-mail.

Kind regards,

Maria Bompa, Kaiti Evgeniou, Rafael Funke, Larissa Hoß, Colm Ó Fuartháin and Robert van Ark

Joyce Hoek, MSc

Nina Schwarzbach, MSc

Sarahanne Field, MSc

Merle Pittelkow, MSc

Dr. Rink Hoekstra

Prof. dr. Don van Ravenzwaaij

Faculty of Behavioural and Social Sciences, Rijksuniversiteit Groningen, the Netherlands.

Reminder Email #1

Dear [title+ name],

A week ago we contacted you because of our survey about “perceptions of the reform movement”, and we highly appreciate your participation. In case you did already fill out the survey: thank you very much! Please disregard this email. Unfortunately, we cannot remove you from our mailing list, since participation is anonymous.

In case you have not filled out the survey, we would kindly like to remind you that participation in our survey is still possible.

You can participate in the survey using the following link:

https://rug.eu.qualtrics.com/jfe/form/SV_8quywigev6mhQa2

In response to previously raised concerns:

- We invited 250 people to this pilot survey. Therefore, it would be difficult to trace back your identity on the basis of demographic data we ask for.
- If you’d like to give more detailed feedback verbally or via email, please do not hesitate to contact us.
- Some said that the survey takes longer than 15 minutes. Please take into consideration that it might take up to 30 minutes depending on how detailed your answers are.

Thank you in advance,

Robert van Ark, Maria Bompa, Kaiti Evgeniou, Colm Ó Fuartháin, Rafael Funke and Larissa Hoß

Research team:

Joyce Hoek, MSc

Nina Schwarzbach, MSc

Sarahanne Field, MSc

Merle Pittelkow, MSc

Dr. Rink Hoekstra

Prof. dr. Don van Ravenzwaaij

Faculty of Behavioural and Social Sciences, Rijksuniversiteit Groningen, the Netherlands

Reminder Email #2

Dear [title+ name],

We would like to remind you one last time about our survey about "perceptions of the reform movement". You still have time to fill it out until December 8th, after which the survey will close. Your participation is still highly appreciated!

In case you did already fill out the survey: thank you very much! Please disregard this email. Unfortunately, we cannot remove you from our mailing list, since participation is anonymous.

You can participate in the survey using the following link:

https://rug.eu.qualtrics.com/jfe/form/SV_8quywigev6mhQa2

In response to previously raised concerns:

- We invited 250 people to this pilot survey. Therefore, it would be difficult to trace back your identity on the basis of demographic data we ask for. In addition, we've decided not to publish the data of this pilot survey on OSF or any other open data platform.
- If you'd like to give more detailed feedback verbally or via email, please do not hesitate to contact us.
- Some said that the survey takes longer than 15 minutes. Please take into consideration that it might take up to 30 minutes depending on how detailed your answers are.

Thank you in advance,

Robert van Ark, Maria Bompa, Kaiti Evgeniou, Colm Ó Fuartháin, Rafael Funke and Larissa Hoß

Research team:

Joyce Hoek, MSc

Nina Schwarzbach, MSc

Sarahanne Field, MSc

Merle Pittelkow, MSc

Dr. Rink Hoekstra

Prof. dr. Don van Ravenzwaaij

Faculty of Behavioural and Social Sciences, Rijksuniversiteit Groningen, the Netherlands

Study Information Form

INFORMATION ABOUT THE RESEARCH

"Perspectives of the replication crisis, science and the reform movement"

Information about the study

Over the last decade, psychology has been experiencing what some people call a replication crisis. This crisis has been shocking for many people inside and outside the field of psychology. In order to counteract the challenges, a movement has emerged promoting replicable and open research practices. The movement has proposed practices, normative changes and policy changes. However, the movement has also received some criticism. Some practices and attitudes the new movement proposes seem to not fit with researchers' research, attitudes or working habits. But where does it clash? Because of these inconsistencies, it is important to ask psychology researchers in the fields affected by the crisis how they experience the crisis, the movement and science in general. The current study is a **pilot study**, which aims to facilitate a critical evaluation of the reform movement's aims and accomplishments.

Why do I receive this information?

The debate about the replication crisis is often dominated by metascience and open science researchers, and excludes the opinions of researchers outside of these movements. We would therefore like to hear your opinion because of your experience as a researcher in social psychology. By participating in this research, you will be able to share your perspective on the replication crisis debate and the proposed solutions.

What does it mean to participate in this study?

We would like to ask you to complete a brief questionnaire, which can be completed in about 15 minutes.

Do I have to participate in this research?

Participation in the research is voluntary. However, your consent is needed. Therefore, please read this information carefully. Ask all the questions you might have, for example because you do not understand something. Only afterwards you decide if you want to participate. If you decide not to participate, you do not need to explain why, and there will be no negative consequences for you. You have this right at all times, including after you have consented to participate in the research.

How will we treat your data?

Data will be processed completely anonymous. You will participate in this study by clicking on the Qualtrics link. We will not ask for your name or email address during the survey, so answers

will not be traceable to you. After data collection and analysis, the full dataset will be made public on OSF for re-use by other researchers.

What else do you need to know?

This pilot study will result in six bachelor theses. Furthermore, the analysis of the data may result in a publication in a scientific journal. The study is supervised by Joyce Hoek: PhD student at Behavioral and Social Sciences, University of Groningen.

This study has received ethics approval by the Ethics Committee of Psychology at University of Groningen (EC code:PSY-2122-S-0016). If you have any questions or concerns regarding your rights as a participant you may contact the committee at ecp@rug.nl

You may always ask questions about the study: now, during the study, and after the end of the study by contacting us at: perceptions.of.reform@rug.nl

Informed Consent Form

Welcome and thank you very much for participating in our survey. For more information about this pilot study, please refer to the Study information form or contact us at: perceptions.of.reform@rug.nl. The study will take approximately 15 minutes, contains 11 sections and is best completed on a computer. Please read the information below and indicate whether you agree with it before continuing with this survey. You have the right to take a screenshot of this information. I have read the information about the research. I have had the opportunity to ask questions about it. I understand what the research is about, what is being asked of me, which consequences participation can have, how my data will be handled, and what my rights as a participant are. I understand that participation in the research is voluntary.I myself choose to participate. I can stop participating at any moment. If I stop, I do not need to explain why. Stopping will have no negative consequences for me.I consent to participating in this study:

o Yes, I consent to participation.

o No, I do not consent to participation.

