When Science Communication Gets Personal

Researchers' Self-Disclosure and Laypeople's Trust in Science

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie (Dr. phil.)
der Ludwig-Maximilians-Universität München

vorgelegt von
Marlene Sophie Altenmüller

aus **München**

2022

Erstgutachter: Prof. Dr. Mario Gollwitzer
Zweitgutachter: Prof. Dr. Rainer Bromme
Tag der mündlichen Prüfung: 12.07.2022

Table of Contents

1 Abstract 1
2 Deutsche Zusammenfassung 2
3 Introduction 8
3.1 Science reception among the lay public 9
3.1.1 Stereotypes about scientists 10
3.2 Trust in science 12
3.2.1 The state of trust in science 13
3.2.2 Researchers' epistemic trustworthiness 14
3.2.3 Credibility of scientific findings 17
3.2.4 Individualized and generalized trust in science 19
3.2.5 Domain- and topic-specificity of trust in science 21
3.2.6 Motivated science reception 23
3.3 Self-disclosure in science communication 24
3.3.1 Self-disclosure 24
3.3.2 Benefits of self-disclosure 26
3.3.3 Challenges for self-disclosure 28
3.4 The present research program 30
4 Summaries of manuscripts 33
4.1 Summary of Manuscript 1 33
4.2 Summary of Manuscript 2 35
4.3 Summary of Manuscript 3 37
5 Manuscript 1: Self-disclosure in science communication 39
6 Manuscript 2: Self-criticism and reform intentions in science 93
7 Manuscript 3: When research is me-search 127

8 Final Discussion 166
8.1 Open Questions and Limitations 168
8.1.1 Effect sizes and heterogeneity 168
8.1.2 Specificity of self-disclosure 169
8.1.3 Limited generalizability 171
8.2 Conceptualization of trust in science 173
8.2.1 Conceptualization of researchers' trustworthiness 173
8.2.2 Conceptualization of credibility 175
8.2.3 Relationship of trustworthiness and credibility 177
8.3 Directions for future research 180
8.3.1 Trust versus distrust 181
8.3.2 Trust in and trust within science 182
8.3.3 (Personalized) narratives in science communication 184
9 Conclusion 187
10 References 190
11 Danksagung 210

1 Abstract

Laypeople's trust in science might be obstructed by their stereotypical views of researchers as highly competent, yet only moderately warm. To counter this perception of lacking warmth, researchers could engage in self-disclosing science communication (e.g., revealing basic personal information, thoughts about one's own work, or one's personal involvement with the research topic), potentially promoting their affective trustworthiness, and, thus, laypeople's trust in science. However, such self-disclosure could also have adverse effects by violating expectations of appropriateness, professionalism and objectivity. This dissertation draws on theory about stereotypes, trust in science, and self-disclosure, and applies it to science communication. It includes three manuscripts and eleven empirical studies demonstrating small and ambivalent effects of researchers' self-disclosure on laypeople's trust in science, it explores potential boundaries of these effects (e.g., self-disclosure content and recipients' attitudes), and it provides important insights into the conceptualization of trust in science. I discuss limitations as well as theoretical and practical implications of this research program and, finally, I highlight further directions for future research.

Keywords: trust in science, trustworthiness, credibility, science communication, self-disclosure.

2 Deutsche Zusammenfassung

Es hat keine weltweite Pandemie benötigt, um die Wissenschaft ins Zentrum des öffentlichen Interesses zu rücken – aber es hat sicherlich geholfen. Unsere Welt ist geprägt von komplexen, globalen Herausforderungen, sei es die anhaltende Coronapandemie oder die drohende Klimakrise. Um diesen zu begegnen, braucht es eine effektive Partnerschaft zwischen der Öffentlichkeit und der Wissenschaft. Solch eine Kooperation kann nur durch vertrauensvolle, effektive Kommunikation gelingen. Das bedeutet, dass es nicht nur um das „was", sondern auch um das „wer" und „wie" in der Wissenschaftskommunikation geht: Die involvierten sozialen Prozesse müssen in den Blick genommen werden. Lai:innen denken jedoch vermutlich im Alltag wenig über Wissenschaft und Forschende nach, obwohl sie in unserer heutigen komplexen Welt bei nahezu allen Entscheidungen, egal ob alltäglich oder außergewöhnlich, von der Expertise anderer abhängig sind. Diese kognitive Arbeitsteilung (z.B. Bromme & Thomm, 2016; Keil, 2005; Keil et al., 2008) führt dazu, dass wir anderen – insbesondere Forschenden – vertrauen müssen, dass sie uns mit verlässlichem Wissen versorgen können.

Das öffentliche Vertrauen in Wissenschaft kann differenziert werden in die personenbezogene *epistemische Vertrauenswürdigkeit*, die Lai:innen Forschenden zuschreiben, und in die sachbezogene *Glaubwürdigkeit*, die sie wissenschaftlichen Befunden zuschreiben (z.B. Hendriks et al., 2016; Sperber et al., 2010). Epistemische Vertrauenswürdigkeit beschreibt, dass Forschende als verlässliche Wissensquellen wahrgenommen werden (z.B. Hardwig, 1991; Hendriks et al., 2015, 2016; Origgi, 2004; Wilholt, 2013). Sie kann unterteilt werden in Expertise (d.h. die notwendige fachliche Kompetenz) und Integrität und Benevolenz (d.h. die

aufrichtige Befolgung professioneller Regeln und Standards sowie die Wohlgesonnenheit gegenüber der Öffentlichkeit). Vertrauenswürdigkeit hat demnach sowohl kompetenzbezogene, kognitiv-rationale Anteile (d.h. Expertise), als auch wärmebezogene, sozio-affektive Anteile (d.h. Integrität und Benevolenz; Hendriks et al., 2015, 2016; Mayer et al., 1995; McAllister, 1995; Morrow et al., 2004). Die Glaubwürdigkeit wissenschaftlicher Befunde bezieht sich im Gegensatz zur personenbezogenen Vertrauenswürdigkeit auf die Überzeugung, dass ein wissenschaftlicher Befund „wahr" ist (z.B. Appelman & Sundar, 2016; Metzger et al., 2003; Rieh & Danielson, 2007; Wathen & Burkell, 2002). Dementsprechend kann sie unterteilt werden in die kognitive Facette der Akzeptanz („Ich halte den Befund für wahr") und die behaviorale Facette der Verhaltensintention („Ich beabsichtige, mich dem Befund entsprechend zu verhalten").

Obwohl das öffentliche Vertrauen in Wissenschaft generell recht stabil und hoch ist (Funk, 2017; Hendriks et al., 2016; Krause et al., 2019; Wissenschaft im Dialog/Kantar, 2021), wird es möglicherweise durch Stereotype über Forschende beeinträchtigt. Viele Lai:innen scheinen Forschende als hoch kompetent, jedoch nur mittelmäßig warm wahrzunehmen (Fiske & Dupree, 2014; Imhoff, Koch, et al., 2018; Rosman & Merk, 2021) – sozusagen als sehr intelligente, jedoch distanzierte "Nerds im Elfenbeinturm." Um diesem Stereotyp mangelnder Wärme zu begegnen, könnten Forschende sich in der Wissenschaftskommunikation an die Öffentlichkeit als warm und sympathisch, also *affektiv* vertrauenswürdig, darstellen. Eine Möglichkeit dies zu tun, besteht darin persönliche Informationen preiszugeben (z.B. Details aus ihrem Leben, selbstkritische Gedanken über ihre Arbeit oder ihre persönliche Involviertheit mit ihrem Forschungsthema). Doch solch eine Selbstoffenbarung kann auch problematisch sein, da sie den

Erwartungen an typische Wissenschaftskommunikation in Bezug auf Angemessenheit, Professionalität und Objektivität widersprechen könnte.

In dieser Dissertation stelle ich drei Manuskripte mit elf empirischen Studien vor, die sich mit der potentiell ambivalenten Wirkung von Selbstoffenbarung durch Forschende in der Wissenschaftskommunikation auf das Vertrauen von Lai:innen in Wissenschaft beschäftigen. Dabei verfolgt das Forschungsprogramm drei übergeordnete Ziele: Erstens wird der generelle Effekt von selbstoffenbarender Wissenschaftskommunikation auf die Wahrnehmung von epistemischer Vertrauenswürdigkeit und Glaubwürdigkeit wissenschaftlicher Befunde überprüft. Zweitens sollen mögliche Grenzen dieses Effekts ergründet werden. Daher werden zum Beispiel verschiedene Inhalte der Selbstoffenbarung untersucht. Manuskript 1 befasst sich mit der Preisgabe rein illustrativer persönlicher Informationen, die nicht direkt mit dem kommunizierten Forschungsprozess in Verbindung stehen und somit sehr grundlegend den Effekt von Selbstoffenbarung abbilden. Manuskript 2 erforscht die Preisgabe von persönlichen Gedanken in Bezug auf die eigene Arbeit (Selbstkritik und Reformabsichten), die also direkter mit der kommunizierten Wissenschaft verbunden sind. Manuskript 3 widmet sich schließlich der Preisgabe von Informationen, die untrennbar mit der kommunizierten Wissenschaft verwoben scheinen: der Offenlegung persönlicher Betroffenheit vom eigenen Forschungsthema („Me-Search"). Zudem berücksichtigt Manuskript 3 die mögliche Abhängigkeit des Effekts von Selbstoffenbarung von Voreinstellungen der Rezipient:innen (motivierte Wissenschaftsrezeption). Drittens verfolgt das Forschungsprogramm das Ziel, empirische Einblicke in die theoretische Konzeptualisierung von Vertrauen in Wissenschaft zu erhalten. Dazu werden die

Befunde reflektiert im Hinblick auf die psychologische Struktur von epistemischer Vertrauenswürdigkeit und Glaubwürdigkeit wissenschaftlicher Befunde sowie ihre Beziehung zueinander.

Die vorliegenden Ergebnisse demonstrieren ambivalente Effekte von Selbstoffenbarung in der Wissenschaftskommunikation: Manuskript 1 zeigt, dass die Preisgabe illustrativer persönlicher Informationen zu kleinen „Trade-Off"-Effekten führen kann: Während selbstoffenbarende Forschende als etwas affektiv vertrauenswürdiger wahrgenommen wurden als nicht selbstoffenbarende Forschende, sank ihre wahrgenommene Expertise. Die Glaubwürdigkeit der Befunde blieb hingegen insgesamt unbeeinflusst und die Ergebnisse einer anschließenden Feldstudie in einem Wissenschaftsmuseum stellten weiterhin in Frage, ob diese Art der Selbstoffenbarung praktische Relevanz in der Wissenschaftskommunikation haben kann. Manuskript 2 zeigt, dass die offene Preisgabe einer selbstkritischen Haltung gegenüber der eigenen Forschung dazu führte, dass Lai:innen die Forschenden positiver beurteilten, als wenn diese sich nicht selbstkritisch äußerten: Ihnen wurde mehr Vertrauenswürdigkeit und ihren Befunden mehr Glaubwürdigkeit zugeschrieben. Manuskript 3 zeigt schließlich, dass die Offenlegung von „Me-Search" sowohl positiv als auch negativ wahrgenommen werden kann, abhängig von bereits bestehenden Einstellungen der Laienrezipient:innen zum Forschungsthema: Hatten die Lai:innen positive Voreinstellungen, wurden die „Me-Searcher" als vertrauenswürdiger und ihre Befunde als glaubwürdiger beurteilt; hatten sie negative Voreinstellungen drehte sich der Effekt um.

Insgesamt scheint Selbstoffenbarung in der Wissenschaftskommunikation allein kein wirksames Mittel zu sein, um stereotype Wahrnehmungen von

Forschenden hinsichtlich vermeintlich mangelnder Wärme zu überwinden. Selbstoffenbarung hatte lediglich kleine und uneinheitliche Effekte auf das öffentliche Vertrauen in Wissenschaft. Das deutet darauf hin, dass es viel eher auf den konkreten Inhalt der offenbarten Information ankommt. So scheint es, dass persönliche Informationen positive Effekte haben, wenn sie beispielsweise „gute" wissenschaftliche Praxis implizieren (z.B. eine selbstkorrektive Haltung) oder wenn sie dazu genutzt werden können, Wissenschaft im Einklang mit bestehenden Einstellungen aufzuwerten (z.B. indem persönlich betroffenen Forschenden besondere Expertise und Motivation zugeschrieben wird). Genauso können die preisgegebenen Informationen aber auch zu negativen Effekten führen, wenn sie beispielsweise „schlechte" wissenschaftliche Praxis implizieren (z.B. mangelnde Selbstkritik) oder zur Abwertung von Wissenschaft genutzt werden können (z.B. indem einem „Me-Searcher" Voreingenommenheit unterstellt wird).

Insgesamt integriert diese Dissertation Theorie und Forschung zu Stereotypen, Vertrauen in Wissenschaft und Selbstoffenbarung und setzt sie in den Kontext der Wissenschaftskommunikation. Die wichtigsten theoretischen Beiträge dieses Forschungsprogramms liegen darin, Grenzen der positiven Effekte von Selbstoffenbarung, insbesondere auf Vertrauen, aufzuzeigen und wichtige Einblicke in die Differenzierung, zugrundeliegende Struktur und Beziehung der verschiedenen Konzepte von Vertrauen in Wissenschaft zu liefern. Die wichtigste praktische Implikation ist, dass Selbstoffenbarung in der Wissenschaftskommunikation allein wenig dazu beiträgt, das öffentliche Vertrauen in Wissenschaft zu stärken. Viel eher wird deutlich, dass Forschende gut abwägen sollten, was sie wie gegenüber welcher Zielgruppe preisgeben. Insgesamt zeigt das Forschungsprogramm viele Ansatzpunkte für zukünftige Forschung in Bezug

auf die Theorie und Praxis von Vertrauen, Wissenschaftskommunikation und auch Selbstoffenbarung auf. Es lohnt sich, die Grundlagen- und Anwendungsforschung zu diesen Themen zu erweitern und zu intensivieren. In dieser Dissertation werden drei vielversprechende Richtungen für zukünftige Forschung besonders hervorgehoben: die klarere Differenzierung von Vertrauen und Misstrauen („Distrust") in Wissenschaft, die Erweiterung des Blicks auf Vertrauen *innerhalb* von Wissenschaft und die Wirkung personalisierter Narrative in der Wissenschaftskommunikation.

Es ist wahrscheinlich, dass die Bewältigung der großen, globalen Herausforderungen der Zukunft maßgeblich von Vertrauen in Wissenschaft und effektiver Wissenschaftskommunikation abhängt. Als Psycholog:innen haben wir die Chance, durch die Erforschung dieser Themenfelder beizutragen.

3 Introduction

Science did not need a global pandemic to demonstrate its worth to the public – but it certainly helped. In light of the first skyrocketing waves of COVID-19 cases, the scientific community rose to the challenge. Quickly, scientists cooperated to develop new vaccines in record time, ran simulation scenarios predicting the virus's mutation and spread, and researched human experiences and behavior in this pandemic. For many, this not only meant intensifying or adapting their known work routines, but growing into a new role as science communicator: Experts from diverse disciplines engaged in outstanding communication efforts to explain the science behind the pandemic and advise politics as well as the public on how to manage this global crisis. Presently, research on science communication hardly gets by without drawing on the COVID-context. However, even before the world's eyes were all turned towards science to end this momentary pandemic situation and long after, effective science communication was and will be crucial. There will be many more global challenges, likely first and foremost combatting the dooming climate crisis, that depend on an effective partnership between society, politics, and science.

So, how do we get there? As in any partnership, communication is key. Yet, as will be elaborated in the following, interpersonal reservations fueled by science stereotypes need to be overcome to reach a trustful exchange of specialized knowledge, so that global as well as local challenges can be met. This requires paying attention to the social processes shaping the relationship between science and the lay public. In other words, it means focusing more on the "who" and "how" than on the "what" in science communication.

This dissertation will outline how laypeople's trust in science is determined not (only) by factual characteristics of scientific content or their underlying "scientific method" but rather by social characteristics of laypeople's interactions with researchers and their findings (cf., Oreskes, 2021). More specifically, I will focus on laypeople's perceptions of researchers revealing personal details, self-criticism and reform intentions, and personal motivations when engaging with the public. To this end, I will present three articles including eleven empirical studies on the effects of such self-disclosure in science communication on trust in science.

3.1 Science reception among the lay public

To understand trust in science, it is necessary to first take a step back and consider how laypeople (i.e., non-scientists and people lacking expert knowledge on the topic in question) perceive and experience science more broadly. In our highly complex world, our understanding of specific (scientific) issues is necessarily limited (i.e., the public's *bounded understanding* of science; Bromme & Goldman, 2014) and we thus need to rely on other people's specialized expertise (i.e., *division of cognitive labor*; e.g., Bromme & Thomm, 2016; Keil, 2005; Keil et al., 2008). For example, consider a week day's morning: When having breakfast and washing up, we trust that consumer items such as food or cosmetics are safe based on assessments by scientists from chemistry, biology, and medicine; when checking the news on our smartphones, we do not know specifics about the underlying technology; and when passing through a tunnel on our way to work, we are confident it will not collapse over our heads because geologists and engineers planned the construction fit for the ground. For the most part, we do not even realize there is science in all aspects of our lives. More likely, non-scientists only think about "science" when directly confronted with researchers and their findings

– typically in a media context (e.g., social media, news media, entertainment media) but also, yet less often, in more specialized science formats (e.g., science museums, science fairs or science events).

Most common science reception situations are characterized by little context or background information on the communicating researchers. For example, in traditional unilateral science communication (e.g., an interview in the news media) usually only the researcher's name and profession will be stated and the audience has no way of asking questions. Thus, in such information-sparse situations, laypeople likely rely on automatic perception processes like categorization and accompanying generalized expectations (i.e., stereotypes) to form an impression about a researcher (e.g., Cao & Banaji, 2016; Devine, 1989; Rubinstein et al., 2018).

3.1.1 Stereotypes about scientists

In social situations, people tend to automatically perceive and categorize others (and themselves) as described by Social Identity Theory (Tajfel, 1982; Tajfel & Turner, 1979) and its extension, Self-Categorization Theory (Turner et al., 1987): When a social category is salient, a set of generalized associations with and expectations of "typical" members of this category (i.e., stereotypes) is activated. For example, when the evening news channel airs an interview with a scientist, laypeople might immediately and automatically categorize the expert as "scientist" and themselves as "non-scientist". Thus, with no further individuating information, laypeople's stereotypical views about "scientists" are activated and affect their perceptions of the interviewed researcher and their communicative content.

While stereotypes about scientists have been becoming more favorable and balanced (e.g., Losh, 2010; Miller et al., 2018) over the last decades, laypeople

might still expect a researcher to be an older man wearing a lab coat who is highly skilled and very intelligent but nerdy, socially inept, and peculiar, maybe even immoral or dangerous (e.g., Losh, 2010; Miller et al., 2018; Rosman & Merk, 2021; Rutjens & Heine, 2016). The Stereotype Content Model (Fiske et al., 2002, 2007), a two-dimensional framework for person perception, helps summarizing these stereotypical views in broader terms: Groups associated with the category of "scientists" are generally perceived as highly competent, yet only moderately warm (i.e., "researchers", "scientists", "professors"; Fiske & Dupree, 2014). Similarly, extended models of person perception such as the ABC Model – a three-dimensional model differentiating between the dimensions agency/competence, beliefs/ideology, and communion/warmth (Abele-Brehm et al., 2021; Koch et al., 2016, 2020, 2021) – mirror these views regarding subgroups of the "scientist" category (i.e., "medical scientists", "physicists", "market research analysts"; Imhoff et al., 2018).