Appendix B - Survey

The survey used in this study begins on the next page.

Reform Movement Pilot Survey

Start of Block 0: Informed Consent

Welcome and thank you very much for participating in our survey. For more information about this pilot study, please refer to the Study information form or contact us at: perceptions.of.reform@rug.nl. The study will take approximately 15 minutes, contains 11 sections and is best completed on a computer. Please read the information below and indicate whether you agree with it before continuing with this survey. You have the right to take a screenshot of this information. I have read the information about the research. I have had the opportunity to ask questions about it. I understand what the research is about, what is being asked of me, which consequences participation can have, how my data will be handled, and what my rights as a participant are. I understand that participation in the research is voluntary. I myself choose to participate. I can stop participating at any moment. If I stop, I do not need to explain why. Stopping will have no negative consequences for me.
I consent to participating in this study:

- ○ Yes, I consent to participation.
- ○ No, I do not consent to participation.

End of Block 0: Informed Consent

Start of Block 1: Demographics

First, we'd like to ask you for some demographic data.

Q1. In what country are you currently working?

▼ Afghanistan ... Zimbabwe

Q2. What is your broad field of expertise?

- ☐ Social psychology
- ☐ Developmental psychology
- ☐ Industrial and organizational psychology/ work psychology
- ☐ Environmental psychology
- ☐ Experimental psychology
- ☐ Personality psychology
- ☐ Clinical (neuro) psychology
- ☐ Cognitive psychology
- ☐ Quantitative psychology
- ☐ Biological psychology
- ☐ Political psychology
- ☐ Other, namely: _____

Q3. What is your current job position?

- ☐ (Undergrad) student
- ☐ Research Assistant
- ☐ Junior researcher
- ☐ PhD student
- ☐ Postdoc
- ☐ Assistant professor/UD
- ☐ Associate Professor/UHD
- ☐ Full professor
- ☐ Other, namely: _____

Q4. How long have you been working in academia? (years)

_____

End of Block 1: Demographics

Start of Block 2: Terms

To have a consistent and shared understanding throughout the survey, we would like to clarify what the terms mean to us. Throughout the survey, you can always go back to these definitions using a pop-up button found at the bottom.

Direct replication: The attempt to conduct a study in a manner as close to the original as possible (the same population, methodology, and statistical analyses).

Conceptual replication: The attempt to test the same theoretical process or effect as an existing study, or understand boundary conditions of given phenomena, but that uses methods that vary in some way from the previous study.

Successful replication: When the replication study yields results which are sufficiently similar to the original study in terms of the strength of the effect and whether the effect goes in the same direction as the original. 'Sufficiently similar' varies, and is usually defined by the replicating author.

Open science: Open science aims to make science more transparent. Open science practices include among others: preregistration, registered reports, open data, open peer review, and open access publishing.

Metascience: The study of research itself, often with the aim of improving its practice. Meta-researchers study the scientific community and its actors, their methods and reporting, reproducibility, evaluation, behavior, and incentives.

Reform movement: There are many different words describing groups of people that are promoting change in science, including 'meta-science movement', 'open science movement' or 'reformer movement'. In the following we summarize people sharing concern with regards to improving science through either meta-scientific or transparent/open science practices as the 'reform movement'.

Q5. Optional: Do you have feedback on these definitions?

_____

_____

From now onwards, we will refer mostly to the reform movement. You can always go back to the definitions if you are unsure about the terms used in the survey.

End of Block 2: Terms

Start of Block 3: Reform movement

The next questions will be about how the aims of the reform movement resonate with you and your research practices.

Q6. Please indicate the extent to which you...

	Not at all	Completely
...identify with the reform movement		

Q7.
Do you agree with this statement: "I am part of the reform movement."?

- ○ Yes
- ○ No
- ○ Don't know

Q8. Optional: Do you have any thoughts with regard to your identification with the reform movement you'd like to add here?

_____

_____

_____

_____

Q9. Optional: Do you have feedback on the questions about identification with the reform movement?

_____

_____

_____

_____

_____

End of Block 3: Reform movement

Start of Block 4: Epistemology/Ontology

We would like to know more about how you think about science and knowledge in general. Please indicate how the following statements relate to your research.

Please indicate how the following statements relate to your research:

	Not at all — Completely
Q10. "For every phenomenon that I study, there are multiple valuable truths."	
Q11. "In my field of research, scientists can ultimately get to/reach the truth."	
Q12. "In my field of research, results depend on the perception of the researcher."	
Q13. "Science should be organized in such a way as to reduce scientists' biases."	
Q14. “In my field of research, the effects are dependent on the **time period** in which these studies took place rather than universal."	
Q15. “In my field of research, the effects are dependent on the **culture** where the study took place rather than universal."	
Q16. “In my field of research, the effects are dependent on the **experimental setup** rather than universal."	
Q17. "It is possible to specify all the boundary conditions that enable a theory to hold true."	
Q18. "Conducting a scientific study requires constant adaptation of the methods used."	
Q19. "The expertise of an individual scientist is important to study a phenomenon."	

Q20. Optional: Do you have any thoughts you’d like to add here?

_____

_____

_____

_____

_____

Q21. Optional: Do you have feedback on the questions about science and knowledge in general?

_____

_____

_____

_____

_____

End of Block 4: Epistemology/Ontology

Start of Block 5: Research Quality

The current survey includes some questions about the quality of research. First, we would like to know what you think of the current state of research quality in your field.

Please indicate the extent to which you agree with the following statement:

	Not at all	Completely
Q22. "I think that research quality in my field is something that needs to be improved."		

Q23. Optional: Can you elaborate?

_____

_____

_____

_____

End of Block 5: Research Quality

Start of Block 6: Replication

The next couple of questions will be about replication.

Please indicate the extent to which you agree with the following statement:

	Not at all — Completely	Not applicable
Q24. "New replication studies should attempt to **generalise** established effects."		☐
Q25. "New replication studies should attempt to **falsify** established effects."		☐
Q26. “New replication studies should attempt to **confirm** established effects."		☐
Q27. "Original researchers of a study should participate in the process of replication."		☐
Q28. "I believe it is important that **direct replications** are conducted in my field."		☐
Q29. "I believe it is important that **conceptual replications** are conducted in my field."		☐

We would now like to ask some questions about replication and research quality.

Please indicate the extent to which you agree with the following statement:

	Not at all — Completely	Not applicable
Q30. "I believe that successful **direct replications** are indicative of research quality in my field."		☐
Q31. "I believe that successful **conceptual replications** are indicative of research quality in my field."		☐

Can you elaborate on your previous two answers?

Q32. Why do you think that successful replication is, or is not, indicative of research quality in your field of research? Please indicate what type of replication you are talking about (i.e., direct, conceptual or any other form)?

_____

_____

_____

_____

_____

Q33. Optional: Which quality indicators other than replication do you think are important in your field of research?

_____

_____

_____

_____

_____

Q34. Optional: Do you have any thoughts you'd like to add here?

_____

_____

_____

_____

_____

Q35. Optional: Do you have feedback on the questions about replication?

_____

_____

_____

_____

_____

End of Block 6: Replication

Start of Block 7: Open Science Ideas

The next couple of questions are about your ideas of open science in general.

Please indicate the extent to which you agree to the following statements:

	Not at all	Completely
Q36. I think that science in general should be transparent and open if possible.		
Q37. Generally, I think that the more transparent and open the research process is, the higher its quality and reliability.		

Q38. Optional: Do you have any thoughts you'd like to add here?

_____

_____

_____

_____

_____

Q39. Optional: Do you have feedback on the questions about open science ideas?

_____

_____

_____

_____

_____

End of Block 7: Open Science Ideas

Start of Block 8: Open Science Practices

The next couple of questions are about your thoughts on the practical application of open science.

Q40. Please give an estimate on how many hours of (informal) training on open science practices you have received.

_____

Please indicate the extent to which you agree with the following statements:

	Very Little — Very Much	Not applicable
Q41. "I feel like I have received sufficient (informal) training on how to practice open science."		▢
Q42. "My working environment/colleagues encourage me to use open science methods to conduct my research."		▢

Q43. Which of the following practices are you currently using in your research?

	Never	Rarely	Sometimes	Mostly	Always	I don't know what this means	Not applicable
Preregistration	○	○	○	○	○	○	○
Registered reports	○	○	○	○	○	○	○
Open access publishing	○	○	○	○	○	○	○
Open data	○	○	○	○	○	○	○
Open materials (code, metadata)	○	○	○	○	○	○	○
Open peer review	○	○	○	○	○	○	○

Q44. Optional: Alternatively, which other open science practice are you currently using in your research?

	Never	Rarely	Sometimes	Mostly	Always
Other practice:	○	○	○	○	○

Q45. Which of the following practices would you like to use (more) in your future research?

- [ ] Preregistration
- [ ] Registered reports
- [ ] Open access publishing
- [ ] Open data
- [ ] Open materials (code, metadata)
- [ ] Open peer review
- [ ] Other, namely: _____
- [ ] None

Q46. Optional: What would you need to practice (open) science the way you'd like to?

_____

_____

_____

_____

_____

End of Block 8: Open Science Practices

Start of Block 9: Critique

From interviews, we gathered some information about how the reform movement is perceived. We will now like to know how much you agree with the next statements.

Q47. Please indicate the extent to which you agree with the following statement:

	Not at all	Completely	Not applicable
"I have the feeling that people in the reform movement understand the practices of my field."			▢

Q48. Optional: Please explain why (not)?

_____

_____

_____

_____

_____

Please indicate the extent to which you agree with the following statement:

	Not at all	Completely	Not applicable
Q49. "I feel like the reform movement addresses the most pressing issues regarding scientific quality in my field."			▢

Q50. Optional: Please explain why (not)?

_____

_____

_____

_____

Please indicate the extent to which you agree with the following statement:

	Not at all	Completely	Not applicable
Q51. “The proposed solutions solve the problems in my field sufficiently.”			☐

Q52. Optional: Please motivate your answer.

_____

_____

_____

_____

Q53. The reform movement prioritizes some solutions over others. Please rank how **you think** the reform movement prioritizes the following issues (1=most priority, 16=least priority):

_____ Preregistration/registered reports
_____ Data/code sharing
_____ Research methods other than inferential (qualitative, descriptive, exploratory)
_____ Improving statistics (bayesian statistics vs NHST etc)
_____ Theory or construct development
_____ Bigger sample sizes
_____ Slow science
_____ Managing competitive culture in academia
_____ More collaboration
_____ More direct replication
_____ More conceptual replication
_____ Increasing diversity within universities
_____ Increasing the importance of societal impact
_____ More freedom to pursue your scientific interests
_____ More job security
_____ Nuanced reporting of results

Q54. Are you sure you finalised the ranking?

- ○ Yes, I am
- ○ No, I am not

Q55. Optional: What problems with regard to the quality of research in your field is the movement missing?

_____

_____

_____

_____

_____

Q56. Optional: Do you have feedback on the questions about the priorities of the reform movement?

_____

_____

_____

_____

_____

End of Block 9: Critique

Start of Block 10: Important Issues To Be Addressed

Q57. In order to improve **research quality** in your field, multiple solutions are suggested. Please rank how important **you think** they are to improve research quality in your field (1=most important, 16=least important).