As person judgments, laypeople's stereotypical views of scientists are closely related to their trust in researchers and their findings. Yet, psychologically, they are not the same and it is necessary to distinguish between stereotypical perceptions of scientists and trust in science (i.e., researcher's trustworthiness and the credibility of their findings). In fact, the Stereotype Content Model, as well as the ABC Model, assume "trustworthiness" as part of the communion/warmth dimension of person perception (Abele-Brehm et al., 2021; Fiske et al., 2002, 2007; Koch et al., 2016, 2020, 2021). However, by doing so, these models disregard aspects of trustworthiness related to competence/agency perceptions that are central for trust in science, as outlined in the following.

3.2 Trust in science

As illustrated above, due to a highly complex and specialized world, laypeople need to rely on others' expertise. This reliance entails trust that these others are, in fact, reliable sources of knowledge and that the information gained from them is valid. In this context, "to trust" thus indicates a willingness to be vulnerable and a risk to accept biased or even incorrect information as "true" which could result in mistakenly adapting one's knowledge and behavior in line with that information (Hendriks et al., 2016; Resnik, 2011; see also Mayer et al., 1995). To minimize the risk of being wrongly informed, laypeople need to vigilantly assess new knowledge and the sources of that knowledge (Origgi, 2012; Sperber et al., 2010). Regarding this assessment, we can differentiate between first-hand ("What is true?") and second-hand ("Who to trust to tell me what is true?") evaluations of knowledge (Bromme et al., 2015; Bromme & Goldman, 2014). Given laypeople's bounded understanding of science and their need to rely on other people's expertise, second-hand evaluations (i.e., perceptions of researchers as reliable sources of knowledge) are likely more decisive for trust in science than first-hand evaluations (i.e., the direct assessments of scientific evidence). In line with this differentiation, the present conceptualization of trust in science distinguishes between laypeople's person-related perceptions of *researchers' epistemic trustworthiness* and their content-related ascriptions of *credibility of researchers' findings*. Additionally, I will detail in the following how trust in science can be determined as either *individualized* trust in specific researchers and their findings or as *generalized* trust in science as the collective of scientists, the social system of knowledge generation, or even the body of scientific knowledge (cf., Hendriks et al., 2016). Then, I will take on a more fine-

grained view considering *topic- and domain-specific trust* in science and how the phenomenon of *motivated science reception* might explain heterogeneity in trust in science. First, however, I will outline the current state of trust in science to demonstrate the relevance of studying trust in science.

3.2.1 The state of trust in science

Why is it relevant to study trust in science? Many people assume that we are currently in a crisis of trust in science (cf., Allea, 2018; Funk, 2017). However, there is little evidence for that notion. The overwhelming majority of studies finds that trust in science has been more stable over the years than trust in other public institutions and ranks higher than trust in most other societal groups or institutions such as the government or the media (e.g., Funk, 2017; Hendriks et al., 2016; Krause et al., 2019). Hendriks and colleagues (2016) provide a comprehensive overview of representative surveys from the United States and several European countries, demonstrating that "by large, the public seems to be very positive about the benefits that Science has to offer society, and […] the public mostly trust scientists to produce reliable knowledge of good quality, not biased and adhering to scientific principles." (p. 149, Hendriks et al., 2016).

Interestingly, there were some notable peaks in trust in science during the first year of the COVID-19 pandemic (e.g., Bromme et al., 2022; Jensen et al., 2021; Pew Research Center, 2022; M. S. Schäfer et al., 2021; Wissenschaft im Dialog/Kantar, 2021). However, recent surveys suggest that these pandemic-related peaks might be slowly regressing back in the direction of the long-standing means; but, as matters stand, they are still continuing on a level of higher public trust in science than before the pandemic (e.g., Bromme et al., 2022; M. S. Schäfer et al., 2021; Wissenschaft im Dialog/Kantar, 2021).

Focusing on public trust in science in Germany, the most important survey is the German Science Barometer, a yearly representative survey on the public's trust in science. Their most recent survey from September 2021 (Wissenschaft im Dialog/Kantar, 2021) shows that 61% of interviewed Germans agreed that they (somewhat or fully) trust science; among younger respondents, trust was even higher (under age 40: 78-80%). However, despite these findings suggesting stable and high trust in science, skeptical voices remain. For example, 39% of interviewed Germans agreed that scientists hold back information about the coronavirus and 26% agreed that they exaggerate the pandemic situation and, regarding their general trust in science, still 32% of respondents were undecided and 6% even did (rather) not trust science at all. This demonstrates that public trust in science cannot be taken for granted but needs constant building and maintenance. For this, research directed at understanding trust in science is necessary. So, how can we psychologically differentiate and study trust in science?

3.2.2 Researchers' epistemic trustworthiness

Epistemic trust is trust that is placed in someone as a reliable source of knowledge; epistemic trustworthiness is thus a characteristic ascribed to the trustee (i.e., a researcher) by the trustor (i.e., a layperson; Hardwig, 1991; Hendriks et al., 2015, 2016; Origgi, 2004; Wilholt, 2013). Theory and research generally conclude that epistemic trustworthiness is composed of different facets. However, there is a wide variety of labels and content such as experience, reliability, responsibility, attractiveness, dynamism, goodwill, or social standing, to name a few (Hendriks et al., 2015; Pornpitakpan, 2004; Resnik, 2011; Wilholt, 2013). Summarizing these various facets, systematic approaches often arrive at three distinct, yet related dimensions, going back as far as Greek philosopher

Aristoteles (e.g., Hendriks et al., 2015, 2016; Mayer et al., 1995): expertise, integrity and benevolence.

These three facets of epistemic trustworthiness can be explained as follows: To be a reliable source of knowledge, one needs to possess relevant *expertise*. If a person has the necessary knowledge and skills to understand a problem (e.g., is intelligent and qualified), they can provide reliable insights on that issue. However, having expertise alone is not sufficient to be a trustworthy source of knowledge, one also needs to possess *integrity*. This means one needs to use a reliable process of acquiring knowledge, for example, by honestly following the rules and standards of one's profession. Finally, one needs to possess *benevolence*; thus, one needs to have the best interest of the trustor (i.e., the lay public, society) in mind, for example, by being responsible.

Laypeople's ascriptions of a researcher's epistemic trustworthiness can be influenced by a wide range of epistemologically valid factors (e.g., having an academic background attesting to relevant expertise; Hendriks et al., 2015). However, there can also be epistemologically invalid or irrelevant factors, among them physical or demographic attributes (e.g., attractiveness, gender, and age) and, importantly, perceiver characteristics (e.g., preexisting attitudes, beliefs, and stereotypes; Origgi, 2012; Pornpitakpan, 2004; Zhu et al., 2016).

Asking laypeople for reasons to trust scientists, researchers' expertise seems most important. For example, 66% of respondents in the Science Barometer 2021 agreed or somewhat agreed that they trust in science and research because researchers possess relevant expertise, followed by integrity (57%) and benevolence (46%; Wissenschaft im Dialog/Kantar, 2021). Laypeople's reasons to distrust scientists, however, are mostly related to the facets of integrity

and benevolence. They worry about researchers' dependency on their funders (48%) and manipulation of findings (25%), while only 19% worried about researchers making mistakes, alluding to researchers' expertise. This data suggests that, while perceptions of researchers' expertise seem relatively solid (mostly as a reason to trust, only little as a reason to distrust), their integrity and benevolence is more ambivalently perceived (as a reason to trust but also as a reason to distrust scientists).

Considering that researchers' expertise is closely linked to their competence, whereas their integrity and benevolence allude to aspects of their warmth, one might draw parallels to laypeople's stereotypical perceptions of researchers as highly competent (i.e., having solid expertise), yet only moderately warm (Fiske & Dupree, 2014; Imhoff, Koch, et al., 2018; Rosman & Merk, 2021), possibly indicating some ambivalent perceptions regarding the warmth dimension. However, comparing the three facets of epistemic trustworthiness with the broader person perception models, the Stereotype Content Model and the ABC Model (Abele-Brehm et al., 2021; Fiske et al., 2002, 2007; Koch et al., 2016, 2020, 2021), it becomes apparent that the latters' subordination of trustworthiness as part of the communion/warmth dimension falls short. Although epistemic trustworthiness indeed entails warmth-related aspects (i.e., integrity and benevolence), a mere subsumption under communion/warmth neglects crucial competence-related aspects (i.e., expertise).

Moreover, covering both warmth-related and competence-related aspects of trustworthiness is consistent with other frameworks of trust, assuming a cognitive-rational, competence-related dimension and a socio-affective, warmth-related dimension of trustworthiness (McAllister, 1995; Morrow et al., 2004; see also Earle

& Siegrist, 2006). In line with this view, some closer inspection of relations between the three dimensions of epistemic trustworthiness reviewed above (i.e., expertise, benevolence, and integrity) suggests that they might be better aggregated into two factors: expertise as a competence-related factor and integrity plus benevolence as a warmth-related factor (see factor structures and correlations in all three manuscripts presented here, particularly Altenmüller, Lange, et al., 2021; Altenmüller, Nuding, et al., 2021). I will come back to this in the Final Discussion chapter.

3.2.3 Credibility of scientific findings

While (epistemic) trustworthiness primarily applies to *interpersonal* perceptions, trust in science also depends on the evaluation of the communication *content* (Hendriks et al., 2016; Sperber et al., 2010; see also first-hand and second-hand evaluations of knowledge: Bromme et al., 2015; Bromme & Goldman, 2014). I use the term "credibility" for such content-related perceptions to clearly distinguish them from person-related perceptions of "trustworthiness." Of note, other conceptualizations use the terms *source, message,* and *media credibility* to make a similar differentiation, thereby also considering the credibility of the communication channel (e.g., television, social media) which was not a focus in the present research program (e.g., Appelman & Sundar, 2016; Metzger et al., 2003).

Credibility of scientific findings (e.g., credibility of new scientific evidence or of the current scientific consensus on a specific question) can be defined as laypeople's acceptance of a claim as "true" (cf., Appelman & Sundar, 2016) leading to its integration in their knowledge system. And, if applicable, laypeople's credibility ascriptions might even go beyond this merely cognitive acceptance of

scientific evidence, informing related behavior, that is, laypeople's intentions to act in line with that evidence (e.g., Rieh & Danielson, 2007; Wathen & Burkell, 2002). However, not every scientific finding suggests direct behavioral consequences (e.g., basic research). So far, there is little research on the structure of content-related credibility. In the three articles presented in this dissertation, the credibility measures include cognitive as well as behavioral elements (e.g., acting in line with the communicated findings, sharing the findings with others, etc.) and I will come back to these conceptual questions regarding credibility in the Final Discussion chapter of this dissertation.

For the most part, person-related trustworthiness and content-related credibility are closely interconnected (e.g., "I trust the source, therefore I believe their statements to be reliable and valid"; see also the *causal chain account of trust*, Siegrist, 1999; Terwel et al., 2009). This fits with theories of persuasion; for example, the Elaboration Likelihood Model (Petty & Cacioppo, 1986) – a model theorizing how information is processed leading to attitude change – assumes persuasive effects of source trustworthiness, in particular when receivers (i.e., laypeople) lack the ability and/or motivation to elaborate the content of the message (i.e., the science) directly. Other times, however, trustworthiness and credibility seem disconnected (e.g., "I trust this person, but I do not believe their knowledge of the topic is reliable and valid information" or "I do not trust the source, but I believe their statements of this subject are true"). This might be explained by characteristics of the communication content: Some information might be inherently credible or not. For example, a claim can be highly logical or a truism or, in contrast, blatantly false or illogical (cf., Sperber et al., 2010). Some claims might even intuitively "feel" credible – a phenomenon called "truthiness"

(e.g., Schwarz & Newman, 2017). In such cases, the trustworthiness of the source is likely less relevant and first-hand evaluations dominate the judgment of "truth." Other factors that have been found to increase the credibility of statements which can be largely unrelated to perceptions of the source of that statement include argument quality and structure, evidence type, and extremity of the claim (Appelman & Sundar, 2016; Metzger et al., 2003; Pornpitakpan, 2004; Wathen & Burkell, 2002).

Overall, despite a wealth of theory on trustworthiness and credibility, there are only few empirical studies that consider a fine-grained picture of trust in science by clearly distinguishing these concepts and their facets. The articles presented in this dissertation aim for more specific insights which might, in turn, inform future research on the structures, relations, and differences of researchers' epistemic trustworthiness and the credibility of their findings.

3.2.4 Individualized and generalized trust in science

In this dissertation, trust in science mainly refers to trust between individuals (i.e., laypeople and specific researchers), thus constituting *individualized or particularized* trust (e.g., Freitag & Traunmüller, 2009; Morrow et al., 2004). But, trust in science can also pertain to other levels of analysis (e.g., Morrow et al., 2004; Resnik, 2011) – most relevantly here, to the trust relationship between individuals and groups or institutions (e.g., the collective of scientists or the social system of science). This level constitutes the macro perspective of a *generalized* "public trust in science" which is, for example, commonly used in public opinion surveys (see above).

How are individualized trust in science and generalized trust in science related? Particular experiences with (un)trustworthiness likely feed into

generalized trust (e.g., Freitag & Traunmüller, 2009). Regarding trust in science, this assumption is supported by correlational as well as causal links between laypeople's perceptions of individual researchers and their findings and laypeople's generalized trust in science as demonstrated in recent research including Manuscripts 1 and 3 of this dissertation (Altenmüller, Lange, et al., 2021; Hendriks et al., 2020; and see correlations in Manuscript 1). Generalized trust in science might therefore be undermined by *negative* experiences but also increased by *positive* experiences with individual researchers and their work. In this context, recent research (including Manuscript 2) suggests that learning about (failed) replications, cases of fraud or researchers' reform intentions in science might affect laypeople's trust in science on the individual as well as the generalized level (Altenmüller, Nuding, et al., 2021; Anvari & Lakens, 2018; Hendriks et al., 2015, 2020; Wingen et al., 2020).

Interestingly, such generalization processes of individual experiences to generalized perceptions bears some resemblance to the intergroup research literature. Intergroup Contact Theory (Allport, 1954; Pettigrew, 1998) assumes that positive experiences with individual members of an out-group are generalized to perceptions of their entire group and, thus, reduce prejudice and discrimination (e.g., Pettigrew & Tropp, 2006). Further, this literature highlights the importance of affective processes for reducing stereotypes (e.g., increasing empathy, reducing anxiety; Pettigrew & Tropp, 2008). Considering the interaction of laypeople with scientists as an *intergroup* situation, this might point to similar stereotype-reducing effects of positive experiences with individual researcher and their findings, particularly via affective routes.

3.2.5 Domain- and topic-specificity of trust in science

Taking a more detailed look at trust in science, one might question how useful broad public opinion surveys on generalized trust in science are (see above). Zooming in, there seems to be a large heterogeneity of trust in more specific subgroups of researchers and their findings (e.g., Pew Research Center, 2020; Rutjens et al., 2018; M. S. Schäfer et al., 2021; Scheitle & Guthrie, 2019). In fact, trust in science is largely *domain-specific*, that is, dependent on the scientific discipline, and *topic-specific*, that is, dependent on the research topic.

Regarding domain-specificity, when studying "trust in science" one not only needs to define "trust" but also "science." As mentioned above, the term "science" can pertain to the collective of scientists, the social system of knowledge generation, or even the body of scientific knowledge (see Hendriks et al., 2016); however, when referring to "science," most literature from English-language regions focuses on "hard" sciences (e.g., STEM disciplines), neglecting that the scientific enterprise generally encompasses many more disciplines, for example, those commonly summarized under "humanities and the arts." Noteworthily, the German language does not distinguish these fields of disciplines so sharply: "Wissenschaft" (i.e., science) refers to all disciplines that use scientific methods or rather the scientific social system for knowledge generation (e.g., consensus via rigorous peer review and debate; cf., Oreskes, 2021). Thus, German surveys or studies referring to trust in "Wissenschaft" are, at least semantically, less domain-specific. However, it is possible that independent of the inclusivity of the umbrella term "science" or "Wissenschaft," some disciplines are perceived as more representative or stereotypical for the scientific endeavor (e.g., domains vary in how "scientific" they are perceive to be; e.g., Allum, 2011; Scheitle & Guthrie,

2019). In fact, when asking people to draw pictures of "scientists," they spontaneously produce images related to the "hard" sciences (e.g., laboratory and field work, men wearing lab coats), suggesting that these ideas dominate stereotypes about "science" (e.g., Ferguson & Lezotte, 2020; Miller et al., 2018). Further, epistemological beliefs are known to be to some degree domain-dependent (e.g., Buehl et al., 2002; Stahl & Bromme, 2007; see also the Theory of Integrated Domains in Personal Epistemology, the TIDE Framework, Muis et al., 2006). Together, this suggests a need to further distinguish trust in science as trust in different fields of science or specific disciplines. Unfortunately, most public opinion surveys and studies on trust in science are not as fine-grained. In this dissertation, the theoretical and empirical contributions aim to broaden the picture beyond STEM-domains. Note, however, that most examples used here come from disciplines related to social sciences.

Beyond domain-specificity, trust in science is likely also topic-specific (see also the extended TIDE Framework; Merk et al., 2018). For example, it is more precise to consider topic-dependent trust than assuming a generalized (dis)trust in science, especially for topics relating to receivers' personal beliefs and ideology: Most prominently, past research focused on science skepticism related to topics like climate change, genetically modified foods (GMOs), and vaccine hesitation – topics where scientific consensus is overwhelmingly clear, yet skepticism is relatively widespread among the lay public (e.g., Pew Research Center, 2020; Rutjens et al., 2018; M. S. Schäfer et al., 2021). Interestingly, some major "science discrediting" campaigns like the oil lobby denying climate change or the tobacco industry downplaying the risks of smoking were actually not campaigns aimed at decreasing trust in science *per se*. Rather, they were directed at discrediting topic-

specific scientists and their findings that contradicted certain political and economic agendas by using science against itself: These companies and affiliated scientists used (seemingly) *scientific* arguments to oppose and question topic-specific scientific consensus (Oreskes & Conway, 2010).

3.2.6 Motivated science reception

While the considerable domain- and topic-specificity of trust in science demonstrates the value of taking the scientific context into account, it does not call for a merely descriptive, more detailed examination of topic- and domain-specific trust in science. Rather, a large share of the heterogeneity in trust in science can be explained by a theory-driven, context-sensitive motivated reasoning approach (Kunda, 1990). Motivated science reception (Hornsey, 2020; Lewandowsky & Oberauer, 2016; Rothmund et al., 2017) refers to the processing and perception of science in a motivated fashion: Instead of a general (dis)trust in science across (or even between) contexts, laypeople's trust in science might be largely driven by a motivation to retain their prior attitudes and belief system (e.g., Evans & Fetterman, 2021; Greitemeyer, 2014; Hutmacher et al., 2022; Nisbet et al., 2015) or to protect their values, self, and identity (e.g., Bender et al., 2016; Nauroth et al., 2015, 2017). For example, Nauroth and colleagues (2015) demonstrated that strongly identified "gamers" discredited science when it threatened their social identity (e.g., by writing hostile comments against studies that argued that violent video games are harmful), particularly targeting the scientific credibility of these studies. And, this was rooted in a need to protect a positive social identity as "gamer." Manuscript 3 in this dissertation considers such motivated science reception processes regarding the social aspects of science communication. It demonstrates how information that researchers reveal about themselves might be

used by laypeople in a motivated manner to up- or devalue researchers and their findings in order to uphold their preexisting attitudes towards the research topic (Altenmüller, Lange, et al., 2021).