_____ More focus on preregistration/registered reports
_____ More focus on data/code sharing
_____ More focus on research methods other than inferential (qualitative, descriptive, exploratory)
_____ More focus on improving statistics (Bayesian statistics and/or NHST etc.)
_____ More focus on theory or construct development
_____ More focus on bigger sample sizes
_____ More focus on slow science
_____ More focus on managing competitive culture in academia
_____ More focus more collaboration
_____ More focus on direct replication
_____ More focus on conceptual replication
_____ Increasing diversity within universities
_____ Increasing the importance of societal impact
_____ More freedom to pursue your scientific interests
_____ More job security
_____ More focus on nuanced reporting of results

Q58. Are you sure you finalised the ranking?

- ◯ Yes, I am
- ◯ No, I am not

Q59. Optional: Did we forget something?

_____

_____

_____

_____

_____

Q60. Optional: Do you have feedback on the questions about the important issues to be addressed?

_____

_____

_____

_____

_____

End of Block 10: Important Issues To Be Addressed

Start of Block 11: Obstacles to Implementation

Researchers also report various obstacles to reforming science. How much do you agree with the following statements?

	Not at all — Completely	Not applicable
Q61. “Open science does not sufficiently take into account privacy issues for studies with sensitive data.”		▢
Q62. “Open sciences practices are too time-consuming.”		▢
Q63. “At this moment, open science practices are not rewarded or incentivised enough.”		▢
Q64. “Practicing open science gives me a competitive **advantage** over other scientists.”		▢
Q65. “Practicing open science gives me a competitive **disadvantage** over other scientists.”		▢
Q66. "The critique about my field of research from the reform movement makes me feel like I have to prove my innocence."		▢
Q67. "The tone of the members of the reform movement should be more nuanced."		▢
Q68. "I am less likely to engage with the propsed reform practices because I feel the reform movement is prejudiced toward my field of research."		▢

Q69. Optional: Do you want to elaborate on any of your answers with regard to obstacles for reform?

_____

_____

_____

_____

Q70. Optional: What other obstacles for changing the practices of your field do you see?

_____

_____

_____

_____

_____

End of Block 11: Obstacles to Implementation

Start of Block 12: Feedback

You've now reached the end of the survey.

Q71. Would you like to give more specific feedback on the survey?

_____

_____

_____

_____

_____

Q72. I have honestly answered the questions above.

- ○ Yes
- ○ No

Q73. I paid attention filling in this survey.

- ○ Yes
- ○ No

Please press → to submit your answers. You cannot change your answers anymore after submitting.

End of Block 12: Feedback

Appendix C - Correlation Coefficient Discussion

This appendix concerns itself with the figures and tables that correspond to the relevant assumption checks for Pearson's *r* and Spearman's rank-order correlations.

For Pearson's correlation, there are three key assumptions that must be met; 1) that a linear relationship exists between the two chosen variables, 2) that both variables are approximately normally distributed, and 3) that no significant outliers exist. To test these assumptions with our data in R, we 1) assessed scatterplots for linear relationships, 2) conducted Shapiro-Wilks tests and generated Q-Q plots to test normality, and 3) identified outliers from descriptive boxplots (see Figure 1). The data violated all three assumptions. First, the scatterplots failed to indicate any clear linear relationships between all relevant variables, as can be seen in Figure C.1. Secondly, all Shapiro-Wilks tests returned significant evidence to reject the null hypothesis, as seen in Table C.1, meaning that a normality assumption should be rejected. The violation of normality assumption was also reflected in the Q-Q plots (see Figure C.2). Finally, outliers were observed for item Q24, Q28, and Q31.

Spearman's rank-order correlation (Spearman's rho) is a commonly used correlation when the assumptions of Pearson's *r* are notably violated (Akoglu, 2018). To test the assumption of monotonicity for Spearman's rho, we inspected the scatterplots in Figure C.1. We determined visually that only the pair Q30|Q31 could meet the monotonicity assumption of Spearman's rho. We used JASP, a statistical software programme, to add a smooth regression line to the scatterplots in order to test the assumption confidently, which supported our visual test (see Figure C.2).