3.3 Self-disclosure in science communication

Considering these theoretical and empirical insights into laypeople's perceptions of scientists and their trust in science, it becomes evident that trust in science might not yet have reached its full potential among the lay public: As outlined above, while researchers' competence-related trustworthiness (i.e., expertise) is perceived as quite solid and high, their warmth-related, affective trustworthiness (i.e., integrity and benevolence) seems more malleable and leaves room for possible improvements in the eyes of the public. Fittingly, social psychology looks back on a wealth of research on interpersonal, affective processes involved in relationship building, communication, and trust. An especially promising avenue for application to the scientific sphere might be research on *self-disclosure*.

3.3.1 Self-disclosure

Self-disclosure describes the process of deliberately revealing personal information about oneself to one or more other people (e.g., Cozby, 1973; Greene et al., 2006; Omarzu, 2000). This broad definition of self-disclosure can be further distinguished in *breadth* (i.e., how much is revealed?), *depth* (i.e., how intimate is the revealed information?), and *quantity/length* (i.e., how often or how long is the person engaging in self-disclosure?). The Theory of Social Penetration (Altman & Taylor, 1973) assumes that systematic changes in self-disclosing communication are necessary for building and maintaining relationships. For this, the breadth and

depth of self-disclosure need to be gradually increased in a reciprocal manner. This can be metaphorically depicted as peeling away layers of an onion (i.e., the "onion model," cf. Carpenter & Greene, 2016): In self-disclosing interactions, first, the outer-layers are penetrated, that means superficial information like biographical facts (e.g., name, age) is exchanged. Going deeper, the information becomes more intimate and attitudes, values, and opinions are revealed. Finally, at the onion's core, personal beliefs, wishes, and fears are disclosed. Thus, the deeper one "peels the onion," the more intimate the revealed information.

Explaining how revealing personal information leads to positive interpersonal outcomes (e.g., forming a relationship), the literature mostly takes two theoretical perspectives (cf., Omarzu, 2000). The *social exchange approach* (e.g., Omarzu, 2000; Worthy et al., 1969) assumes that exchanging intimate information is a rewarding experience (while also holding potential costs) and spurs reciprocal disclosure. The *social strategy approach* (e.g., Omarzu, 2000; Tardy & Smithson, 2018) assumes that self-disclosure is strategically applied to serve specific functions (e.g., self-expression, social validation, relationship development, or social control). Based on these theoretical approaches, Omarzu (2000) proposed the Disclosure Decision Model. It outlines when, what, how, and to whom individuals will disclose personal information depending on their goals (i.e., social strategy) and weighing possible rewards against potential risks (i.e., social exchange).

Typically, self-disclosure mainly concerns the process of disclosing personal information in communication. Sometimes, however, it has also been conceptualized as a personality characteristic: In his review, Cozby (1973) summarizes some early research on interindividual differences in the level of self-

disclosure (e.g., dependent on gender or cross-cultural differences). Interestingly, results on such interindividual differences have been mixed (see, for example, Dindia & Allen, 1992, on gender differences in self-disclosure). In this dissertation, I investigate the effect of engaging in the *process* of self-disclosure. However, I also consider laypeople's general expectations of researchers' disclosure (see Manuscript 1) which could be interpreted as lay theoretical assumption of such trait-like self-disclosure. These baseline expectations for usual self-disclosure based on communicators' membership in the social category of "scientists" could, in turn, impact how their engagement in the process of self-disclosure is perceived.

3.3.2 Benefits of self-disclosure

There is a wealth of evidence for the positive effects of self-disclosure on person perception across various social situations. Most prominently, self-disclosure leads to increases in warmth-related judgments like feelings of closeness and liking or ascriptions of empathy (e.g., Aron et al., 1997; Collins & Miller, 1994; Kadji & Schmid Mast, 2021; Sprecher, Treger, & Wondra, 2013; Sprecher, Treger, Wondra, et al., 2013; Tal-Or & Hershman-Shitrit, 2015). For example, Collins and Miller (1994) demonstrate in their meta-analysis that intimate self-disclosers are liked more than low self-disclosers, that people disclose more to those they like, and that disclosing leads to increased liking for the recipient of that disclosure. Of particular interest for this dissertation are the effects of self-disclosure on trust. In this regard, studies also demonstrate a positive link of self-disclosure and trust in diverse contexts from school, work place, health to economic interactions (e.g., Huang, 2015; Kim et al., 2016; Mazer et al., 2009; Nazione et al., 2019; Nifadkar et al., 2019; Wheeless & Grotz, 1977).

Considering these benefits of self-disclosure for interpersonal evaluations and predictions from the Disclosure Decision Model (Omarzu, 2000), researchers might use self-disclosure strategically to increase laypeople's perceptions of them as approachable, warm, and likable – thus, increasing perceptions of affective trustworthiness. This might be facilitated by self-disclosure providing individuating information about a researcher, and thereby lessening laypeople's reliance on their general stereotypes about scientists in this communication situation (cf., Rubinstein et al., 2018). In fact, a few studies have adopted this research focus on self-disclosure in science communication in recent years: Jarreau and colleagues (2019) demonstrated that researchers sharing selfies were perceived as warmer and more trustworthy compared to researchers who only showed their work equipment. Saffran and colleagues (2020) demonstrated that using first-person (vs. neutral) communication increased perceived authenticity (which, they argue, is closely linked to integrity and benevolence).

Coming back to parallels to research on intergroup contact (see above), contact interventions that build on self-disclosure have been shown to be successful in reducing stereotypes and promoting positive interpersonal evaluations that were, then, generalized from the individual experience to perceptions of the whole group (e.g., Ensari & Miller, 2002; Kotzur et al., 2019). Thus, construing science communication as an intergroup situation, one could assume that, in a similar fashion, researchers' self-disclosure might improve laypeople's perceptions of individual researchers and their findings and, in turn, their generalized trust in science.

3.3.3 Challenges for self-disclosure

The evidence seems encouraging that researchers' self-disclosure might hold potential for boosting public trust in science; however, the context of science communication also implies specific challenges for the effectiveness of self-disclosure. Science communication, especially the traditional formats which are still the most common (e.g., interviews or opinion pieces in the news media), rely on unilateral communication: This means, a researcher communicates their findings towards a broad audience with only scarce opportunities for exchange between recipients and communicators. Thus, engaging in self-disclosure in such a communication format constitutes *broadcasted self-disclosure* (e.g., Bazarova, 2012; Rains & Brunner, 2018; Schlosser, 2020): unilaterally revealing personal information to a large audience of (mostly) anonymous strangers. This poses the problem for researchers to adopt a level of self-disclosure that is perceived as *appropriate* by the audience of their science communication.

In general, positive effects of self-disclosure likely only unfold when the self-disclosure is seen as appropriate (e.g., Brewer & Mittelman, 1980; Caltabiano & Smithson, 2010; Chaikin & Derlega, 1974; Derlega et al., 1976). Consequently, scholars have questioned the linear association of self-disclosure and positive person evaluation outcomes (i.e., "the more, the better"), rather suggesting a curvilinear relationship (e.g., Cozby, 1972, 1973; Gelso & Palma, 2011): Too extreme self-disclosure is perceived as inappropriate and maladjusted; thus, leading to more negative outcomes than moderate self-disclosure. Another explanation might be that the more one discloses, the likelier dissimilarities are revealed between the interaction partners, leading to less liking (i.e., "less is

more," Norton et al., 2007; note, however, that this conclusion has been challenged, e.g., Sprecher, Treger, & Wondra, 2013).

In particular, science communication as broadcasted form of communication might complicate achieving perceived appropriateness. Research on social media communication suggests that audiences react more sensitively to broadcasted self-disclosure (Rains & Brunner, 2018) and perceive it as more intimate (Bazarova, 2012) than the same self-disclosure in private communication. Additionally, traditional science communication formats lack the element of reciprocity and, thus, an important source of information about appropriateness: In reciprocal interactions, the communicators typically increase intimacy in a gradual fashion and, by taking turns in disclosing, it is possible to implicitly come to a mutual understanding of the consensual, thus appropriate, level of intimacy (cf., Dindia, 2002; Sprecher, Treger, Wondra, et al., 2013). Most importantly, the appropriateness of self-disclosure in science communication might be hindered by the very same thing self-disclosure could help to overcome: laypeople's stereotypical perceptions of scientists.

As scientists are typically perceived as moderately warm but highly competent (e.g., Fiske & Dupree, 2014; Imhoff, Koch, et al., 2018; Rosman & Merk, 2021), laypeople might have high expectations of scientists' communication to be very professional, objective, and distanced. Therefore, self-disclosure in science communication could violate these expectations and seem inappropriate. Further, engaging in self-disclosure might implicate that researchers do not know how to correctly and appropriately fulfill their professional role as "scientist" (cf., Schwabe et al., in prep.). Consequently, self-disclosing researchers might be perceived as unprofessional, incompetent, and unqualified (i.e., lacking expertise).

3.4 The present research program

Taken together, this dissertation integrates theoretical assumptions from research on stereotypes, trust in science, and self-disclosure, and applies them to the context of science communication. In short, laypeople's trust in science might be obstructed by their stereotypical views of scientists as lacking warmth which could be countered by researchers engaging in self-disclosing science communication, potentially promoting their affective trustworthiness, and, thus, laypeople's trust in science. However, at the same time, such self-disclosure could also impair laypeople's trust in science, particularly their perceptions of researchers' expertise. The present research program follows three superordinate goals: 1) testing the effects of researchers' self-disclosure in science communication on laypeople's trust in science, 2) exploring potential boundary conditions of self-disclosure in science communication, and 3) gaining empirical insights into trust in science.

First, while prior research indicates predominantly positive effects of self-disclosing communication on warmth-related person perception in general, and, thus, on perceptions of researchers' (affective) trustworthiness in particular, there might be challenges for these benefits of self-disclosure that are specifically associated with the context of science communication (i.e., laypeople's stereotypical views of scientists and appropriateness), particularly regarding researchers' perceived expertise. Thus, researchers' self-disclosure in science communication might have trade-off effects on laypeople's trust in science: Increases in warmth-related trustworthiness might go hand in hand with decreases in competence-related trustworthiness. This idea is in line with findings on compensatory social judgment (e.g., Judd et al., 2005; Yzerbyt et al., 2008):

Warmth and competence perceptions regarding stereotypes can be negatively related, that means, increases in one dimension lead to decreases in the other. Thus, the first goal of this dissertation is to investigate the potentially ambivalent effects of researchers' self-disclosure on laypeople's trust in science (i.e., perceptions of researchers' trustworthiness and the credibility of scientific findings) in a general manner.

Second, more specifically, whether self-disclosure in science communication has positive or negative effects on laypeople's trust in science might also depend on the specific information that is revealed by the communicating researcher and what it means for the receiving layperson. This dissertation explores such boundaries of self-disclosure in science communication. For example, researchers' self-disclosure might be of general nature (e.g., personal facts about their non-academic life) or specifically relevant to the communicated science (e.g., critical thoughts about their scientific work, personal involvement with the research topic). Thus, in this dissertation, I consider laypeople's perception of researchers who disclose content of varying association with the communicated science, as outlined in the next chapter. In the Final Discussion, I will elaborate further on the issue of self-disclosure content. Regarding receiver characteristics, research on motivated science reception suggests that disclosed information could be interpreted in line with preexisting attitudes. Thus, in this dissertation, I will also consider the role of laypeople's attitudes for the reception of researchers' self-disclosure in science communication on trust in science.

Third, this research program aims for a better understanding of the concept of trust in science to inform theory and practice on the matter. The included

manuscripts will provide insights regarding the conceptualization of trust in science as researchers' epistemic trustworthiness – including expertise and affective trustworthiness (i.e., integrity and benevolence) – as well as the credibility of scientific findings – including a cognitive (i.e., acceptance) and behavioral facet (i.e., behavioral intentions. In the Final Discussion, I will reflect on what we can learn from the present studies about (laypeople's) trust in science.

4 Summaries of manuscripts

This dissertation covers three manuscripts that include eleven experiments investigating the effects of researchers' self-disclosure in science communication on laypeople's trust in science (i.e., their perceptions of researchers' epistemic trustworthiness and the credibility of their findings). The three manuscripts deal with different degrees of interconnection between the disclosed information and the communicated science. In Manuscript 1, we first investigated the basic effects of researchers' self-disclosure, that is, revealing personal information which is of predominantly neutral valence and functions as a mere illustration to the communicated science. It is unrelated to the underlying scientific processes producing the communicated science (e.g., conduction of the presented study). In Manuscript 2, we then tested the effects of researchers' self-disclosure which is directed at the scientific processes generating their findings, yet not interwoven with the research topic itself: researchers' self-corrective stance towards their work (i.e., disclosing doubts about past findings and reform intentions for future work). In Manuscript 3, we analyze the effects of researchers' self-disclosure which implies a potentially full blending of their personal motivations and their research topic and can also be perceived as colluding with the underlying scientific processes (i.e., conducting research on idiosyncratically relevant topics, so-called "me-search"). In the following, the manuscripts and their findings are summarized.

4.1 Summary of Manuscript 1

Manuscript 1 includes six online experiments (overall $N = 2414$) and a field study ($N = 480$) on the basic effects of researchers' self-disclosure (i.e., framing the same information as either personal or neutral) across various science

communication scenarios and formats (i.e., scientific presentations, videos, and social media, as well as an audio guide through a science museum exhibit). We experimentally manipulated self-disclosure (vs. no self-disclosure) in between-subject designs and tested its effects on laypeople's perceptions of closeness towards the communicating researchers, ascriptions of affective trustworthiness (i.e., integrity and benevolence) and expertise as well as their acceptance of the communicated findings as "true" and behavioral intentions (i.e., the facets of credibility).

While designs and samples were quite similar across the online studies, the results and effect sizes varied considerably. Thus, we integrated these findings in a mini meta-analysis: Here, small positive meta-analytical effects of self-disclosure on closeness and affective trustworthiness, as well as a small negative meta-analytical effect of self-disclosure on expertise emerged, while self-disclosure had no meta-analytical effects on acceptance or behavioral intentions. Thus, this supports the notion of a trade-off: Researchers might benefit from self-disclosure in terms of their perceived affective trustworthiness, while they might lose in terms of their perceived expertise.

In the field study, laypeople visited the marine science exhibition in a large science museum in Germany while listening to an interactive audio guide that varied in self-disclosure (vs. no self-disclosure) and the role ascribed to the guide (marine scientist vs. museum guide) or a neutral control condition. Results from this field study further question whether the effects of researchers' self-disclosure on laypeople's trust in science have practical relevance: In this real and engaging science communication scenario with very high ecological validity, we did not find

any effects of self-disclosure (vs. no self-disclosure) on perceptions of closeness, trustworthiness or credibility.

Across all studies, we also considered a range of contextual factors (i.e., researchers' gender, researchers' discipline, opportunity for exchange, role of discloser). However, these factors did not influence the effects of self-disclosure on any of our outcome variables.

Taken together, basic self-disclosure in science communication (i.e., framing illustrative information as personal) might only have small, volatile, and counteractive effects on laypeople's perceptions of researchers' trustworthiness and likely no effect on the perceived credibility of their findings. The practical relevance of these effects is questionable.

4.2 Summary of Manuscript 2

Manuscript 2 includes two studies (overall $N = 702$) investigating researchers' disclosure of doubts about their own findings (i.e., self-criticism) and reform intentions for their future work routines, and its effects on laypeople's ascriptions of researchers' epistemic trustworthiness, the anticipated credibility of their future findings, and laypeople's willingness to further engage with the researchers and their findings. Both studies are set in traditional science communication scenarios (i.e., printed interviews on research about group processes).

In Study 1, we manipulated self-criticism and reform intentions in a 2 (self-criticism vs. no self-criticism) × 3 (reform intentions: no vs. minor vs. major) between-subjects design. Results showed a small positive effect of expressing self-criticism (vs. no self-criticism) on ascriptions of affective trustworthiness (i.e., integrity and benevolence), while ascriptions of expertise, credibility, or

willingness to engage with science remained unaffected. Announcing reform intentions (minor and major reform intentions compared to no reform intentions) had large positive effects on ascriptions of affective trustworthiness, expertise, and credibility, as well as small effects on laypeople's willingness to further engage with science. In Study 1, expressing even minor reform intentions led to increases in trustworthiness and credibility, suggesting that it might be enough to signal some willingness to implement reforms.

Study 2 replicated and scrutinized these findings. It was split into two parts (i.e., two interviews presented in randomized order) and each part had three experimental conditions (between-subjects). This time, the design included control conditions: Part 2A manipulated self-criticism (i.e., self-criticism vs. no self-criticism vs. no information) and part 2B manipulated reform intentions (i.e., reform intentions vs. no reform intentions vs. undecided). Regarding self-criticism, Study 2 revealed that the effects were largely driven by the no self-criticism condition. That means, expressing *no* self-criticism led to a *decrease* in ascriptions of affective trustworthiness, expertise, credibility, and laypeople's willingness to engage with science compared to the no information control group or explicitly engaging in self-criticism. Thus, scientists revealing an overconfident and uncritical attitude towards their own research might decrease laypeople's trust in science. Regarding reform intentions, expressing intentions to implement reforms led to an increase in ascriptions of trustworthiness (i.e., affective trustworthiness and expertise) and credibility compared to being undecided; actively negating reforms decreased ascribed expertise and credibility compared to being undecided. Further, laypeople reported a higher willingness to engage with science when researchers intended reforms (vs. no reforms).

Taken together, across both studies, researchers' self-disclosure regarding their thoughts directed at the scientific processes underlying their work increased laypeople's epistemic trustworthiness ascriptions, credibility perceptions, and their willingness to further engage with science – but only if the disclosed information implied a self-corrective stance towards science (i.e., being self-critical and willing to reform one's work routines). Practically, these findings suggest that researchers' hesitation and fear that self-correction may damage their reputation is unwarranted.

4.3 Summary of Manuscript 3

Manuscript 3 consists of two studies (overall $N = 621$) on the effects of researchers' self-disclosure of personal motivations underlying their research interests (i.e., idiosyncratic relevance or "me-search") on laypeople's ascriptions of epistemic trustworthiness and credibility. This time, we also considered motivated science reception by including laypeople's preexisting attitudes towards the research topic (here: LGBTQ and veganism). As science communication scenario, we asked participants to inspect fictional public research proposals including statements by the conducting researchers. Of note, they only saw the proposal for a research project but no outcomes were presented. In both studies, we used a two-group, between-subjects design manipulating researchers' personal affection by the research topic within the researcher statement in the fictional proposal (i.e., Study 1 – homosexual vs. not homosexual, Study 2 – vegan vs. not vegan). Further, we measured laypeople's attitudes towards the research topic in question.

Findings from both studies demonstrate that laypeople's perceptions of researchers' disclosure of engaging in "me-search" can be positive as well as negative depending on laypeople's preexisting attitudes: In line with motivated

science reception, “me-searchers” were perceived as more trustworthy and their future findings were anticipated to be more credible when participants had positive attitudes towards the research topic. When they had negative attitudes, this pattern reversed. Further, Study 2 demonstrated that these motivated perceptions were generalized to evaluations of the entire field of research.

Taken together, these findings suggest that such self-disclosure (i.e., disclosing “me-search”) does not have effects *per se* on trust in science but rather laypeople might use the revealed information in a motivated manner, either to preemptively *upvalue* researchers and their findings in terms of trustworthiness and credibility when laypeople expect the (not yet known) science to support their preexisting attitudes or to *downvalue* researchers and their findings when they expect the findings to contradict their attitudes.

5 Manuscript 1: Self-disclosure in science communication

This manuscript is under review for publication:

Altenmüller, M. S., Kampschulte, L, Verbeek, L., & Gollwitzer, M. (under review). Science Communication gets Personal: Ambivalent Effects of Self-Disclosure in Science Communication on Trust in Science.

[manuscript printed as submitted]

Abstract

In an attempt to display themselves as warm, approachable, and trustworthy, researchers might reveal personal details about themselves (i.e., self-disclosure) when communicating their science to the public. Here, we test whether self-disclosure in science communication can actually increase public trust in science. We present six online experiments (overall *N*=2431), integrate their results in a mini meta-analysis and report a field experiment in a science museum (*N*=480): In sum, our findings suggest that self-disclosure leads to small, but measurable increases in laypeople's feelings of closeness towards researchers and perceptions of researchers' warmth-related trustworthiness; yet, self-disclosure also leads to decreases in competence-related trustworthiness perceptions. The credibility of scientific findings was, overall, unaffected by self-disclosing communication. Findings from the field study further question whether self-disclosure in science communication has any practical relevance.

[129 words]

Key words: self-disclosure, science communication, closeness, trustworthiness, credibility.

Science Communication gets Personal: Ambivalent Effects of Self-Disclosure in Science Communication

Marlene's guinea pigs are called Noxi and Neri. They hate each other, but they share a love for cucumber. This information is irrelevant to this article and it might be unprofessional to start a scientific article with such personal details about one of the authors; but, then again, it reveals something about the people behind this research and, thus, might make Marlene, the researcher, appear closer, more approachable, and, thus, more trustworthy. Revealing intimate information about oneself—self-disclosure—implies a signal to the reader: I trust you with this information, so you can trust me.

Self-disclosure, which is broadly defined as revealing personal information to another person, has long been known to promote positive evaluations of one another and is often seen as central mechanism in interpersonal relationship building (e.g., Social Penetration Theory, Altman & Taylor, 1973). Research on self-disclosure in interpersonal interactions (e.g., among colleagues, acquaintances, or even strangers) showed that it increases feelings of closeness and liking (e.g., Aron et al., 1997; Collins & Miller, 1994; Sprecher, Treger, & Wondra, 2013; Sprecher, Treger, Wondra, et al., 2013) and trust in the self-disclosing person (e.g., Huang, 2015; Kim et al., 2016; Wheeless & Grotz, 1977). The present research tests whether self-disclosure in science communication, that is, including personal information in the communication about research-related facts, can increase the public's trust in researchers and their findings along similar lines.

Using the famous two-dimensional space of person perception (i.e., agency/competence and communion/warmth; Abele-Brehm & Wojciszke, 2018;

Fiske et al., 2002), laypeople often perceive researchers as highly competent, yet only moderately warm (Fiske & Dupree, 2014; Rosman & Merk, 2021). In other words, scientists are viewed by many in a stereotypical fashion: as highly intelligent and competent; yet also as distanced, cold, socially unskilled, and maybe even immoral (e.g., Losh, 2010; Rutjens & Heine, 2016). This stereotypical perception poses a challenge for building and maintaining trust in science: After all, trustworthiness perceptions are rooted both in competence-related (i.e., expertise) *as well as* warmth-related judgments (i.e., integrity and benevolence; Fiske & Dupree, 2014; Hendriks et al., 2015; McAllister, 1995). Thus, to be perceived as trustworthy, it is not enough to present oneself as highly skilled and professional; researchers should also present themselves as warm, approachable, benevolent and integer (Benson-Greenwald et al., 2021; Fiske & Dupree, 2014; Hendriks et al., 2016).

Anecdotally, science communicators are aware of this challenge and act accordingly: For example, on Twitter some science communication channels focus on sharing personal insights in work routines as well as every aspect of life as a scientist (e.g., "*@RealScientists*"), science museums increasingly work with personalization features (e.g., exploring the personal backgrounds of historic scientists and bringing them "back to life" via staged scenes or reenactments; Garner & Rossmanith, 2021), and health care works are encouraged to reveal personal experiences with vaccinations when talking to their patients (e.g., Pan American Health Organization, 2020).

However, researchers who share intimate details about their personal lives might also violate laypeople's stereotypical expectations of how researchers *should* communicate their scientific findings: namely, in a highly professional, objective,

distanced way. Research on self-disclosure, too, highlights the importance of perceived appropriateness of self-disclosure for a given communication situation (e.g., Caltabiano & Smithson, 2010; Cozby, 1973; Omarzu, 2000; Zhang et al., 2009). For example, Cozby (1973) reviewed evidence showing that inappropriate (i.e., too much) self-disclosure might be seen as maladjusted; thus, suggesting a curvilinear relationship of self-disclosure and positive person evaluation. Therefore, self-disclosure in science communication may increase warmth at the cost of decreased competence ascriptions (see also compensatory social judgment; Judd et al., 2005; Yzerbyt et al., 2008). Further, a self-disclosing researcher might be perceived as unaware of how to fulfill his role as science communicator appropriately, indicating a lack of competence and professionalism.

Here, we test whether self-disclosure in science communication not only leads to an increase in ascribed benevolence and integrity (i.e., the warmth-related, affective facet of trustworthiness) but to a decrease in ascribed expertise (i.e., the competence-related facet of trustworthiness). Moreover, given the possibility of a trade-off between benevolence and integrity on the one hand and expertise on the other hand, this raises the question of how researchers' self-disclosure impacts the perceived credibility of their scientific findings. In science communication, credibility perceptions can be conceptualized with two facets (Altenmüller, Lange, et al., 2021; Altenmüller, Nuding, et al., 2021). First, credibility refers to laypeople's acceptance of a scientific claim as "true" (i.e., the cognitive facet of credibility; c.f., Appelman & Sundar, 2016), and, second, credibility can result in a layperson's intention to behave in accordance with a scientific claim or its implications (i.e., the behavioral facet of credibility; c.f. Rieh & Danielson, 2007; Wathen & Burkell, 2002). It is possible that a trade-off between

benevolence/integrity and expertise due to self-disclosure might also apply to perceptions of credibility.[1]

In this article, we present a research program testing the effects of researchers' self-disclosure in science communication on (1) laypeople's perceived closeness towards these researchers, (2) their ascriptions of trustworthiness (i.e., integrity and benevolence and expertise), and (3) the perceived credibility of the communicated scientific findings (i.e., acceptance and behavioral intentions). Based on our theoretical reasoning, we expect researchers' self-disclosure to increase laypeople's perceived closeness and ascriptions of affective trustworthiness (i.e., integrity and benevolence), while we expected ascriptions of expertise to remain unaffected or even decrease. In addition, we expect positive effects on credibility. We tested the effects of self-disclosure in six online experiments using different science communication scenarios (i.e., scientific presentations, videos, and social media) as well as a field study conducted in a science museum.

We report all data exclusions, manipulations, and all measures: We summarize all measured variables in Table 5.1. In Table 5.2, we provide more details about the methods and design of all studies (i.e., date of sample collection, sample details, design, sensitivity power analyses, dependent variables). Table 5.3 – 5.9 display means and standard deviations per experimental group for each study. Additionally, we provide information on data exclusions, all study materials, correlations among measured variables, open data and analysis scripts in an online

[1] In our preregistrations, we assumed positive effects of self-disclosure on affective trustworthiness (i.e., integrity and benevolence) and credibility given that prior research has accumulated stronger empirical evidence for positive effects than for potential negative effects (on expertise).

supplement (https://osf.io/s39u2/). The crucial effect tested in all studies is a (between-participants) difference between a self-disclosure and a no self-disclosure condition, that is, a mean difference or contrast between these two conditions. According to sensitivity analyses (see Table 5.2), all studies reported here had enough statistical power to detect even a small effect (i.e., d=.31 or smaller). All studies were designed and conducted in accordance with ethical guidelines and legal regulations regarding data protection. In all online studies, participants were compensated by chances in voucher lotteries (total value per study: 50-225€, depending on sample size goals and length of study); in the field study, participants received free entry into the science museum.

Online Experiments

Study 1

The first study investigated the effects of a researcher's self-disclosure on laypeople's perceptions of closeness and their ascriptions of trustworthiness (i.e., expertise and affective trustworthiness) in a standard science communication scenario (i.e., a scientific talk) using a two-group between subject design. To manipulate self-disclosure without producing demand effects, we used a relatively subtle manipulation: all participants saw a research talk given by a (fictious) researcher, but several pieces of illustrative information in this talk was either framed as *personal information* (in the self-disclosure condition) or *neutral information* (in the control condition).

Methods. After obtaining informed consent, we measured demographic variables (age, gender, occupation; see Table 5.1 for more details and procedural overview). Following, we told participants that on the next six pages of the questionnaire they would see presentation slides by a (fictious) researcher called

Dr. Lauber (no gender was indicated) on the effects of light conditions on music experience at a scientific conference called "Science Impulses." We asked them to look closely at the slides. Participants were randomly assigned to one of two experimental conditions which were identical in presented content, yet differed in regard to the subtext of illustrating pictures: We framed this subtext as either neutral (i.e., *no self-disclosure* condition; e.g., "*View from a window*" or "*This is how it might look in a hotel*") or as disclosing personal information (i.e., *self-disclosure condition;* e.g., "*View from my window*" or "*This is how it looked like in my last hotel*") in three of these six presentation slides. Table 5.3 summarized all manipulated texts. Note, the manipulated information was associated with the illustrations, not with the described scientific process. After that, we measured participants' responses regarding our dependent and control variables, as detailed in Table 5.1 and Table 5.2.

Results and Discussion. We included 323 participants in our analyses (see Table 5.2 for details). The experimental conditions in Study 1 did not differ regarding control variables (engagement frequency: p=.691; engagement experiences: p=.668; expectation of scientists' disclosure: p=.956). Speaking for the effectiveness of our self-disclosure manipulation, our participants perceived the researcher to disclose significantly more personal information in the self-disclosure condition than in the control condition, $t(321)$=-9.14, p<.001, d=-1.02, 95% CI_d [-1.25, -.0.78].

Mean differences between conditions are reported in Table 5.4. Participants in the self-disclosure condition ascribed significantly higher affective trustworthiness to the researcher, $t(321)$=-3.39, p<.001, d=-0.38 [-0.60; -0.16], but not higher expertise, $t(321)$=0.92, p=.358, d=0.10 [-0.12; 0.32], and, they felt

significantly closer to the researcher, $t(321)=-2.30$, $p=.022$, $d=-0.26$, 95% CI_d [-0.48, -.0.04] compared to participants in the control condition. Thus, findings from Study 1 support the assumption that even subtle forms of self-disclosure in science communication can increase laypeople's feelings of closeness towards a researcher and ascriptions of affective trustworthiness (i.e., benevolence and integrity); yet, self-disclosure did not affect general perceptions of competence or ascriptions of expertise.

Study 2

Study 2 was designed to replicate the findings from Study 1 and to test the effects of self-disclosure on perceived credibility of the communicated scientific findings (including laypeople's acceptance of these findings as "true" and their behavioral intention to act in line with these findings).[2] Study 2 was preregistered (https://osf.io/gqsbm).

Methods. Study 2 was designed very similarly to Study 1: After providing informed consent, participants again saw presentation slides. This time, the slides came from an alleged blitz talk on the effects of coffee consumption on music experience by a researcher named Dr. Meyer. Again, participants were randomly assigned to one of two experimental conditions (self-disclosure vs. control). The manipulation was designed in the same fashion as in Study 1 (c.f., Table 5.3), but we tried to strengthen the self-disclosure by mixing the previously used rather subtle self-disclosing information with more intimate disclosure of feelings and weaknesses (e.g. "*After drinking coffee I am more nervous while giving a*

[2] Another preregistered aim was to test whether the effects of self-disclosure on credibility were mediated by epistemic trust. As this is not the focus of the present paper and, due to the absence of effects of self-disclosure on credibility or trustworthiness, we did not conduct the preregistered mediation analyses.

presentation" vs. "*After drinking coffee some people are more nervous while giving a presentation*"). In Study 2, the manipulation was applied to four out of six slides. After participants viewed the slides, dependent and control variables were measured (see Table 5.1 and 5.2).

Results and Discussion. We included 344 participants in our analyses (see Table 5.2). The experimental conditions in Study 2 did not differ regarding control variables (engagement frequency: p=.404; engagement experiences: p=.753; pro-science attitude: p=.393; expectation of scientists' disclosure: p=.123; political orientation: p=.815). Again, participants perceived the researcher to disclose significantly more personal information in the self-disclosure condition than in the neutral condition, $t(342)$=-6.47, p<.001, d=-0.70, 95% CI_d [-0.92, -.0.48].

Descriptive statistics are reported in Table 5.5. Contrary to Study 1, the two experimental conditions neither differed on affective trustworthiness, $t(342)$=-0.78, p=.435, d=-0.08, 95% CI_d [-0.30, 0.13], nor on ascriptions of expertise $t(342)$=0.61, p=.544, d=0.07, 95% CI_d [-0.15, 0.28]. Similarly, there were no significant differences on credibility: Both groups reported very similar acceptance, $t(342)$=0.45, p=.654, d=0.05, 95% CI_d [-0.16, 0.26], and behavioral intentions, $t(342)$=-0.28, p=.778, d=-0.03, 95% CI_d [-0.24, 0.18]. Following up, we explored the effect of our self-disclosure manipulation on closeness: As in Study 1, participants felt significantly closer to the self-disclosing scientist than the neutral scientist, $t(342)$=-2.10, p=.036, d=-0.23, 95% CI_d [-0.44, -0.01].

There are a number of explanations for the divergent results between Studies 1 and 2. First, the education level (i.e., highest obtained degree) was higher in Study 2 than in Study 1 (96% were currently enrolled in a university program or already held an academic degree). Especially students who likely

interact with researchers on a daily basis might be less receptive to our manipulation compared to laypeople who have less contact with communicating scientists. Second, our manipulation of self-disclosure was stronger (i.e., less subtle) than in Study 1. Considering the inverted u-shape (i.e., curvilinear) association of self-disclosure with other variables (e.g., Cozby, 1973; Gelso & Palma, 2011), it is possible that with this level of intimacy we surpassed the degree of self-disclosure perceived as appropriate for such a scenario which would have attenuated possible positive effects.

Study 3

We conducted a third study to account for these issues. This time, we excluded currently enrolled students from participation and designed three experimental conditions (*no*, *medium* and *extreme* self-disclosure) to scrutinize potentially backfiring effects of inappropriately high levels of self-disclosure on our DVs. We still expected an increase in affective trustworthiness and credibility for moderate compared to no (or extreme) self-disclosure (replicating our results from Study 1).[3] Study 3 was also preregistered (https://osf.io/mktv9).

Methods. Again, Study 3 was designed very similarly to the previous studies. After providing informed consent, participants viewed presentation slides from an alleged blitz talk about the effects of weather conditions on music experience by a researcher called Dr. Winkler. This time, participants were randomly assigned to one of three experimental conditions: *no*, *moderate* or *extreme* self-disclosure. As before, content was held constant while pictures were

[3] As in Study 2, we also preregistered mediation analyses of the effect of self-disclosure on credibility via trust. As this is not the focus of the present paper, these analyses are not reported here.

framed as either purely illustrative and non-personal (*no* self-disclosure control group; e.g., "*View from a window: dark, depressing, rainy winter weather*"), as moderately personal (*moderate* self-disclosure; e.g., "*View from my window: dark, depressing, rainy winter weather*"), or as personal and intimate (*extreme* self-disclosure; e.g. "*View from my window: dark, rainy winter weather makes me feel a bit depressive…*"). The manipulation was applied to four out of six presented slides. Dependent variables were the same as in Study 2 (see Table 5.1 and Table 5.2).

Results and Discussion. We included 618 participants in our analyses (see Table 5.2). Again, the experimental conditions did not differ regarding control variables (engagement frequency: p=.217; engagement experiences: p=.322; pro-science attitude: p=.667; expectation of scientists' disclosure: p=.730). The researcher in the extreme self-disclosure condition was perceived to disclose the most personal information and the researcher in the no self-disclosure condition was perceived to disclose the least personal information, $F(2, 615)=86.50$, $p<.001$, $\eta^2=.22$, 95% CI_{η^2} [.16, .27] (no < moderate < extreme, both: $p<.001$).

Descriptive statistics are reported in Table 5.6. The experimental groups did not differ significantly on affective trustworthiness, $F(2, 615)=1.21$, p=.299, $\eta^2<.01$, 95% CI_{η^2} [.00, .02], but they did differ on expertise $F(2, 615)=3.12$, p=.045, $\eta^2=.01$, 95% CI_{η^2} [0.00, 0.03].[4] Further, the groups did not significantly differ in participants' acceptance of the findings as "true," $F(2, 615)=2.53$, p=.081, $\eta^2<.01$, 95% CI_{η^2} [0.00, 0.03], but they did differ regarding participants' behavioral

[4] Tukey-HSD tests revealed that participants ascribed significantly less expertise in the *extreme* condition than in the *no* self-disclosure condition (p=.034), while the comparisons with the *moderate* condition were non-significant (both: p>.409).

intentions $F(2, 615)=3.57$, $p=.029$, $\eta^2=0.01$, 95% CI_{η^2} [0.00, 0.03].[5] Again, we explored the effect of self-disclosure on perceived closeness: Unlike Studies 1 and 2, we found no effect on perceived closeness, $F(2, 615)=0.62$, $p=.539$, $\eta^2<0.01$, 95% CI_{η^2} [0.00, 0.01]. For optimal comparability to the other studies reported here, we deviated from the preregistration and followed up with Helmert contrasts (i.e., testing the no self-disclosure condition against the combined moderate and the extreme conditions): We did not detect an effect on affective trustworthiness, $t(616)=-1.43$, $p=.153$, $d=-0.12$, 95% CI_d [-0.29, 0.05], but a negative effect on expertise, $t(616)=2.16$, $p=.031$, $d=0.18$, 95% CI_d [0.02, 0.35]. Regarding credibility, there were significant negative effects on both acceptance, $t(441.44)=2.45$, $p=.025$, $d=0.20$, 95% CI_d [0.03, 0.37], and behavioral intentions, $t(616)=2.23$, $p=.026$, $d=0.19$, 95% CI_d [0.02, 0.36]. For perceived closeness, we did not find a significant effect regarding this comparison, $t(616)=-0.42$, $p=.675$, $d=-0.04$, 95% CI_d [-0.20, 0.13].

In Study 3, we thus found no positive effects of self-disclosure on laypeople's perceptions of the researcher's affective trustworthiness but rather a negative effect on expertise and both dimensions of credibility (acceptance and behavioral intentions). Further, we did not detect any inverted u-shaped associations of extent of self-disclosure and our variables of interest. This time, while the manipulation check indicates that participants were aware of the self-disclosure, we did not even replicate the effect on perceived closeness. Considering sample characteristics, there is no clear evidence that these aspects

[5] Tukey-HSD tests revealed that participants reported significantly lower intentions to act in line with the communicated findings in the *moderate* condition compared to the *no* self-disclosure condition ($p=.021$), while the comparisons with the *extreme* condition were non-significant (both: $p>.312$).

might explain the different pattern of results compared to Studies 1 or 2: Participants were older (but age did not correlate with affective trustworthiness, expertise, acceptance or behavioral intentions above r=.09) and they reported less experiences with science in the last 12 months (but still 61% held a university degree and the reported frequency of contact with science was similar to the previous samples). Together, the divergent results between Studies 1 to 3 increased our doubts about a robust effect of self-disclosure on trust in science. Thus, to scrutinize this robustness (i.e., generalizability to other scientific domains), we conducted three additional studies using other science communication scenarios.

Study 4

In a fourth study, we wanted to test the effects from Study 1, 2 and 3 in a different research field (medical context) and communication context (video at an alleged science exhibit). Similar to Studies 1 and 2, we realized only two conditions: self-disclosure and no self-disclosure. However, we also varied the researcher's gender to explore possible interactions with the effect of self-disclosure in science communication (gender biases might play a role for the perceptions of self-disclosure; e.g., Kadji & Schmid Mast, 2021). Study 4 was preregistered (https://osf.io/rtua5).