Figure C.1

Scatter plot matrix for all relevant quantitative survey items

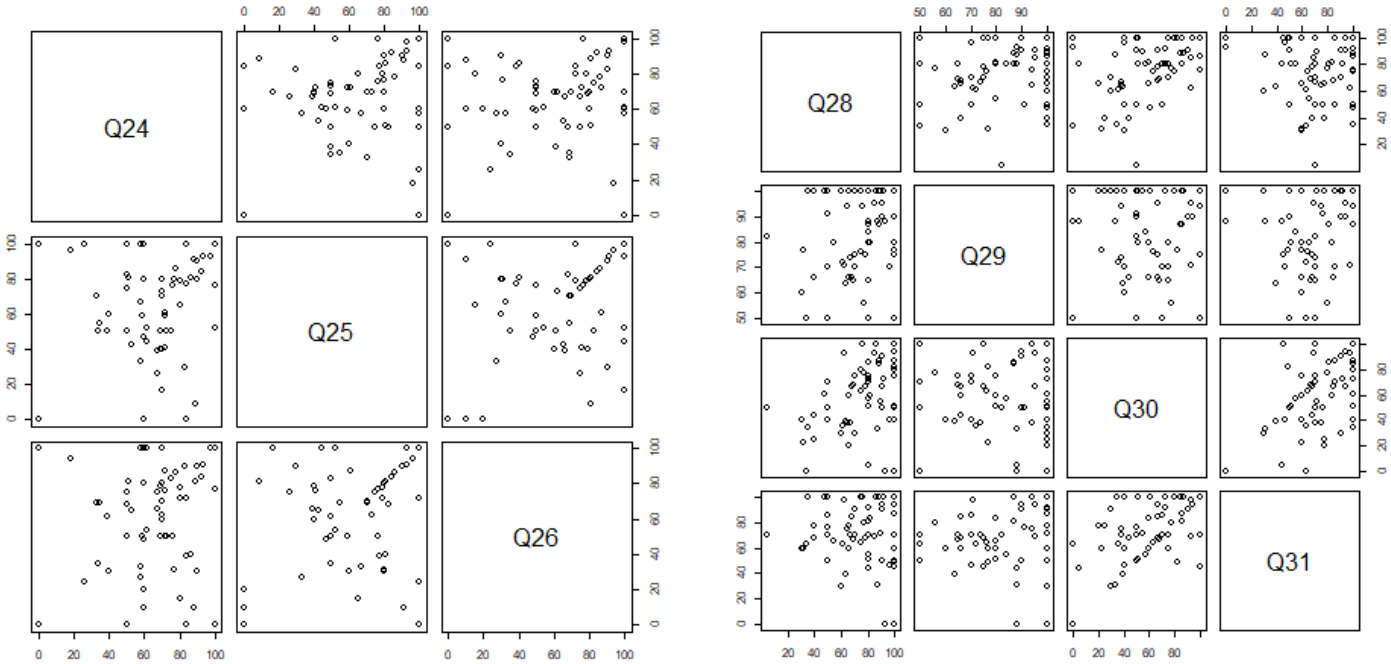


Table C.1

Results for the Shapiro-Wilks test for normality assumption in the chosen items

Item	Statistic (*W*)	p-value
Q24	0.95	0.00836
Q25	0.927	0.000595
Q26	0.93	0.000823
Q28	0.904	4.76E-05
Q29	0.866	1.96E-06
Q30	0.954	0.0154
Q31	0.909	0.000112

Figure C.2

QQ plots for all relevant quantitative survey items

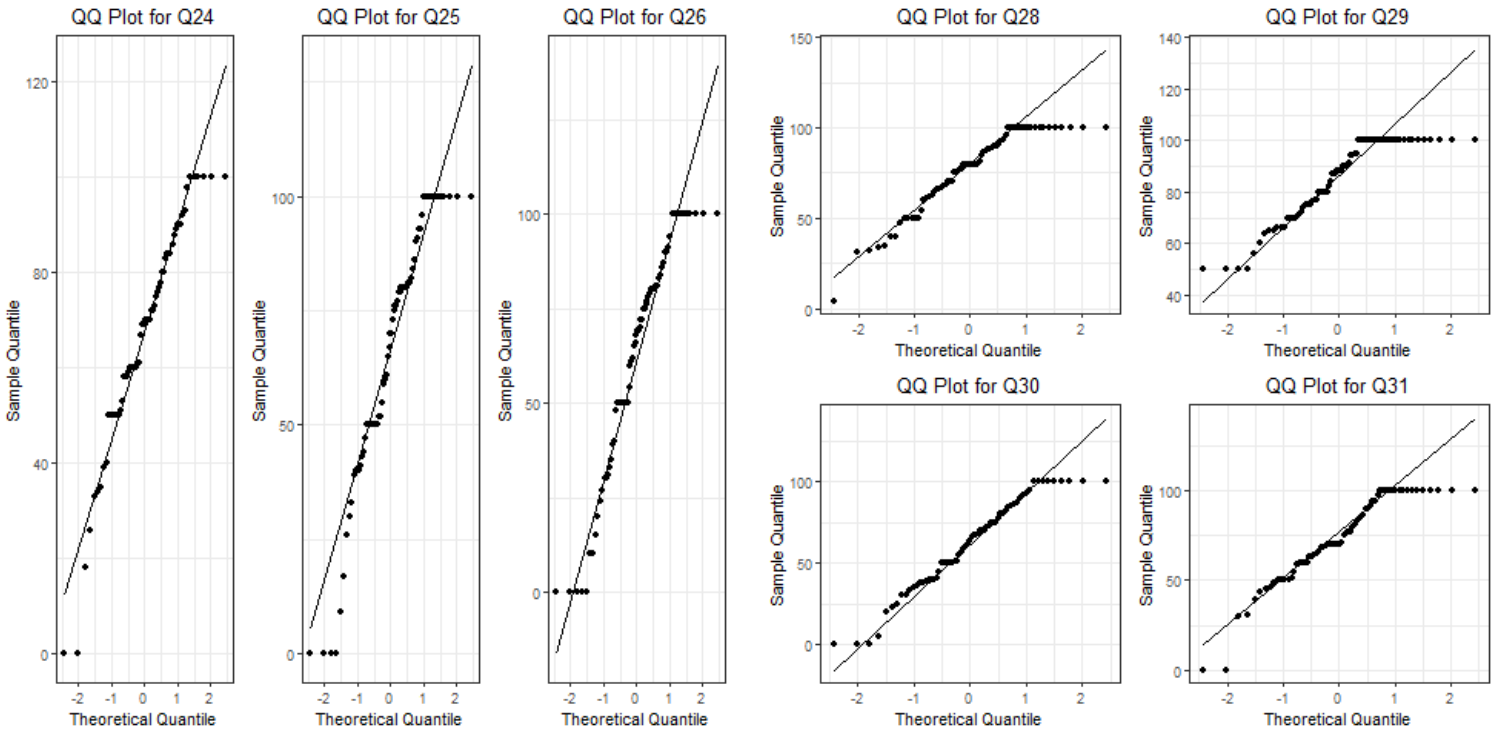
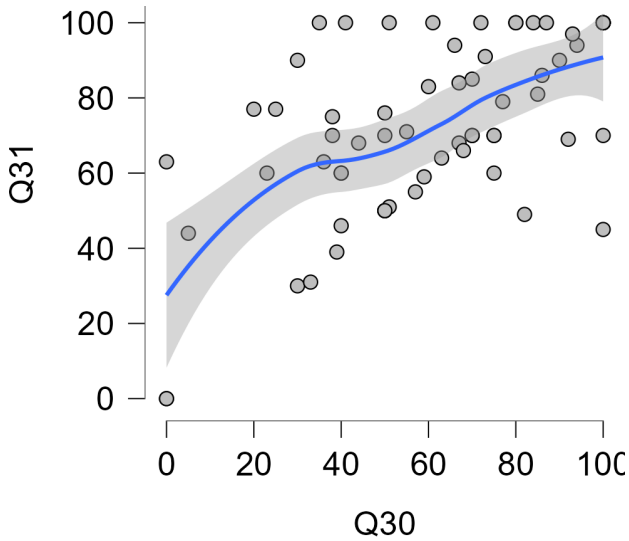


Figure C.3

Scatterplot for Q30-Q31 with regression line for the relevant three item pairs



Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish journal of emergency medicine*, *18*(3), 91-93.

Appendix D - R code for Quantitative Analysis

```
#install.packages("ggplot2")
#install.packages("tidyverse")
#install.packages("hrbrthemes")
#install.packages("viridis")
#install.packages("reshape2")
#install.packages("grid")
#install.packages("gridExtra")

library(ggplot2)
library(tidyverse)
library(hrbrthemes)
library(viridis)
library(reshape2)
library(grid)
library(gridExtra)

## Setting working directory and preparing data for analysis
setwd("/Users/colmofuarthain/R")
mydata = read.csv("RealData.csv", header = TRUE)
mydata2 = mydata[, c("fieldofexpertise","Q99_1", "Q99_5","Q99_6",
"Q99_7","Q99_8","Q110_1","Q110_5", "Q123", "Q124")]
mydata2 = dplyr::filter(mydata2, grepl('24', fieldofexpertise)) # exclude non-social
psychologists
mydata2 = dplyr::filter(mydata2, Q123 == "1" | Q123 == "" & Q124 == "1" | Q124 == "")
#Excluding possible dishonest responding and lack of attention
mydata2 = mydata2[, c("Q99_1", "Q99_5","Q99_6", "Q99_7","Q99_8","Q110_1","Q110_5")] #
removing exclusion columns
colnames(mydata2) = c('Q24','Q25','Q26','Q28','Q29','Q30','Q31')
mydata2 = as.data.frame(sapply(mydata2,as.numeric))


## Demographics
# Preparing Data
demogr = mydata[, c("fieldofexpertise","workingcountry", "currentjob",
"workinacademiayears", "Q123", "Q124")]
demogr = dplyr::filter(demogr, grepl('24', fieldofexpertise)) # exclude non-social
psychologists
demogr = dplyr::filter(demogr, Q123 == "1" | Q123 == "" & Q124 == "1" | Q124 == "")
#Excluding possible dishonest responding and lack of attention
```

```
demogr = demogr[, c("fieldofexpertise","workingcountry", "currentjob",
"workinacademiayears")] # removing exclusion columns

# In which country are participants working
workingcountry_NL = sum(demogr$workingcountry == '144') # Number of participants
working in the Netherlands
workingcountry_notNL = sum(demogr$workingcountry != '144') # Number of participants
working outside the Netherlands

# How many years have participants been in academia
yearsinacademia = (as.data.frame(sapply(demogr$workinacademiayears,as.numeric))) #
Putting the Data as numerical
median_yearsinacademia = apply(yearsinacademia, 2, median, na.rm = TRUE)
all_quartiles_yearsinacademia = apply(yearsinacademia, 2, quantile, na.rm = TRUE)
Quartiles_1_yearsinacademia = all_quartiles_yearsinacademia[2,]
Quartiles_3_yearsinacademia = all_quartiles_yearsinacademia[4,]

# What is their current profession
num_UndergradStudent = sum(demogr$currentjob == '4')
num_ResearchAssistant = sum(demogr$currentjob == '7')
num_JuniorResearcher = sum(demogr$currentjob == '8')
num_PhDStudent = sum(demogr$currentjob == '9')
num_Postdoc = sum(demogr$currentjob == '10')
num_AssistantProf_UD = sum(demogr$currentjob == '11')
num_AssociateProf_UHD = sum(demogr$currentjob == '12')
num_FullProf = sum(demogr$currentjob == '13')
num_Other = sum(demogr$currentjob == '14')
#CurrentJob_Counts = c(num_UndergradStudent, num_ResearchAssistant,
                      #num_JuniorResearcher, num_PhDStudent, num_Postdoc,
                     # num_AssistantProf_UD, num_AssociateProf_UHD, num_FullProf)


## Calculating medians, IQR, quartiles, and a minimum value

# Medians
d_Medians = apply(mydata2,2,median, na.rm = TRUE)

# IQRs, along with Q1 and Q3
d_IQRs = apply(mydata2,2,IQR, na.rm = TRUE)
all_quartiles = apply(mydata2,2, quantile, na.rm = TRUE)
Quartiles_1 = all_quartiles[2,]
Quartiles_3 = all_quartiles[4,]

# Minimum value
complete_mydata2_min29 = mydata2[complete.cases(mydata2),5]
complete_mydata2_min28 = mydata2[complete.cases(mydata2),4]
min(complete_mydata2_min29)
min(complete_mydata2_min28)
```

```
## Boxplot chart
mydata2_m = melt(mydata2) #Change data structure for function ggplot
mylabels = c("Q24. New replication studies should \n attempt to generalise
established effects.",
             "Q25. New replication studies should \n attempt to falsify established
effects.",
             "Q26. New replication studies should \n attempt to confirm established
effects.",
             "Q28. I believe it is important that direct \n replications are
conducted in my field.",
             "Q29. I believe it is important that conceptual \n replications are
conducted in my field.",
             "Q30. I believe that successful direct replications \n are indicative of
research quality in my field.",
             "Q31. I believe that successful conceptual replications \n are
indicative of research quality in my field.")