Methods. Study 4's design was again similar to the previous studies. However, we changed the stimulus material: After providing informed consent and stating their demographics (see Table 5.1), participants viewed a video with subtitles (no sound) that was allegedly shown at a science exhibition for the general public about current health research. In the video, an alleged medical scientist reported current research results about the prevention of skin diseases.

The scientist's name was either implying a female scientist (*"Dr. Michaela Schneider"*) or a male scientist (*"Dr. Michael Schneider"*). Self-disclosure was manipulated via subtitles to avoid confounding effects of different voices: Again, we framed information as either personal (e.g., *"In winter, I get a very dry scalp"*) or neutral (e.g., "*In winter, many people get a dry scalp"*). Participants were randomly assigned to the gender (*male* vs. *female*) × self-disclosure (*self-disclosure* vs. *no self-disclosure*) conditions. Next, we applied our measures 3 (see Table 5.1 and 5.2).

Results and Discussion. We included 330 participants in our analyses (see Table 5.2). The experimental conditions in Study 4 did not differ regarding control variables (engagement frequency: p=.069; engagement experiences: p=.378; pro-science attitude: p=.430; expectation of researchers' disclosure: p=.553) and participants perceived the self-disclosing scientist to be sharing significantly more personal information than the non-self-disclosing scientist, F(1, 326)=184.23, p<.001, η^2=.36 CI_{η^2} [.28, .43], while neither the researcher's gender (p=.350) nor the interaction effect with self-disclosure were significant (p=.646).

Descriptive statistics are reported in Table 5.7. In 2×2 univariate ANOVAs[6], we detected a significant main effect of self-disclosure on affective trustworthiness, F(1, 326)=6.15, p=.014, η_p^2=.02, 95% $CI_{\eta p^2}$ [0.00, 0.06], d=-0.27, 95% CI_d [-0.49, -0.06][7], but not on expertise, F(1, 326)=0.45, p=.503, η_p^2<.01, 95% $CI_{\eta p^2}$ [0.00, 0.02],

[6] In a 2×2 MANOVA, we found a significant multivariate effect of self-disclosure, F(1, 326)=2.98, p=.020, Pillai-V = .04, η_p^2=.04, 95% $CI_{\eta p^2}$ [0.00, 0.07], we did not find any significant multivariate effects of the researcher's gender, F(1, 326)=0.63, p=.644, Pillai-V=.01, η_p^2<.01, 95% $CI_{\eta p^2}$ [0.00, 0.02], nor an interaction, F(1, 326)=0.50, p=.738, Pillai-V=.01, η_p^2<.01, 95% $CI_{\eta p^2}$ [0.00, 0.02].

[7] In Study 4, we calculated Cohen's d based on the two-group-comparisons: self-disclosure vs. no self-disclosure (across researcher's gender).

d=0.07, 95% CI_d [-0.14, 0.29], nor on acceptance, F(1, 326)=0.03, p=.859, η_p^2<.01, 95% $CI_{\eta p^2}$ [0.00, 0.01], d=0.02, 95% CI_d [-0.20, 0.24] or behavioral intentions, F(1, 326)=0.01, p=.930, η_p^2<.01, 95% $CI_{\eta p^2}$ [0.00, 0.01], d<0.01, 95% CI_d [-0.21, 0.23]. None of our outcome variables were influenced by the researcher's gender (all: p>.287) or the interaction (all: p>.245). As before, we also analyzed the effects on perceived closeness: In line with results from Studies 1 and 2, we found a significant main effect of self-disclosure on perceived closeness, F(1, 325)=10.75, p=.001, η_p^2=.03, 95% $CI_{\eta p^2}$ [0.01, 0.08], d=-0.36, 95% CI_d [-0.58, -0.15]. The researcher's gender (p=.951) and the interaction (p=.311) had no significant effects on closeness.

In Study 4, we applied our experimental design to another science-to-public communication scenario with content from medical science and were able to replicate the small positive effect of a researcher's self-disclosure (compared to no self-disclosure) on their perceived affective trustworthiness which we found in Study 1 (but not in Studies 2 or 3). However, we did not replicate the small negative effects on the researcher's perceived expertise and/or credibility of the communicated findings which we found in Study 3 (but not in Studies 1 and 2). Furthermore, we detected an effect of self-disclosure on perceived closeness (like in Studies 1 and 2, but not Study 3).

The researcher's gender did not make any difference for the effects of self-disclosure on any of our outcome variables and it did not have any main effects either. This could mean that researchers' gender does not play a role for the effectiveness of engaging in self-disclosure when communicating their science or for the perception of their trustworthiness and the credibility of their research in general. However, it is likely that many of our participants did not effectively notice the gender manipulation, which was very subtle (44% could not recall the

indicated gender). Thus, these results do not warrant any clear conclusions regarding the effect of researchers' gender in this context.

Study 5

In a fifth study, we manipulated the scientific discipline in which the (self-disclosing vs. not self-disclosing) researcher communicated his or her science directly. After all, stereotypical views about scientists likely pertain to scientist from the so-called "hard" sciences such as physics, biology, or chemistry (e.g., asking for drawings of "scientists", laypeople produce images related to laboratory work; Ferguson & Lezotte, 2020; Miller et al., 2018). Assuming that the effect of self-disclosure in science communication aims at debunking stereotypical perceptions of scientists, we expected the effects of self-disclosure to be more pronounced in the "hard" sciences compared to the social sciences. Further, we again explored possible interaction effects of self-disclosing communication and researchers' gender by using a stronger manipulation than in Study 4. Moreover, we applied our experimental design to another common science-to-public communication situation: social media. Again, this study was preregistered (https://osf.io/yztq2).

Methods. We used a similar design as before. This time, after obtaining informed consent and information about their demographics (see Table 5.1), we asked participants to imagine scrolling through social media when they come across a post by a news channel regarding a recent study on autism-spectrum disorder in children and adolescents. The post consisted of three pictures each showing one statement on the study's main findings by one of the three alleged study authors. We told participants that the study was conducted by an interdisciplinary team from either the "hard" sciences (a biologist, a

neuroscientist, and a pharmacist) or the social sciences (an educationalist, a communication scientist, and a sociologist). In each picture, we manipulated self-disclosure (e.g., *"[…] This result fits with my personal experiences from my practical activity with autistic children and their families"*) or no self-disclosure (e.g., *"[…] This result fits with experiences from practice with autistic children and their families"*). Further, we also randomized the researchers' gender: Either the study was allegedly conducted by an all-female or an all-male research team. This information was conveyed in the introductory text (like in Study 4) but also on every single picture frame by using gendered descriptions common in German language (e.g., "*Soziologin*", which can be translated to "female sociologist"). Then, we applied our measures (see Table 5.1 and Table 5.2).

Results and Discussion. We included 410 participants in our analyses (see Table 5.2). The experimental conditions in Study 5 did not differ regarding control variables (engagement frequency: p=.763; engagement experiences: p=.521; general trust in and credibility of social and "hard" sciences: MANOVA, p=.436; expectations of scientists' disclosure in social sciences: p=.755, and in "hard" sciences: p=.084). Participants perceived the self-disclosing researchers to be sharing significantly more personal information than the non-self-disclosing researchers, $F(1, 402)=197.33$, $p<.001$, $\eta^2=.33$, 95% CI$_{\eta^2}$ [.26, .40], while neither the researchers' disciplines, nor gender, nor any interaction effects were significant (all: $p \geq .128$).

Descriptive statistics are reported in Table 5.8.[8] We analyzed all dependent

[8] Adding researchers' indicated gender as factor, did not change the pattern of the results from the preregistered 2×2 analyses, thus, we only report the 2×2×2 analyses. We also conducted a 2×2×2 MANOVA: We found no main effect of self-disclosure, F(1, 402)=1.09, p=.361, Pillai-V = .01, ηp^2=.01, 95% CIηp^2 [0.00, 0.03]; however, there was a main effect

variables in 2×2×2 ANOVAs. Regarding affective trustworthiness, we found no main effect of self-disclosure, $F(1, 402)=0.04$, $p=.840$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], $d=0.02$, 95% CI_d [-0.17, 0.21][9], no main effect of discipline, $F(1, 402)=3.59$, $p=.074$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.03], but a main effect of researchers' gender, $F(1, 402)=16.02$, $p<.001$, $\eta_p^2=.04$, 95% $CI_{\eta p^2}$ [0.01, 0.08]. There were no significant interaction effects (self-disclosure × discipline, $p=.987$; self-disclosure × gender, $p=.325$; discipline × gender, $p=.721$; self-disclosure × discipline × gender, $p=.507$; all $\eta_p^2<.01$). Similarly, regarding expertise, we found no main effect of self-disclosure, $F(1, 402)=1.10$, $p=.296$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], $d=0.10$, 95% CI_d [-0.09, 0.30], no main effect of discipline, $F(1, 402)=0.04$, $p=.841$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], but a main effect of researchers' gender, $F(1, 402)=6.48$, $p=.011$, $\eta_p^2=.02$, 95% $CI_{\eta p^2}$ [0.00, 0.05]. Again, there were no significant interaction effects (self-disclosure × discipline, $p=.598$; self-disclosure × gender, $p=.925$; discipline × gender, $p=.190$; self-disclosure × discipline × gender, $p=.769$; all $\eta_p^2<.01$). On acceptance, there were no significant main effects of either self-disclosure, $F(1, 402)=0.11$, $p=.916$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], $d=-0.01$, 95% CI_d [-0.20, 0.18], of discipline, $F(1, 402)=0.43$, $p=.514$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02], nor of researchers' gender, $F(1, 402)=4.49$, $p=.071$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.03], and there were no significant interaction effects (self-disclosure × discipline, $p=.224$; self-disclosure × gender, $p=.890$; discipline × gender, $p=.205$; self-disclosure × discipline × gender, $p=.622$; all $\eta_p^2<.01$).

---

of discipline, F(1, 402)=2.45, p=.046, Pillai-V=.02, ηp^2=.02, 95% CIηp^2 [0.00, 0.05] and a main effect of researchers' gender, F(1, 402)=4.33, p=.002, Pillai-V=.04, ηp^2=.04, 95% CIηp^2 [0.01, 0.08]. The interaction effects were all non- significant ($p>.074$; $\eta p^2<.02$).

[9] As in Study 4, we calculated Cohen's d based on the two-group-comparisons: self-disclosure vs. no self-disclosure (across researchers' discipline).

However, on behavioral intentions, while the main effect of self-disclosure, $F(1, 402)=0.65$, $p=.423$, $\eta_p^2<.01$ [0.00, 0.02], and of discipline, $F(1, 402)=0.18$, $p=.674$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], were also not significant, we found a significant main effect of researchers' gender, $F(1, 402)=4.73$, $p=.030$, $\eta_p^2=.01$, 95% $CI_{\eta p^2}$ [0.00, 0.04], and the interaction effect of self-disclosure × discipline, $F(1, 402)=4.06$, $p=.045$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.04]. Once more, the other interaction effects were not significant (self-disclosure × gender, $p=.913$; discipline × gender, $p=.618$; self-disclosure × discipline × gender, $p=.825$; all $\eta_p^2<.01$). Following up on the significant interaction of self-disclosure × discipline, Tukey-HSD post hoc tests did not reveal significant differences between the conditions (all: $p>.194$).

In sum, we found no main effects of self-disclosure on trustworthiness or credibility in Study 5. Results for behavioral intentions suggest that there might be an interaction effect of researchers' self-disclosure and their discipline, but this effect was negligibly small and did not stand up in post hoc tests. Interestingly, participants ascribed more trustworthiness and reported higher behavioral intentions when confronted with female (vs. male) researchers (results remained significant when controlling for participant's gender). However, this should be interpreted carefully because—as indicated by responses on the manipulation checks—72% of our participants did not remember whether or which gender was indicated for the researchers.

Study 6

Like Study 5, Study 6 was designed to follow up on results from Studies 1 to 4. For the sake of consistency, we focus on the effects of self-disclosure on closeness and trustworthiness, deviating from this study's preregistration

(https://osf.io/fa25n). Here, we investigated self-disclosure in another science-to-public communication scenario in social media and extended the scope of this line of research to possible effects of framing the communication scenario as reciprocal (vs. purely unilateral). In dyadic interactions, self-disclosure is mostly understood as a reciprocal process in which both interlocutors take turns disclosing personal information (e.g., Cozby, 1972; Sprecher, Treger, Wondra, et al., 2013). This is, of course, not the case in most science communication formats. Thus, the opportunity to interact with the researcher via social media might strengthen the effects of a researcher's self-disclosure on laypeople's trust in science. To test this assumption, we manipulated the opportunity for an interpersonal exchange with the researcher.

Methods. After obtaining informed consent and collecting demographics (see Table 5.1), we presented screenshots of real Twitter posts on the topic of stress research from a big science communication channel and informed participants that, on this account, every week another researcher tweets about their scientific work. Then, we manipulated the opportunity to interact with the researcher: We either told participants that the channel was only meant for informational purposes and no exchange was possible due to a deactivated comment section (i.e., *no exchange*) or that the channel was meant for exchange and one could converse with the researchers in the comment section (i.e., *exchange*). Further, we manipulated self-disclosure by showing participants real tweets either including original self-disclosing content (e.g., *"For example, I am more focused when stressed"*) or without that content. We then measured the dependent and control variables (see Table 5.1 and Table 5.2).

Results and Discussion. We included 389 participants in our analyses (see

Table 5.2). The experimental groups neither differed with regard to their general engagement with science (engagement frequency: p=.492; engagement experiences: p=.554), nor their general expectation of scientists' disclosure (p=.213). Again, our self-disclosure manipulation led participants to perceive the self-disclosing scientist to be sharing significantly more personal information than the neutral scientist, $F(1, 385)=153.42$, $p<.001$, $\eta^2=.21$, 95% CI_{η^2} [.15, .28], while neither the exchange main effect (p=.101) nor the self-disclosure × exchange interaction (p=.186) were significant.

Descriptive statistics are reported in Table 5.9. On affective trustworthiness, there was neither a main effect of self-disclosure, $F(1, 385)=0.36$, p=.552, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02], d=0.06, 95% CI_d [-0.14, 0.26],[10] nor of exchange, $F(1, 385)=0.01$, p=.919, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], or an interaction effect, $F(1, 385)=1.31$, p=.252, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02]. However, on expertise, we found a main effect of self-disclosure, $F(1, 385)=11.76$, $p<.001$, $\eta_p^2=.03$, 95% $CI_{\eta p^2}$ [0.01, 0.07], d=-0.35, 95% CI_d [-0.55, -0.15]; yet again, no effects of exchange, $F(1, 385)=0.49$, p=.486, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02], or an interaction effect, $F(1, 385)=0.05$, p=.822, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01]. Regarding closeness, neither self-disclosure, $F(1, 385)=0.567$, p=.452, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02], d=0.08, 95% CI_d [-0.12, 0.28]; nor exchange, $F(1, 385)=1.62$, p=.204, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.03], or the interaction, $F(1, 385)=0.08$, p=.778, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], had significant effects. The same pattern emerged for participants' willingness to engage with science: The effects of self-disclosure, $F(1, 385)=0.56$, p=0.454, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02], d=-0.08, 95% CI_d [-0.27, 0.12],

[10] As in Study 4 and 5, we calculated Cohen's d based on the two-group-comparisons: self-disclosure vs. no self-disclosure (across exchange).

exchange, $F(1, 385)=0.34$, $p=.562$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.02], and the interaction, $F(1, 385)=0.03$, $p=.873$, $\eta_p^2<.01$, 95% $CI_{\eta p^2}$ [0.00, 0.01], were not significant.

Thus, in Study 6, a communicating researcher's self-disclosure did not have significant trust-increasing effects (on affective trustworthiness) among laypeople but rather decreased perceived trustworthiness by lowering their ascriptions of the researcher's expertise. Further, while self-disclosure led to higher perceptions of the researcher's warmth, we did not find the effect on perceived closeness (as in Studies 1, 2, and 4). There were no effects of self-disclosure on perceptions of the researcher's competence or participants' willingness to further engage with the researcher and the communicated science.

Framing the science communication scenario (i.e., real science communication channel on Twitter) as an opportunity for exchange did not make any difference for ascriptions of trustworthiness, closeness, warmth, competence or even for participants' willingness to further engage with the researcher and their work. However, it is likely that our manipulation was not effective enough, as participants did not really have the opportunity to interact with the researcher. Further, data on the respective manipulation check also suggests that the exchange manipulation in this study might have been too subtle (36% did not correctly remember the manipulation). Thus, these findings regarding opportunity for exchange and its interaction with self-disclosure in science communication should be taken as merely preliminary.

Mini Meta-Analysis

Reflecting the overall picture, some findings from Study 1 to 6 support a possible trade-off effect: In some cases, self-disclosing personal information in

science communication led laypeople to feel significantly closer to the communicating researchers (Studies 1, 2 and 4) and they sometimes even ascribed higher affective trustworthiness to them (Studies 1 and 4); yet, in other cases, self-disclosure led to decreases in perceptions of expertise (Studies 3 and 6) or even less credibility ascribed to the researchers' findings (Study 3). However, even though our samples and manipulations were rather similar across most of the studies presented here, our results and effect sizes vary considerably. Thus, we decided to integrate our findings from these six studies in a fixed-effects mini meta-analysis (assuming the existence of one "true" population effect of self-disclosure; Goh et al., 2016; Viechtbauer, 2010).

We calculated standardized mean effect sizes for closeness, affective trustworthiness, expertise, acceptance and behavioral intentions in Studies 1-6 weighted by sample sizes. The meta-analytic results are summarized in forest plots in Figures 5.1 to 5.5. Regarding closeness, the estimated average standardized difference was significant in the expected direction, $\hat{\theta}$=-0.17, 95% $CI_{\hat{\theta}}$ [-0.26, -0.08][11], z=–3.69, p<.001. Similarly, regarding affective trustworthiness, the estimated average standardized difference was also significant, $\hat{\theta}$=-0.14, 95% $CI_{\hat{\theta}}$ [-0.22, -0.06], z=–3.36, p<.001. Looking at the meta-analytical effect on expertise, the direction reversed: The estimated average standardized difference was significant, $\hat{\theta}$=0.15, 95% $CI_{\hat{\theta}}$ [0.07, 0.23], z=3.63, p<.001. The average standardized differences regarding both credibility dimensions were non-significant: Acceptance, $\hat{\theta}$=0.08, 95% $CI_{\hat{\theta}}$ [-0.02, 0.17], z=1.54, p=.122, and

[11] Note, due to the coding of conditions a negative numerical effect means a positive effect of self-disclosure on the respective outcome variable.

behavioral intentions, $\hat{\theta}$=0.05, 95% $CI_{\hat{\theta}}$ [-0.06, 0.14], z=0.80, p=.423.

In sum, the mini meta-analysis thus supports the trade-off assumption for self-disclosure in science communication (albeit with small effect sizes): Overall, participants felt closer to self-disclosing scientists and ascribed higher affective trustworthiness to them. At the same time, they perceived these scientists to have less expertise. The credibility of the communicated scientific findings (i.e., acceptance of a finding as "true" and behavioral intentions to act in line with these findings) was, meta-analytically, unaffected by the researchers' self-disclosure.

Field Experiment

In our highly controlled online studies, the observed effects were small, but the mini meta-analysis suggests that they are measurable notwithstanding. This raises the question whether such self-disclosure can actually have practical relevance for "real" science communication. Thus, we applied our theoretical rational to a field setting in a large German science museum. Additionally, we wanted to scrutinize whether the effects of self-disclosure on trust in science were specific to researchers (i.e., dependent on stereotypes about scientists) or more generalizable to other science communicators (e.g., science museum guides). Therefore, we developed an audio guide for an exhibition on historic and current marine research, in which we varied self-disclosure (*self-disclosure* vs. *no self-disclosure*) and role of the speaker (*marine researcher* vs. *museum guide*). Additionally, there was a neutral, impersonal audio tour ("control" condition).