item_obs = c(sum(mydata2$Q24 > -1, na.rm = TRUE),
             sum(mydata2$Q25 > -1, na.rm = TRUE),
             sum(mydata2$Q26 > -1, na.rm = TRUE),
             sum(mydata2$Q28 > -1, na.rm = TRUE),
             sum(mydata2$Q29 > -1, na.rm = TRUE),
             sum(mydata2$Q30 > -1, na.rm = TRUE),
             sum(mydata2$Q31 > -1, na.rm = TRUE))
mylabels = paste(mylabels, "\n (N =", item_obs, ")") # Adding the number of
observations to the labels

ggplot(mydata2_m, aes(x = variable, y = value, fill = variable)) +
  geom_boxplot(alpha = 0.8, varwidth = TRUE, outlier.colour = "red", outlier.fill =
"red", outlier.size = 3) +
  stat_boxplot(geom ='errorbar') +
  geom_jitter(width = 0.08) +
  scale_fill_brewer(palette = "Dark2") +
  scale_x_discrete(labels = mylabels) +
  xlab("Questions") +
  ylab("Agreement score \n \n (0 = Not at All, 100 = Completely)") +
  xlab("") +
  theme(plot.margin = unit(c(0.5,0.5,0.5,2),"cm"), legend.position = "none",
axis.text.x = element_text(angle = 50, hjust=1, size = 10.5))

## Assumptions check for Pearson correlation coefficient
# Checking linearity assumption by means of scatterplots, visualised in a matrix

# Q24-26
png(file = "Linear Check Q24-26 scatterplot matrix.png")
pairs(~Q24+Q25+Q26, data = mydata2, main = "Scatter Plots, Q24-Q26")
dev.off()
```

```
# Q28-31
png(file = "Linear Check Q28-Q31 scatterplot matrix.png")
pairs(~Q28+Q29+Q30+Q31, data = mydata2, main = "Scatter Plots, Q28-Q31")
dev.off()


## Checking normality assumption by means of QQ plots and Shapiro-Wilks Test
# Shapiro-Wilks Test - when it is not significant, we can assume normality
Shap_Q24 = shapiro.test(mydata2$Q24)
Shap_Q25 = shapiro.test(mydata2$Q25)
Shap_Q26 = shapiro.test(mydata2$Q26)
Shap_Q28 = shapiro.test(mydata2$Q28)
Shap_Q29 = shapiro.test(mydata2$Q29)
Shap_Q30 = shapiro.test(mydata2$Q30)
Shap_Q31 = shapiro.test(mydata2$Q31)

# Q24-26 side by side QQ plots
qq24 = ggplot(mydata2, aes(sample = Q24)) + stat_qq() + ggtitle("QQ Plot for Q24") +
xlab("Theoretical Quantile") + ylab("Sample Quantile")+stat_qq_line() + theme_bw() +
theme(plot.title = element_text(hjust = 0.5))
qq25 = ggplot(mydata2, aes(sample = Q25)) + stat_qq() + ggtitle("QQ Plot for Q25") +
xlab("Theoretical Quantile") + ylab("Sample Quantile") + stat_qq_line() + theme_bw()
+ theme(plot.title = element_text(hjust = 0.5))
qq26 = ggplot(mydata2, aes(sample = Q26)) + stat_qq() + ggtitle("QQ Plot for Q26") +
xlab("Theoretical Quantile") + ylab("Sample Quantile") + stat_qq_line() + theme_bw()
+ theme(plot.title = element_text(hjust = 0.5))
png(file = "Normality Check Q24-26 matrix.png")
grid.arrange(qq24, qq25, qq26, ncol = 3, nrow = 1,top = textGrob("Normality Check
Q24-26",gp = gpar(fontsize=20,font=1)))
dev.off()

# Q28-31 side by side QQ plots
qq28 = ggplot(mydata2, aes(sample=Q28)) + stat_qq() + ggtitle("QQ Plot for Q28") +
xlab("Theoretical Quantile") + ylab("Sample Quantile") + stat_qq_line()+theme_bw() +
theme(plot.title = element_text(hjust = 0.5))
qq29 = ggplot(mydata2, aes(sample=Q29)) + stat_qq() + ggtitle("QQ Plot for Q29") +
xlab("Theoretical Quantile") + ylab("Sample Quantile") + stat_qq_line() + theme_bw()
+ theme(plot.title = element_text(hjust = 0.5))
qq30 = ggplot(mydata2, aes(sample=Q30)) + stat_qq() + ggtitle("QQ Plot for Q30") +
xlab("Theoretical Quantile") + ylab("Sample Quantile") + stat_qq_line() + theme_bw()
+ theme(plot.title = element_text(hjust = 0.5))
qq31 = ggplot(mydata2, aes(sample=Q31)) + stat_qq() + ggtitle("QQ Plot for Q31") +
xlab("Theoretical Quantile") + ylab("Sample Quantile") + stat_qq_line() + theme_bw()
+ theme(plot.title = element_text(hjust = 0.5))
png(file = "Normality Check Q28-Q31 matrix.png")
grid.arrange(qq28, qq29, qq30, qq31,ncol = 2, nrow = 2,top = textGrob("Normality
Check Q28-31",gp = gpar(fontsize=20,font=1)))
dev.off()
```

```
## Spearman's correlation coefficient

# Assumptions of Pearson's coefficient were not met. For the pair of variables
# that was thought to have a monotonic
# relationship, Spearman's correlation coefficient is computed.

# Q30 & Q31
mydata2_3031 <- mydata2[,c(6,7)]
mydata2_complete_3031 <- mydata2_3031[complete.cases(mydata2_3031),]
cor(mydata2_complete_3031, method = "spearman"

# Scatterplot for Q30 & Q31
png(file = "Q30-Q31 scatterplot.png")
plot(mydata2$Q30, mydata2$Q31, xlab = "Q30: Direct indicative of research quality" ,
     ylab = "Q31: Conceptual indicative of research quality")
dev.off()
```