Methods. Participants were recruited beforehand and booked a date for their museum visit or were approached spontaneously while visiting the museum (see Table 5.2 for further details). Testing took place on weekdays (Thursdays and Fridays) and on weekends between October 2020 and November 2021.

Participants were welcomed to the museum and instructed on the study procedure. After they gave their informed consent, they walked through the exhibition by themselves using their own smartphones and headphones to play the audio guide. The audio guide was app-based, took about 15-20 minutes and contained short interactive elements in between audio tracks (e.g., little quizzes). Participants could proceed the tour at their own pace and were randomly allocated to one of five conditions: There was a neutral control condition with no personalized elements (i.e., no introduction of speaker, no personal addressing of the listener). The other four conditions resulted from orthogonally varying *self-disclosure* (i.e., some, mostly illustrative information in the audio text was framed as either personal or neutral information, see Study 1-5) in combination with the role of the *guide* (i.e., either the guide introduced herself as marine scientists or as museum guide and repeatedly mentioned her role throughout the tour; e.g., "*us researchers…*" vs. "*us museum guides…*"). In all audio tours, we kept the presented content, the interactive elements, and the audio voice constant (i.e., same speaker).

After participants completed the audio tour, they came back for further instructions and received a link to the online questionnaire. Data from the questionnaire were matched with data about the experimental condition by asking for a specific codeword (however, questionnaire links were also marked with an identifier to facilitate correct matching). Next, we measured dependent, control and exploratory variables (see Table 5.1 and Table 5.2). After completing the questionnaire, participants were fully debriefed. For this study, participants were compensated by free entry to the museum at the day of their study participation or, if visitors participated spontaneously, they gained a free ticket for another day.

Results and Discussion. We included 480 participants in our analyses (see Table 5.2). The experimental groups neither differed with regard to their general engagement with science (engagement frequency: p=.108; engagement experiences: p=.582), nor their general expectation of scientists' (p=.493) or museum guides' (p=.088) disclosure. Our self-disclosure manipulation led participants to perceive the self-disclosing audio speakers to be sharing significantly more personal information than the non-self-disclosing speakers and the non-self-disclosing speakers to be sharing significantly more than the control audio speaker (control < no self-disclosure < self-disclosure; all: p<.001), F(2, 473)=277.00, p<.001, η^2=.54, 95% CI_{η^2} [.48, .59].

Descriptive statistics are reported in Table 5.10. We were interested in four specific contrasts: 1) control versus experimental conditions; 2) self-disclosure versus no self-disclosure conditions (i.e., a main effect of self-disclosure); 3) scientist versus museum guide conditions (i.e., a main effect of guide), 4) self-disclosing scientist condition versus the three other experimental conditions (i.e., an ordinal self-disclosure × guide interaction effect). For the dependent variables liking (all: p>.344) and interest (all: p>.344) as well as closeness (all: p>.085), affective trustworthiness (all: p>.189), acceptance (all: p>.753), and behavioral intentions (all: p>.503), these contrasts were all not significant. Only for expertise, contrast 1 significantly differed from zero, t(475)=2.31, p=.022, d=0.21, 95% CI_{η^2} [0.03, 0.39]; indicating that participants ascribed higher expertise to the audio guide in the experimental conditions than in the control condition. This effect was however small (and testing it on an adjusted significance level would render it non-significant). The other contrasts for expertise were not significant (all: p>.053).

Together, these results suggest that there is little value (or risk) to

disclosing personal information in science communication: Neither self-disclosure nor role of guide (i.e., scientist or museum guide) or their interaction had any significant effects on laypeople's enjoyment of the tour (i.e., liking) or how much it sparked their interest; it did not change their trustworthiness ascriptions to the speaker (affective trustworthiness and expertise) or their perceptions of the credibility of marine research (acceptance and usefulness of findings). In fact, the only effect we found was a small effect on expertise regarding the comparison of the control with the experimental conditions: Introducing the audio speaker with a name and photo and personally addressing the listener lead to higher expertise ascriptions than not "putting a face" to the audio voice. Engaging in more self-disclosure (vs. no self-disclosure) did not increase this effect, suggesting that it might be enough to provide some superficial personal information in such a science communication scenario.

This field data points to a lack of practical relevance of self-disclosure in science communication. The realization of our stimulus material as audio guide through a real exhibition in a science museum presents a very authentic application of self-disclosure in science communication and was developed in close cooperation with museum experts. Thus, our field study has very high ecological validity. However, naturally, there are limitations to this experimental field data that might also account for the non-significant effects. Most importantly, participants listened to the audio guide unsupervised in the exhibition by themselves and, thus, we cannot ensure their unrestricted concentration. In line with this, participants reported having been distracted by other museum visitors or other exhibition objects which were not part of the tour. This is also reflected in some of the data; for example, many participants in the museum guide condition

(34%) mistook the guide for a marine researcher although the speaker's role was stated many times in the audio (analyzing only results of participants who correctly identified the guide's role did not change the pattern of results). This shows parallels to an earlier study in which objects exhibited in the museum were presented to the visitors either as original objects or as replicas (Hampp & Schwan, 2015). Here, it was shown that the majority of participants (63%) could no longer say whether the highlight object shown was an original or a replica shortly after visiting the exhibition ("Most of them did not question the status of the objects as they considered the display of originals to be customary in the museum.", p. 170, Hampp & Schwan, 2015). In general, museums are perceived by the public as highly trustworthy institutions, ranking second after friends and family, and are significantly more trustworthy than, for example, news media or the government (Wilkening, 2021). Therefore, it can be assumed, that visitors expect original objects and credible information in museums, granting only subordinate importance to the evaluation of the knowledge sources in the museum (e.g., the audio guide). Thus, it might not be of great relevance whether the information is conveyed by a real scientist or by a museum guide. However, one might argue that such distractions and presumptions are a realistic part of science communication in practice. Therefore, our findings from this field study question whether self-disclosure in science communication can actually have practical relevance in an applied setting like a science museum for outcomes like enjoyment and interest or trust in researchers and scientific findings.

General Discussion

In this article, we present six experimental online studies, a mini meta-analysis, and a field experiment investigating the effects of researchers' self-

disclosure in science communication on laypeople's perceptions of researchers and their findings, namely feelings of closeness towards researchers, trustworthiness ascriptions, and the perceived credibility of their findings. Taken together, our results suggest that self-disclosure in science communication has measurable, yet small—and, thus, practically irrelevant—effects: The meta-analysis shows evidence of small positive effects on laypeople's closeness towards self-disclosing researchers and on their affective trustworthiness (i.e., integrity and benevolence) ascriptions. However, at the same time, self-disclosure can also affect laypersons' perceptions about researchers' expertise. Thus, self-disclosure might bring researchers down from their stereotypical distanced ivory tower but, by doing so, it also makes them seem less like the stereotypical smart geniuses. Additionally, self-disclosure did, overall, not affect the perceived credibility of the communicated scientific findings, the effects on trustworthiness were small and volatile across all studies, and the lack of significant findings from the field study further question whether the effects of self-disclosure are actually meaningful when put into practice.

We also considered some contextual factors for self-disclosure in science communication: researchers' gender (Studies 4 and 5), researchers' discipline (Study 5), opportunity for exchange (Study 6) and role of discloser (Field Study). However, these factors did not influence the effects of self-disclosure. Together, it seems self-disclosure in and of itself is not very promising as a booster for public trust in science or even as a remedy for the relationship with science-skeptical groups. It is likely that, in science communication, the content of the information revealed is much more important than the act of self-disclosure on its own.

Looking at our methods and designs, we tried to keep most of the

manipulations of self-disclosure free of aspects that might just signal warmth (e.g., photo from one's window vs. photos of one's kids) in order to test the basic effect of disclosing any personal information. We did this in order to test our hypotheses conservatively. That said, sharing more inherently "warmth-related" personal details about oneself could actually lead to more consistent positive responses: For example, Zahry and Besley (2021) recently showed that scientists directly communicating prosocial intentions and emotional appeals were perceived as warmer than those using neutral appeals and Jarreau et al. (2019) demonstrated that researchers sharing selfies on social media were perceived as warmer and more trustworthy. Another kind of self-disclosure that seems to be received quite well in the scientific context is related to personal thoughts and feelings towards one's own work: For example, researchers disclosing doubts about past findings and intentions to reform their future work routines were perceived as more likable and trustworthy and their work as more credible (Altenmüller, Nuding, et al., 2021). Likely, such science-related self-disclosure is perceived as more appropriate as it still relates to the researchers' professional sphere.

Interestingly, inspecting the correlations across all studies (see "supplementary tables," https://osf.io/s39u2/), perceived similarity to a researcher was similarly related to trustworthiness and credibility as perceived closeness. This suggests that self-disclosure could be more effective when it concerns details that make it easier for laypeople to find similarities (and not only closeness) between themselves and researchers (e.g., Meijnders et al., 2009; Zorn et al., 2022): For example, researchers could reveal personal memories of their time at school when talking to teenagers at a school event or they could reveal personal experiences as parents of teenagers when talking to the high schoolers' parents.

However, revealing sensitive personal information is always a risk (c.f., Omarzu, 2000), especially when communicating potentially polarizing scientific findings. Thus, self-disclosure might pose other challenges than "only" appearing unprofessional: For example, disclosing being personally affected by one's own research (i.e., doing "me-search") has been shown to have the potential to increase as well as decrease researchers' trustworthiness and the credibility of their findings (Altenmüller, Lange, et al., 2021): Laypeople use this information to adapt their perceptions of such researchers and their findings in line with pre-existing attitudes (i.e., motivated science reception). Even more problematically, personal information might be used for threats and attacks against researchers (e.g., Nogrady, 2021). Thus, researchers should consider well which details about their lives they are willing to disclose to the public.

Conclusion

Coming back to Marlene's guinea pigs Noxi and Neri, our findings suggest that starting this article with such irrelevant and purely illustrative insights into one of the authors' lives had likely little more effect than eliciting a small smile or eye-roll: Our results across six online studies and a field experiment, suggest that engaging in self-disclosure when communicating science can somewhat increase laypeople's feelings of closeness and perceptions of researchers' warmth-related trustworthiness (i.e. integrity and benevolence), but at the same time, decrease perceptions of competence-related trustworthiness (i.e., expertise), while the credibility of the communicated science remains unaffected. However, the effects of such general self-disclosure were small and inconsistent and are likely not relevant for practical application. Future research should focus on the effects of revealing more specific information (e.g., similarity cues). More practically,

science communicators cannot rely on positive effects of disclosing any personal information and should consider well whether and which personal details they want to share with their audience.

References

Abele-Brehm, A., & Wojciszke, B. (2018). *The Agency—Communion Framework.* (W. B. Abele A.E., Ed.). Routledge.

Altenmüller, M. S., Lange, L. L., & Gollwitzer, M. (2021). When research is me-search: How researchers' motivation to pursue a topic affects laypeople's trust in science. *PLOS ONE*, *16*(7), e0253911. https://doi.org/10.1371/journal.pone.0253911

Altenmüller, M. S., Nuding, S., & Gollwitzer, M. (2021). No harm in being self-corrective: Self-criticism and reform intentions increase researchers' epistemic trustworthiness and credibility in the eyes of the public. *Public Understanding of Science*, *30*(8), 09636625211022181. https://doi.org/10.1177/09636625211022181

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. Holt, Rinehart & Winston.

Appelman, A., & Sundar, S. S. (2016). Measuring Message Credibility: Construction and Validation of an Exclusive Scale. *Journalism & Mass Communication Quarterly*, *93*(1), 59–79. https://doi.org/10.1177/1077699015606057

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*(4), 596–612. https://doi.org/10.1037/0022-3514.63.4.596

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The Experimental Generation of Interpersonal Closeness: A Procedure and Some Preliminary Findings. *Personality and Social Psychology Bulletin*, *23*(4), 363–377. https://doi.org/10.1177/0146167297234003

Asbrock, F. (2010). Stereotypes of Social Groups in Germany in Terms of Warmth and Competence. *Social Psychology*, *41*(2), 76–81. https://doi.org/10.1027/1864-9335/a000011

BBVA Foundation. (2011). *International Study on Scientific Culture. Understanding of Science.* BBVA.

Benson-Greenwald, T. M., Trujillo, A., White, A. D., & Diekman, A. B. (2021). Science for Others or the Self? Presumed Motives for Science Shape Public Trust in Science. *Personality and Social Psychology Bulletin*, 01461672211064456. https://doi.org/10.1177/01461672211064456

Caltabiano, M. L., & Smithson, M. (2010). Variables Affecting the Perception of Self-Disclosure Appropriateness. *The Journal of Social Psychology*. https://www.tandfonline.com/doi/abs/10.1080/00224545.1983.9712017

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, *116*(3), 457–475. https://doi.org/10.1037/0033-2909.116.3.457

Cozby, P. C. (1972). Self-Disclosure, Reciprocity and Liking. *Sociometry*, *35*(1), 151–160. https://doi.org/10.2307/2786555

Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin*, *79*(2), 73–91. https://doi.org/10.1037/h0033950

Ferguson, S. L., & Lezotte, S. M. (2020). Exploring the state of science stereotypes: Systematic review and meta-analysis of the Draw-A-Scientist Checklist. *School Science and Mathematics*, *120*(1), 55–65. https://doi.org/10.1111/ssm.12382

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Dupree, C. (2014). Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences*, *111*(Supplement_4), 13593–13597. https://doi.org/10.1073/pnas.1317505111

Garner, J., & Rossmanith, E. (2021). Using Participation and Empathy to Inspire Positive Change: A Transcontinental Conversation. *Journal of Museum Education*, *46*(1), 48–60. https://doi.org/10.1080/10598650.2020.1852491

Gelso, C. J., & Palma, B. (2011). Directions for research on self-disclosure and immediacy: Moderation, mediation, and the inverted U. *Psychotherapy*, *48*(4), 342–348. https://doi.org/10.1037/a0025909

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, *10*(10), 535–549. https://doi.org/10.1111/spc3.12267

Hampp, C., & Schwan, S. (2015). The Role of Authentic Objects in Museums of the History of Science and Technology: Findings from a visitor study. *International Journal of Science Education, Part B*, *5*(2), 161–181. https://doi.org/10.1080/21548455.2013.875238

Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOS ONE*, *10*(10), e0139309. https://doi.org/10.1371/journal.pone.0139309

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Trust in Science and the Science of Trust. In B. Blöbaum (Ed.), *Trust and Communication in a Digitized World* (pp. 143–159). Springer International Publishing. https://doi.org/10.1007/978-3-319-28059-2_8

Huang, L.-S. (2015). Trust in product review blogs: The influence of self-disclosure and popularity. *Behaviour & Information Technology*, *34*(1), 33–44. https://doi.org/10.1080/0144929X.2014.978378

Jarreau, P. B., Cancellare, I. A., Carmichael, B. J., Porter, L., Toker, D., & Yammine, S. Z. (2019). Using selfies to challenge public stereotypes of scientists. *PLOS ONE*, *14*(5), e0216625. https://doi.org/10.1371/journal.pone.0216625

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between

judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. https://doi.org/10.1037/0022-3514.89.6.899

Kadji, K., & Schmid Mast, M. (2021). The effect of physician self-disclosure on patient self-disclosure and patient perceptions of the physician. *Patient Education and Counseling*, *104*(9), 2224–2231. https://doi.org/10.1016/j.pec.2021.02.030

Kim, T.-Y., Lee, D.-R., & Wong, N. (2016). Supervisor Humor and Employee Outcomes: The Role of Social Distance and Affective Trust in Supervisor. *Journal of Business & Psychology*, *31*(1), 125–139. https://doi.org/10.1007/s10869-015-9406-9

Losh, S. C. (2010). Stereotypes about scientists over time among US adults: 1983 and 2001. *Public Understanding of Science*, *19*(3), 372–382. https://doi.org/10.1177/0963662508098576

McAllister, D. J. (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. *Academy of Management Journal*, *38*(1), 24–59. https://doi.org/10.5465/256727

Meijnders, A., Midden, C., Olofsson, A., Öhman, S., Matthes, J., Bondarenko, O., Gutteling, J., & Rusanen, M. (2009). The Role of Similarity Cues in the Development of Trustin Sources of Information About GM Food. *Risk Analysis*, *29*(8), 1116–1128. https://doi.org/10.1111/j.1539-6924.2009.01240.x

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Nogrady, B. (2021). 'I hope you die': How the COVID pandemic unleashed attacks on scientists. *Nature*, *598*, 250–253.

Omarzu, J. (2000). A disclosure decision model: Determining how and when individuals will self-disclose. *Personality and Social Psychology Review*, *4*(2), 174–185. https://doi.org/10.1207/S15327957PSPR0402_05

Pan American Health Organization. (2020). *Communicating about Vaccine Safety: Guidelines to help health workers communicate with parents, caregivers, and patients.*

Phelan, S., Bauer, J., & Lewalter, D. (2018). Visit motivations: Development of a short scale for comparison across sites. *Museum Management and Curatorship*, *33*(1), 25–41. https://doi.org/10.1080/09647775.2017.1389617

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2017). A Short Scale for Assessing the Big Five Dimensions of Personality: 10 Item Big Five Inventory (BFI-10). *methods, data*, 17 Pages. https://doi.org/10.12758/MDA.2013.013

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, *41*(1), 307–364. https://doi.org/10.1002/aris.2007.1440410114

Rosman, T., & Merk, S. (2021). Teacher's Reasons for Trust and Distrust in Scientific Evidence: Reflecting a "Smart But Evil" Stereotype? *AERA Open*, *7*, 233285842110285. https://doi.org/10.1177/23328584211028599

Rutjens, B. T., & Heine, S. J. (2016). The Immoral Landscape? Scientists Are Associated with Violations of Morality. *PLOS ONE*, *11*(4), e0152798. https://doi.org/10.1371/journal.pone.0152798

Specht, I. (2013). *Wahrnehmung und Verarbeitung kontroverser Informationen im Museum*. Technische Universität München.

Sprecher, S., Treger, S., & Wondra, J. D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. *Journal of Social and Personal Relationships*, *30*(4), 497–514. https://doi.org/10.1177/0265407512459033

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology*, *49*(5), 860–866. https://doi.org/10.1016/j.jesp.2013.03.017

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*, 1–48. https://doi.org/10.18637/jss.v036.i03

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134–144. https://doi.org/10.1002/asi.10016

Wheeless, L. R., & Grotz, J. (1977). The measurement of trust and its relationship to self-disclosure. *Human Communication Research*, *3*(3), 250–257. https://doi.org/10.1111/j.1468-2958.1977.tb00523.x

Wilkening, S. (2021). *Museums and Trust*.

Wissenschaft im Dialog/Kantar. (2020). *Wissenschaftsbarometer 2020*.