Appendix E - Codebook including definitions and example quote

Table E.1

Codebook, definitions and exemplars used for the thematic analysis

Codes	Definition	Exemplars
Successful single replication is not conclusive	The success/failure of a single replication (both direct and conceptual) study should not tell us anything about the quality of the study.	*"...a successful or unsuccessful replication may not necessarily mean that a phenomenon is not true, but reveal more nuances to our understanding of what we study."*
Replication is a learning and quality process	Replication as a whole is a credible research quality method. Replication has important benefits for the field of social psychology that does not depend on the results of single studies, which themselves are not so beneficial and meaningful	*"It's not the results of the replication that matters. What matters is that we do them and learn from them."*
Incentives and bias for failed replications	The idea that the incentives for failed replications may outweigh successful replications for researchers, which may consciously or unconsciously affect their replication efforts	*"The underlying problem is that failed replications are seen as more newsworthy than successful replications, so that replicators can have more impact if their replication attempts fail."*
Replication should not have purposes	Here is the idea that when replication has goals, such as finding out if the outcome of a single study is or is not reproducible, it limits the productive output of this enterprise.	*"I believe replicators (or scientists in general) should not have such goals. The goal should be to establish whether a particular effect replicates, and the replicator should be open to all possible outcomes."*
Direct replication for reliability	Successful direct replication studies can say something about the reliability (a measure) of our results	*"Id say direct replication would be the first step, to ensure the reliability of the effect…."*
Direct replication for robustness	Successful direct replication studies can say something about the robustness (a characteristic) of our field and our theories.	*"Direct replication rate should be diagnostic of the robustness of findings published in a field."*
Direct replication indicative of quality of methodology	Successful direct replication studies can say something about the quality of the measures and methodology we use in social psychology	*"Direct replications also have their value and can indicate the stability of the work and the quality of research protocols (can someone replicate the work)."*
Direct replication is uninformative	Direct replication does not provide any new information and cannot prove anything, so it yields uninformative results	*"There are a lot of factors which might influence a direct replication to not be successful (...) a failed direct replication does not tell us all that much about the effect."*
Direct replication reinforces original biases and mistakes	As direct replication aims to follow the original study in as close a manner as possible, what ends up happening is that the mistakes and biases of the original researcher also end up in the new replication study, eliminating the possibility of improving the quality of our research.	*"Using the same materials/populations as before, as in direct replications, only further reinforces/empowers the potential biases involved in the original research."*
Direct replication not applicable in Social Psychology	Due to the context-sensitivity of the social sciences, the conditions between two different studies will always be different even if the methods are followed as closely as possible. Thus, direct replication cannot say anything about the research quality of our results, and can be considered not applicable in the field.	*"… there are cases when direct replication is difficult because of changed context or meaning…"*

Conceptual replication for generalizability	Conceptual replication increases the generalizability (for example, contexts, different populations, and different operationalizations) of phenomenon that have been studied	*"Conceptual replication is more important for generalizability than direct ones because it gives a sense that the way we study things can be applied to other contexts or samples or methods."*
Conceptual replication for validity	Conceptual replication for (construct) validity, in the sense that conceptual replication measures phenomena in different manners and thereby increases our confidence that we are effectively capturing the phenomenon we purport to.	*"In my opinion, conceptual replications are extremely important. Without them, it is difficult to be certain that the study examines the phenomenon of interest."*
Conceptual replication for theory boundary conditions	The success or failure of conceptual replication studies can inform us on the boundaries of our theories.	"*... conceptual replications can add important information on boundary conditions and extensions."*
Conceptual replication for theory development	Conceptual replication can be a more effective form of replication for building theories, which is considered a desired facet for science in social psychology.	*"For an effect to be meaningful it should be present in more than one study. Conceptual replications are thus important as they indicate in what context something is and is not present, which we can build on theoretically."*
Conceptual replication and context-sensitivity	Conceptual replication is more appropriate for social psychology due to the context sensitive nature of the field.	*"Especially conceptual replications are important as social psychological theories can be quite time, culture and context dependent."*
Conceptual replication overcomes methodological limitations and bias	By testing in a different manner than the original study, conceptual replication overcomes methodological limitations and bias in original studies.	*"Conceptual replications can overcome the methodological limitations or unique methodological features of the previous studies."*
Conceptual over direct replication	Responses that indicate that conceptual replication is distinctly more important than direct replication	*"Conceptual replications do the same [as direct replication] AND are indicative of whether the result is something generalizeable [sic] and not specific to the exact methods used in the original experiment."*
Both replication types are similarly important	To code when a response indicates that both replication types are similarly important.	*"I'd say direct replication would be the first step, to ensure the reliability of the effect, followed by conceptual replication for validity."*
Both replication types are uninformative	To code when a response indicates that both replication types, even if successful, do not provide any information or meaning.	*"Successful replications, whether direct or conceptual, can be meaningless if the original phenomenon/effect is not of theoretical value."*
Direct over conceptual replication	To code when a response indicates that direct replication is distinctly more important than conceptual replication.	*"Direct replication rate should be diagnostic of the robustness of findings published in a field. Conceptual replications are limited by the (typically) unclear correspondence in validity of measurements across studies purporting to test the same hypothesis."*
Nature of study determines which type of replication	When the nature of study determines which type of replication.	*"I think it depends heavily on the nature of the study. If we're talking about decision making processes, observing behavior in experiments, etc., direct replications have value. If we're talking about field experiments, conceptual replication might be more fruitful, as of, for example, cultural differences."*

Social psychology and context-sensitivity	The idea that social psychology as a field deals with context-sensitivity in terms of phenomena and dynamic systems.	*"Basically a complexity perspective confronts you with the possibility of fundamental uncertainty. Replication might be only possible in more stable situations of complex systems, and hence is not a good concept to study more turbulent stages in social systems."*
Objectivity and truth as foundations for science	The idea that in science objectivity and (single) truth are valuable and realistic cornerstones to hold	*"Replication is important, because if we cannot make any replicable observations, then it is very unlikely that any of our explanations captures parts of the truth."*
Research is subjective by nature	This code defines the idea that an ideal of an objective researcher and science is not realistic, and acknowledging that subjectivity of the scientists will always be contained within science and our findings	*"I think that there will always be an element of subjectivity in the kind of research we do ..."*
Universal and stable effects exist in science	The idea that there exists objects and effects in psychological science that are stable and universal.	*''I think there are phenomena in reality that have stable characteristics but that can never be described in any words."*