Yzerbyt, V. Y., Kervyn, N., & Judd, C. M. (2008). Compensation Versus Halo: The Unique Relations Between the Fundamental Dimensions of Social Judgment. *Personality and Social Psychology Bulletin*, *34*(8), 1110–1123. https://doi.org/10.1177/0146167208318602

Zahry, N. R., & Besley, J. C. (2021). Can scientists communicate interpersonal warmth? Testing warmth messages in the context of science communication. *Journal of Applied Communication Research*, *49*(4), 387–405. https://doi.org/10.1080/00909882.2021.1880017

Zhang, S., Shi, Q., & Hao, S. (2009). The appropriateness of teacher self-disclosure: A comparative study of China and the USA. *Journal of Education for Teaching*, *35*(3), 225–239. https://doi.org/10.1080/02607470903091278

Zorn, T. J., Mata, A., & Alves, H. (2022). Attitude similarity and interpersonal liking: A dominance of positive over negative attitudes. *Journal of Experimental Social Psychology*, *100*, 104281. https://doi.org/10.1016/j.jesp.2021.104281

Table 5.1

Overview of all measures in Study 1-6 and the Field Study

Nr	Measured variables	Scale	Reliability	Source
Study 1	Demographics (age, gender, occupation)			
	Trustworthiness (expertise and affective trustworthiness)*	14 opposite adjective pairs; 7-point bipolar scale	Expertise: α=.95; affective trustworthiness: α=.92	Hendriks et al., 2015
	Warmth and competence	6 items, 4 distractors; 6-point scale	Competence: α=.95; warmth: α=.92	Asbrock, 2010
	Manipulation check (self-disclosure)	1 item; 5-point scale		
	Perceived closeness			Aron et al., 1992
	General public engagement with science (PES): engagement frequency & experiences	5 items; 5-point scale; and 15-item multiple choice	Frequency: α=.70	BBVA Foundation, 2011
	General expectation of scientists' disclosure (GED)	2 items, 7 distractors; 5-point scale		
	Demographics (academic discipline)			
Study 2	Perceived closeness and similarity	1 item each, 6-point scale		
	Trustworthiness (expertise and affective trustworthiness)	14 opposite adjective pairs, 7-point bipolar scale	Expertise: α=.94; affective trustworthiness: α=.89	Hendriks et al., 2015
	Credibility (acceptance and behavioral intentions)	2 items, 5-point scale		

	Pro-science attitude	3 items, 5-point scale	α=.58	
	PES: engagement frequency & experiences	5 items, 5-point scale; 15-item multiple choice	Frequency: α=.69	BBVA Foundation, 2011
	Manipulation check (self-disclosure)	1 item, 5-point scale		
	GED	1 item, 6 distractors, 5-point bipolar scale		
	Demographics (age, gender, occupation, academic discipline, political orientation)	Political orientation: 7-point bipolar scale		
Study 3	Perceived closeness and similarity	1 item each, 6-point scale		
	Trustworthiness (expertise and affective trustworthiness)	14 opposite adjective pairs, 6-point bipolar scale	Expertise: α=.91; affective trustworthiness: α=.86	Hendriks et al., 2015
	Credibility (acceptance and behavioral intentions)	2 items, 6-point scale		
	Pro-science attitude	3 items, 6-point scale	α=.59	
	PES: engagement frequency & experiences	5 items, 5-point scale; 15-item multiple choice	Frequency: α=.73	BBVA Foundation, 2011
	Manipulation check (self-disclosure)	1 item, 6-point scale		
	GED	1 item, 8 distractors, 6-point bipolar scale		
	Demographics (age, gender, occupation, academic discipline)			

Study 4	Demographics (age, gender, occupation, academic discipline)			
	Attention check	4 items, multiple choice		
	Perceived closeness and similarity	1 item each, 6-point scale		
	Trustworthiness (expertise and affective trustworthiness)	14 opposite adjective pairs, 6-point bipolar scale	Expertise: α=.92; affective trustworthiness: α=.91	Hendriks et al., 2015
	Credibility (acceptance and behavioral intentions)	3 items each, 6-point scale	Acceptance: α=.78; behavioral intention: α=.71	
	Manipulation check (self-disclosure)	1 item, 6-point scale		
	Manipulation check (gender)	3 items, single choice		
	Pro-science attitude	3 items, 6-point scale	α=.59	
	PES: engagement frequency & experiences	5 items, 5-point scale; 15-item multiple choice	Frequency: α=.78	BBVA Foundation, 2011
	GED	2 item, 7 distractors, 6-point bipolar scale	α=.55	
	Use-me item	1 item, yes/no		
Study 5	Demographics (age, gender, occupation, academic discipline)			
	Attention check	3 items single choice		
	Trustworthiness (expertise and affective trustworthiness)	14 opposite adjective pairs, 6-point bipolar scale	Expertise: α=.94; affective trustworthiness: α=.93	Hendriks et al., 2015

	Credibility (acceptance and behavioral intentions)	3 items each, 6-point scale	Acceptance: α=.83; behavioral intention: α=.63	
	Manipulation check (discipline)	2 items, single choice		
	Manipulation check (gender)	5 items, single choice		
	Manipulation check (self-disclosure)	1 item, 6-point scale		
	Typicality of disciplines	6 items, 6-point bipolar scale		
	GED social & „hard" sciences	2 item and 7 distractors each, 6-point bipolar scales	Social sciences: α=.55; "hard" sciences: α=.41	
	General trust in social & "hard" sciences	7 items each, 6-point scales	Social sciences: α=.77; "hard" sciences: α=.66	Wissenschaft im Dialog/Kantar, 2020
	General credibility of social & "hard" sciences	5 items each, 6-point scales	Social sciences: α=.62; "hard" sciences: α=.55	
	PES: engagement frequency & experiences	5 items, 5-point scale; 15-item multiple choice	Frequency: α=.71	BBVA Foundation, 2011
	Use-me item	1 item, yes/no		
Study 6	Demographics (age, gender, occupation, academic discipline)			
	Attention check	4 items, single choice		
	Warmth and competence	5 items each, 6-point scale	Competence: α=.76; warmth: α=.81	Asbrock, 2010

	Trustworthiness (expertise and affective trustworthiness)	14 opposite adjective pairs, 6-point bipolar scale	Expertise: α=.92; affective trustworthiness: α=.91	Hendriks et al., 2015
	Willingness to further engage with the content	9 items, 6-point scale	α=.92	Altenmüller, Nuding, et al., 2021
	PES: engagement frequency & experiences	5 items, 5-point scale; 15-item multiple choice	Frequency: α=.80	BBVA Foundation, 2011
	GED	2 item, 7 distractors, 6-point bipolar scale	α=.51	
	Manipulation check (exchange)	4 items, single choice		
	Manipulation check (self-disclosure)	1 item, 6-point scale		
	Use-me question	1 item, yes/no		
Field Study	Liking	6-point visual smiley scale		
	Interest	6 items, 6-point scale	α=.81	Specht, 2013
	Perceived closeness and similarity	1 item each, 6-point scale		
	Trustworthiness (expertise and affective trustworthiness)	14 opposite adjective pairs, 6-point bipolar scale	Expertise: α=.91; affective trustworthiness: α=.91	Hendriks et al., 2015
	Credibility (acceptance and behavioral intentions)	3 items each, 6-point scale	Acceptance: α=.71; behavioral intention: α=.55	
	Planning of visit	3 items, single choice		
	Visit situation	7 items, multiple choice		

Visit motivation	18 items, 5-point scale		Phelan et al., 2018
Museum visit habits	2 items, single choice		
First visit	1 item, yes/no		
PES: engagement frequency & experiences	5 items, 5-point scale; 15-item multiple choice	Frequency: α=.75	BBVA Foundation, 2011
GED (scientists & museum guides)	Each: 2 item, 7 distractors, 6-point bipolar scale	Scientists: α=.60; museum guides: α=.62	
Manipulation check (self-disclosure)	1 item, 6-point scale		
Manipulation check (guide)	3 items, single choice		
Speakers' familiarity	3 items, yes/no/don't know		
Demographics (age, gender, occupation, academic discipline)			
Big Five personality	10 items, 5-point scale		Rammstedt et al., 2017
Openness	12 items, 5-point scale	α=.81	Danner et al., 2016
Use-me item	1 item, yes/no		
Feedback	open-ended text field		

Note: Measured variables are sorted in order of presentation in each study. Reliability is Cronbach's α. Full materials, including all original German items and English translations, are provided as online supplement (https://osf.io/s39u2/).
* Mirroring previous findings in which we used the METI (Altenmüller, Lange, et al., 2021; Altenmüller, Nuding, et al., 2021), we used a two-factor solution, collapsing the facets integrity and benevolence (due to high intercorrelations, *r*=.81, and factor analysis suggesting two factors) into one variable which we refer to as "affective trustworthiness", (factor 1) and expertise (factor 2).

Table 5.2

Overview of study characteristics, Study 1-6 and Field Study

	Study 1	**Study 2**	**Study 3**	**Study 4**	**Study 5**	**Study 6**	**Field Study**
Conducted in	May – Jun 2019	Oct 2019	Feb – Apr 2020	Oct 2020	Apr – May 2021	Apr – May 2021	Oct 2020 – Nov 2021
Sample Size	323	344	618	330	410	389	480
Sample gender	60% female, 39% male, 1% other	71% female, 28% male, 0.3% other	65% female, 35% male, 0.3% other	72% female, 28% male, 0.6% other	70% female, 29% male, 0.7% other	67% female, 32% male, 1% other	62% female, 36% male, 2% other
Sample age	*M* = 29.36, *SD* = 9.24, range: 17-68	*M* = 24.39, *SD* = 8.46, range: 18-80	*M* = 43.24, *SD* = 14.34, range: 16-78	*M* = 33.53, *SD* = 15.86, range: 16-78	*M* = 30.70, *SD* = 13.94, range: 16-86	*M* = 29.15, *SD* = 13.38, range: 16-79	*M* = 27.76, *SD* = 11.30, range: 16-73
Sample education	84% currently studying or already holding university degree (35% social sciences, 13% natural sciences, 17% humanities, 11% life sciences, 11% engineering, 13% other)	96% currently studying or already holding university degree (47% social sciences, 15% natural sciences, 14% humanities, 12% life sciences, 4% engineering, 8% other)	77% currently studying or already holding university degree (61% social sciences, 15% natural sciences, 14% humanities, 12% life sciences, 4% engineering, 8% other)	77% currently studying or already holding university degree (50% social sciences, 12% natural sciences, 15% humanities, 8% life sciences, 7% engineering, 8% other)	81% currently studying or already holding university degree (39% social sciences, 13% natural sciences, 14% humanities, 13% life sciences, 8% engineering, 13% other)	77% currently studying or already holding university degree (46% social sciences, 14% natural sciences, 14% humanities, 10% life sciences, 8% engineering, 9% other)	83% currently studying or already holding university degree (36% social sciences, 17% natural sciences, 13% humanities, 12% life sciences, 13% engineering, 10% other)

Design	2 groups (SD vs nSD), between subjects	2 groups (SD vs nSD), between subjects	3 groups (extreme SD, moderate SD, nSD), between subjects	2 (SD vs nSD) x 2 (male vs female), between subjects	2 (SD vs nSD) x 2 ("hard" vs social sciences) x 2 (male vs female), between subjects	2 (SD vs nSD) x 2 (exchange vs no exchange), between subjects	2 (SD vs nSD) x 2 (museum guide vs researcher) + 1 (control), between subjects
Sensitivity (80% power; $\alpha = .05$)	$d = 0.31$ (independent *t*-test: SD vs nSD)	$d = 0.30$ (independent *t*-test: SD vs nSD)	$d = 0.24$ (independent *t*-test: nSD vs moderate + extreme SD), $f = 0.13$ (omnibus ANOVA, 3 groups)	$d = 0.31$ (independent *t*-test: SD vs nSD), $f^2 = 0.02$ (2×2 MANOVA)	$d = 0.28$ (independent *t*-test: SD vs nSD), $f^2 = 0.01$ (2×2×2 MANOVA)	$d = 0.28$ (independent *t*-test: SD vs nSD), $f^2 = 0.02$ (2×2 MANOVA)	$d = 0.29$ (independent *t*-test: SD vs nSD), $f^2 = 0.01$ (global effects MANOVA, 5 groups), $f^2 = 0.02$ (2×2 MANOVA)
Dependent variables	Closeness	Closeness	Closeness	Closeness	--	Closeness	Closeness
	Trustworthiness (expertise and affective trustworthiness)	Trustworthiness (expertise and affective trustworthiness)	Trustworthiness (expertise and affective trustworthiness)	Trustworthiness (expertise and affective trustworthiness)	Trustworthiness (expertise and affective trustworthiness)	Trustworthiness (expertise and affective trustworthiness)	Trustworthiness (expertise and affective trustworthiness)
	--	Credibility (acceptance and behavioral intentions)	Credibility (acceptance and behavioral intentions)	Credibility (acceptance and behavioral intentions)	Credibility (acceptance and behavioral intentions)	--	Credibility (acceptance and behavioral intentions)
	--	--	--	--	--	Willingness to engage with science	Liking & interest

Note: SD = self-disclosure, nSD = no self-disclosure.

Table 5.3

All self-disclosure manipulation texts used in Study 1

Slide Nr.	Self-disclosure	No self-disclosure
Slide 2	View from my window This is how it looked in my last hotel	View from a window This is how it might look in a hotel
Slide 3	Do my sunglasses have something to do with how I perceive music? This and other questions I asked myself during my last holiday, for example, when I took this photo at the beach of Tenerife.	Do sunglasses have something to do with how music is perceived? This and other questions one might ask oneself during the holidays, for example, when taking this photo at the beach of Tenerife.
Slide 5	Maybe you will soon only meet me wearing sunglasses and headphones?	Maybe you will soon only meet people wearing sunglasses and headphones?

Note. Texts accompanied illustrative photos.

Table 5.4

Means and standard deviations by experimental groups, Study 1

Variable	Self-disclosure	*M*	*SD*
affective trustworthiness	no	4.71	0.87
	yes	5.05	0.90
expertise	no	4.72	1.30
	yes	4.58	1.31
warmth	no	3.96	1.06
	yes	4.35	1.05
competence	no	3.67	1.09
	yes	3.84	0.96
closeness	no	2.30	1.31
	yes	2.65	1.46
perceived disclosure (MC)	no	2.29	0.98
	yes	3.30	1.06

Note. M and *SD* are used to represent mean and standard deviation, respectively.

Table 5.5

Means and standard deviations by experimental groups, Study 2

Variable	Self-disclosure	*M*	*SD*
affective trustworthiness	no	4.86	0.87
	yes	4.93	0.87
expertise	no	4.50	1.32
	yes	4.42	1.26
acceptance	no	2.75	1.06
	yes	2.70	1.06
behavioral intentions	no	3.08	1.40
	yes	3.12	1.35
closeness	no	3.04	1.22
	yes	3.32	1.32
perceived disclosure (MC)	no	2.27	0.86
	yes	2.89	0.90

Note. M and *SD* are used to represent mean and standard deviation, respectively.

Table 5.6

Means and standard deviations by experimental groups, Study 3

Variable	Self-disclosure	*M*	*SD*
affective trustworthiness	no	4.96	0.87
	moderate	5.09	0.85
	extreme	5.04	0.84
expertise	no	4.94	1.19
	moderate	4.78	1.29
	extreme	4.62	1.37
acceptance	no	3.91	1.34
	moderate	3.63	1.50
	extreme	3.64	1.49
behavioral intentions	no	4.15	1.57
	moderate	3.73	1.66
	extreme	3.96	1.61
closeness	no	2.74	1.24
	moderate	2.73	1.25
	extreme	2.85	1.29
perceived disclosure (MC)	no	2.37	1.19
	moderate	3.17	1.30
	extreme	4.01	1.27

Note. M and *SD* are used to represent mean and standard deviation, respectively.

Table 5.7

Means and standard deviations by experimental groups, Study 4

Variable	Self-disclosure	Researcher's gender: male	Researcher's gender: female
affective trustworthiness	no	5.19 (0.95)	5.13 (0.87)
	yes	5.36 (1.02)	5.48 (0.90)
expertise	no	5.97 (0.92)	5.71 (0.99)
	yes	5.77 (1.15)	5.77 (1.00)
acceptance	no	4.52 (1.09)	4.46 (0.96)
	yes	4.51 (1.09)	4.43 (0.85)
behavioral intentions	no	4.59 (0.88)	4.50 (1.08)
	yes	4.54 (1.10)	4.53 (0.88)
closeness	no	2.68 (1.38)	2.53 (1.30)
	yes	3.03 (1.42)	3.19 (1.43)
perceived disclosure (MC)	no	1.73 (0.88)	1.91 (0.98)
	yes	3.51 (1.34)	3.57 (1.27)

Note. Values represent means, values in brackets represent standard deviations.

Table 5.8

Means and standard deviations by experimental groups, Study 5

Variable	Self-disclosure	Disciplines „hard" sciences	Disciplines social sciences
affective trustworthiness	no	4.90 (1.12)	5.08 (1.10)
	yes	4.87 (1.06)	5.06 (1.02)
expertise	no	5.28 (1.20)	5.19 (1.26)
	yes	5.09 (1.16)	5.13 (1.19)
acceptance	no	3.36 (1.22)	3.58 (1.26)
	yes	3.51 (1.08)	3.45 (1.13)
behavioral intentions	no	2.59 (1.24)	2.76 (1.07)
	yes	2.89 (1.01)	2.63 (1.10)
perceived disclosure (MC)	no	2.76 (1.42)	2.64 (1.43)
	yes	4.67 (1.26)	4.54 (1.38)

Note. Values represent means, values in brackets represent standard deviations.

Table 5.9

Means and standard deviations by experimental groups, Study 6

Variable	Self-disclosure	Exchange no	Exchange yes
affective trustworthiness	no	4.70 (0.70)	4.62 (0.65)
	yes	4.66 (0.70)	4.75 (0.73)
expertise	no	4.83 (0.89)	4.92 (0.73)
	yes	4.54 (1.05)	4.58 (0.93)
closeness	no	3.50 (1.21)	3.70 (1.32)
	yes	3.63 (1.29)	3.76 (1.48)
willingness to engage	no	3.63 (1.20)	3.69 (1.18)
	yes	3.52 (1.28)	3.61 (1.32)
warmth	no	4.32 (0.77)	4.25 (0.80)
	yes	4.39 (0.72)	4.55 (0.82)
competence	no	4.57 (0.72)	4.58 (0.65)
	yes	4.46 (0.81)	4.50 (0.84)
perceived disclosure (MC)	no	2.65 (1.14)	2.69 (1.40)
	yes	3.75 (1.23)	4.12 (1.05)

Note. Values represent means, values in brackets represent standard deviations.

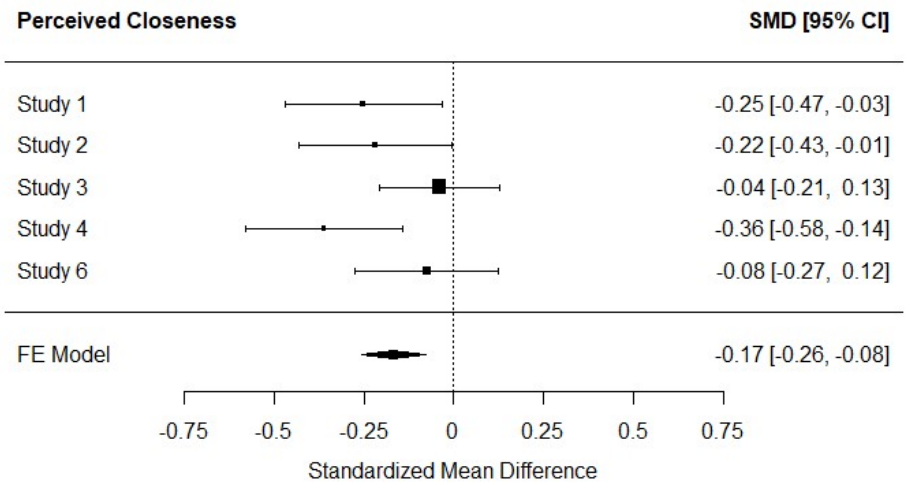
Table 5.10

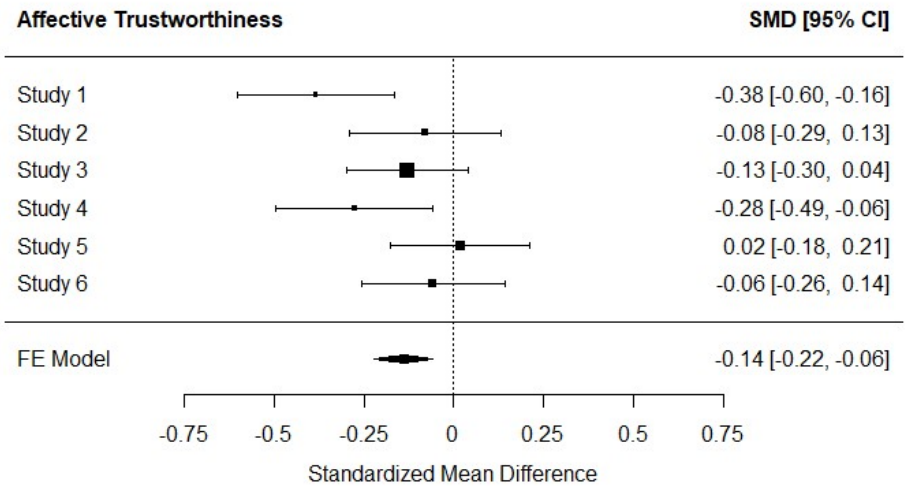
Means and standard deviations by experimental groups, Field Study

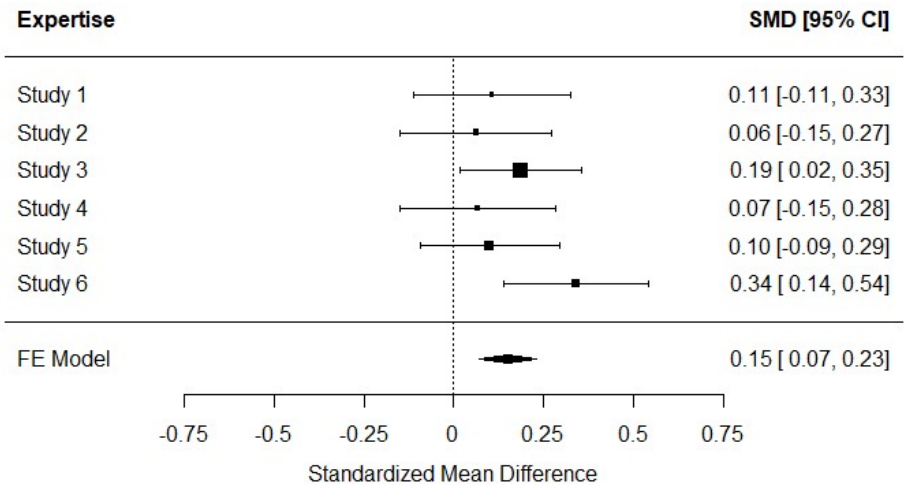
Variable	Self-disclosure	Audio guide role: scientist	Audio guide role: museum guide	neutral control group
liking	no	4.60 (0.84)	4.67 (0.90)	
	yes	4.61 (0.85)	4.55 (0.83)	
				4.61 (0.83)
interest	no	4.63 (0.76)	4.70 (0.78)	
	yes	4.66 (0.82)	4.66 (0.83)	
				4.66 (0.69)
affective trustworthiness	no	5.14 (0.66)	5.12 (0.62)	
	yes	5.23 (0.70)	5.25 (0.63)	
				5.09 (0.59)
expertise	no	5.24 (0.74)	5.26 (0.67)	
	yes	5.16 (0.78)	5.06 (0.77)	
				4.98 (0.73)
acceptance	no	5.12 (0.66)	5.13 (0.63)	
	yes	5.09 (0.71)	5.10 (0.62)	
				5.10 (0.67)
behavioral intentions	no	5.19 (0.69)	5.19 (0.66)	
	yes	5.25 (0.66)	5.19 (0.71)	
				5.26 (0.70)
closeness	no	3.61 (1.28)	3.76 (1.08)	
	yes	3.98 (1.38)	3.69 (1.26)	
				3.61 (1.23)
perceived disclosure (MC)	no	3.18 (1.12)	2.78 (1.25)	
	yes	4.86 (0.85)	4.85 (0.88)	
				1.98 (1.11)

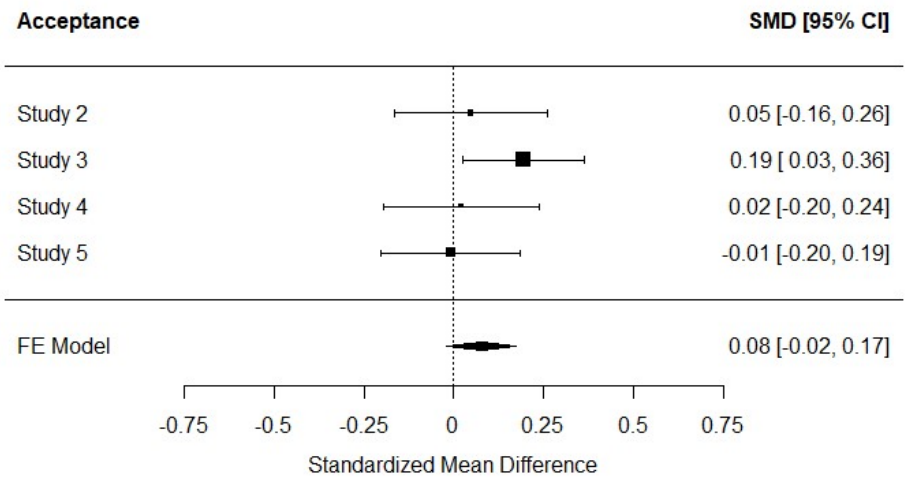
Note. Values represent means, values in brackets represent standard deviations.

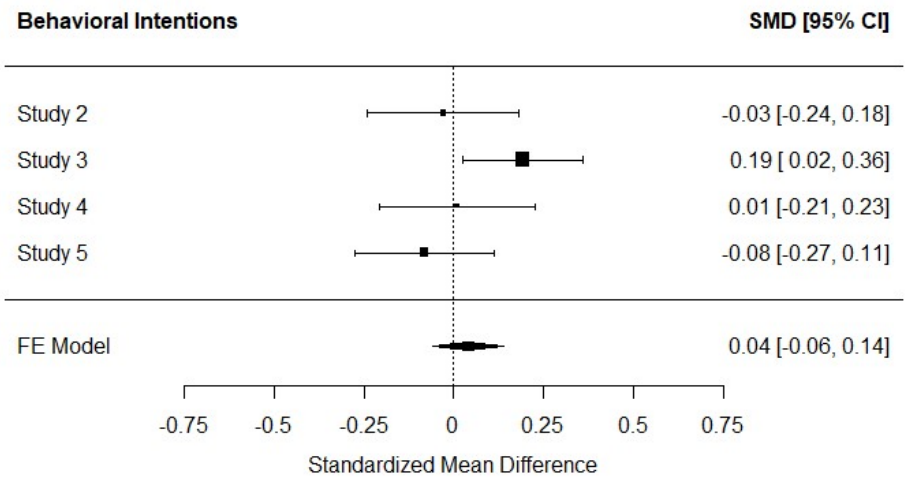
Figures 5.1 – 5.5








Acceptance
SMD [95% CI]
Study 2
0.05 [-0.16, 0.26]
Study 3
0.19 [0.03, 0.36]
Study 4
0.02 [-0.20, 0.24]
Study 5
-0.01 [-0.20, 0.19]
FE Model
0.08 [-0.02, 0.17]
-0.75
-0.5
-0.25
0
0.25
0.5
0.75
Standardized Mean Difference


Behavioral Intentions
SMD [95% CI]
Study 2
-0.03 [-0.24, 0.18]
Study 3
0.19 [0.02, 0.36]
Study 4
0.01 [-0.21, 0.23]
Study 5
-0.08 [-0.27, 0.11]
FE Model
0.04 [-0.06, 0.14]
-0.75
-0.5
-0.25
0
0.25
0.5
0.75
Standardized Mean Difference

6 Manuscript 2: Self-criticism and reform intentions in science

This manuscript has been published:

Altenmüller, M. S., Nuding, S., & Gollwitzer, M. (2021). No harm in being self-corrective: Self-criticism and reform intentions increase researchers' epistemic trustworthiness and credibility in the eyes of the public. *Public Understanding of Science*, *30*(8), 962–976. https://doi.org/10.1177/09636625211022181

[manuscript printed as accepted]

Abstract

Science should be self-correcting. However, researchers often hesitate to admit errors and to adopt reforms in their own work. In two studies (overall *N*=702), we test whether scientific self-criticism and reform intentions expressed by researchers damage or rather improve their reputation in the eyes of the public (i.e., perceivers). Across both studies, such self-correction (compared to no self-correction) increases perceivers' epistemic trustworthiness ascriptions, credibility perceptions, and willingness to further engage with science. Study 2 revealed that these effects were largely driven by the no self-criticism condition. In addition, researchers' commitment to implementing reforms had positive effects and rejecting reforms had negative effects on perceptions, irrespective of the extent of these reforms. These findings suggest that researchers' fear that self-criticism and expressing reform intentions may damage their reputation may be unfounded. [130 words]

Key words: self-criticism, reforms, trust, credibility, open science.

No Harm in Being Self-Corrective: Self-Criticism and Reform Intentions Increase Researchers' Epistemic Trustworthiness and Credibility in the Eyes of the Public

Humans fail and err all the time in their daily lives. Yet, admitting failures and errors is psychologically costly: doing so threatens one's self-image and the desire to make a favorable impression upon others. While there are certainly contexts in which people's hesitation to admit potential errors and to change their work routines is psychologically comprehensible, science is no such context: here, a self-corrective mindset (i.e., admitting flaws and intending to improve one's routines) is crucial. The basic idea of scientific progress is that by constantly correcting previous work and improving future work, researchers increase the likelihood of detecting the "truth" (or, at least, a robust phenomenon) and, thus, their own and other stakeholders' confidence that their findings are trustworthy and credible.

However, empirical findings suggest that the ever-doubtful and self-correcting scientist is an ideal and that, when it comes to admitting flaws and failures, researchers behave just as "normal" people do: they hesitate to do so for the sake of protecting their self-image and their reputation (Bishop, 2018; Fetterman & Sassenberg, 2015; Rohrer et al., 2021; van der Bles et al., 2020). Thus, the question is whether these concerns are justified. How does the general public react to researchers who admit that their work may have been faulty and that they are willing to implement reforms aimed at improving the quality of their future work? Would the public place more vs. less trust in researchers who admit such prior faults and/or who express reform intentions, and does the extent of these intended reforms play a role here? The two studies presented here were

designed to find answers to these questions.

Trust in science can be conceptualized on different levels: trust in science as a whole and trust in individual researchers and their work. Judging a scientist to be a reliable source of knowledge is known as epistemic trust, which includes a cognitive aspect—*expertise*—and two affective aspects—*benevolence* and *integrity* (Fiske & Dupree, 2014; Hendriks et al., 2015; Mayer et al., 1995; McAllister, 1995; Neal et al., 2012). Expertise means that one is perceived as able and competent with regard to their (scientific) work; benevolence means that one is perceived as having the best in mind for others/society; and, finally, integrity means that one is perceived as adhering to prescriptive rules and principles. An additional element of trust in science, which is not directly reflected by any of the three interpersonal facets mentioned above, is *credibility*—people's willingness to accept a scientific finding as true and integrating it in their own understanding of the world. Both epistemic trustworthiness and credibility are relevant for maintaining general trust in science on a societal level.

Previous findings suggest that self-criticism has positive as well as negative effects on people's trust in scientists. For instance, Hendriks et al. (2016) found that expertise ascribed to a science blogger was lower when the blogger admitted (vs. did not admit) an error in one of their blog entries, while perceived integrity as well as benevolence were higher. Notably, the error that the blogger admitted was not related to the research *per se* (i.e., the methodological quality of the study), but rather to how they communicated about it (i.e., overgeneralization of results in a science journalism piece). It is unclear, however, how admitting doubts about one's own research might influence the public's trust in scientists. In addition, it is unclear whether such self-criticism also affects credibility judgments. Here, effects

are plausible in both ways: On the one hand, admitting doubts about past work might imply incompetence and, hence, decrease expertise and credibility judgments. On the other hand, being self-critical and noticing and disclosing potential flaws in one's previous work might indicate a more attentive approach to future research projects, which leads to more confidence that this work will produce sound scientific results. Additionally, a self-critical approach to one's research demonstrates commitment to a certain scientific attitude of constantly challenging and updating scientific knowledge, even at the cost of questioning oneself.

Regarding the effect of reform intentions on the public's trust in science, previous studies yielded an inconsistent pattern: recent studies support the idea that successful replications increase laypeople's trust in science (Hendriks et al., 2020; Wingen et al., 2020). However, learning about *specific* reforms can have null or even backfiring effects. For instance, Wingen et al. (2020) found that increasing transparency (e.g., by means of preregistrations, open data, and open materials) or providing explanations for the "replicability crisis" in psychology had no effect on laypeople's trust in psychological science. Anvari and Lakens (2018) even found that participants expressed less trust in psychological science after learning about suggested reforms. The authors discuss three explanations for this backfiring effect: first, respondents may have been negatively surprised that the proposed reforms are not already common practice; second, their manipulation of reform implementation might have been problematic (i.e., their "reform" vignette began by talking about replication failures, while the other vignettes began by talking about the history of psychology, which might have elicited a stronger negative response in the reform condition); third, participants might have judged the

reforms to be too weak. These inconclusive results call for more research on the effect of reform intentions on epistemic trustworthiness and credibility judgments.

Expanding our main focus on trustworthiness and credibility,[12] we will also investigate the effect of self-criticism and reform intentions on participants' *willingness to engage with science*. Disclosure of uncertainties and doubts has been found to not only have a mixed impact on perceptions of trustworthiness and credibility (e.g., Hendriks et al., 2016; Jensen, 2008; van der Bles et al., 2020) but also to increase the public's interest in science and in new technologies (Retzbach & Maier, 2015). Thus, self-correction in science—expressing self-criticism and reform intentions—might also influence laypeople's interest and make them want to engage more with science.

In the present paper, we report two preregistered studies investigating the effects of researchers' self-criticism and reform intentions on their epistemic trustworthiness, the credibility of their future findings, and the public's willingness to engage further with these researchers and their findings. For both studies, we report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures (Simmons et al., 2012). All materials, the anonymized data, and analyses are available online (https://osf.io/yhsbp/).

Study 1

In Study 1, we compare the effect of self-criticism (yes vs. no) and reform intentions in varying degrees: As previous studies regarding the perception of reforms focused on reforms in general (Anvari & Lakens, 2018) or on specific reform approaches (Wingen et al., 2020), we manipulate the extent of these

[12] Note: *Willingness to engage* was explored in Study 1 and preregistered only in Study 2.

reforms irrespective of their specific content (see Methods; no, minor or major reforms). This way, we account for two issues related to previous research in this area: First, we investigate the effect of the extent to which a researcher promises to implement reforms in their research on laypeople's trust and credibility perceptions (Anvari & Lakens, 2018). Second, we try to get a more generalizable picture of the effects of reform intentions (as the findings will not be specific to distinct reforms, but rather get at the general willingness to implement reforms of differing degrees).

In Study 1, we also explore possible interaction effects of self-criticism and reform intentions on epistemic trust, credibility judgments, and willingness to engage. First, it is possible that self-criticism of previous work and reform intentions are perceived independently from each other, and, thus, yield two main effects on trust, credibility, and willingness to engage. Second, it is also possible that expressing doubts about one's prior work makes reform intentions more reasonable and more credible, which would result in a synergetic interaction between self-criticism and reform intentions on trust, credibility, and willingness to engage. Third and finally, self-criticism of past work might lead to reform intentions for future work being perceived as a mandatory consequence; therefore, self-criticism followed by a refusal to implement future reforms may be perceived as inconsistent, and expressing reform intentions preceded by full-blown confidence in one's prior results may be perceived as "cheap talk" and, thus, lead to particularly low levels of trust, credibility, and willingness to engage (i.e., an ordinal interaction between self-criticism and reform intentions).

Method

Experimental Manipulation. We conducted an online study using a 2 (self-

criticism: yes/no) × 3 (reform intentions: no/minor/major) full between-subject design. After obtaining informed consent, participants were asked to read an alleged online interview with Dr. Romberg[13], a psychological researcher, who talks about a past study he conducted. At one point during the interview, Dr. Romberg stated (without being asked) either that "...*looking back on this study today, I admittedly doubt these findings. The results are probably not quite right, because according to my current methodological knowledge this study has some weaknesses*" (self-criticism) or that "...*looking back on this study today, I do not doubt these findings. The results are probably right, because according to my current methodological knowledge this study has no weaknesses*" (no self-criticism).

Next, being asked explicitly about the on-going "open science" reform discussion in psychological science, Dr. Romberg describes some reforms in lay-terms (transparency through open data and preregistrations) and explains how they should work (enhanced reproducibility and constructive exchange, early detection of mistakes, higher reliability of findings). Then the interviewer asks Dr. Romberg how he judges these reforms in regard to his own research, to which he either states: "*To be honest, I do not think these reforms are necessary for research on group processes. Therefore, I won't apply any of these currently discussed reforms in my future research*" (no reform intentions); "*To be honest, I think these reforms are partly necessary for research on group processes.*

[13] We made "Dr. Romberg" (Study 1), as well as "Dr. Kugler" and "Dr. Ecker" (Study 2) male researchers because of linguistic simplicity in the German language, yet, we assume the manipulations would work just as well with female researchers. We avoided overemphasizing gender by using the neutral "Dr. Romberg/Kugler/Ecker" more that "he/his/him".

Therefore, I will apply some of these currently discussed reforms in my future research" (minor reform intentions); or "*To be honest, I think these reforms are necessary for research on group processes. Therefore, I will apply many of these currently discussed reforms in my future research*" (major reform intentions).

Dependent Variables. After completing two attention check questions ("Which optimal group sizes did Dr. Romberg's study show?"; "Which topic does Dr. Romberg want to investigate next?"), participants rated Dr. Romberg's *trustworthiness* with the Muenster Epistemic Trustworthiness Inventory (METI; Hendriks et al., 2015), consisting of 14 opposite adjective pairs measuring expertise (e.g., competent – incompetent, Cronbach's α=.94) and integrity and benevolence (e.g., honest – dishonest, Cronbach's α=.95)[14] on 6-point bipolar scales. Next, they rated the perceived *credibility* of Dr. Romberg's future research on four items developed specifically for the purpose of this study based on theoretical assumptions (e.g., Anvari & Lakens, 2019), including cognitive as well as behavioral indicators of credibility (6-point Likert scale ranging from 1="not at all" to 6="absolutely;" e.g., "I think, future research findings by Dr. Romberg will be credible", "I will try to consider future research findings by Dr. Romberg in my daily life"; Cronbach's α=.83). Finally, participants' *willingness to further engage* with Dr. Romberg's research ("I intend to register for the free account to be able to read the rest of the article"), support for public funding ("Dr. Romberg's future research deserves public funding"), and likeability ("I like Dr. Romberg") were

[14] Contrary to the proposed three-factor solution by Hendriks et al. (2015), factor analyses conducted with the present data suggest a two-factor solution with expertise (factor 1) and integrity/benevolence (factor 2), corresponding to the idea of a cognitive and an affective dimension of trustworthiness. Deviating from our pre-registration in Study 1, epistemic trustworthiness will be analyzed on these two dimensions instead of three.

assessed on 6-point Likert scales ranging from 1="not at all" to 6="absolutely".

Other Measures. In a final section, participants completed two manipulation comprehension check questions ("Did Dr. Romberg criticize his own previous study about group size?"; "Does Dr. Romberg want to apply reforms for his future research?") by selecting either "yes" or "no." If participants believed Dr. Romberg wanted to apply reforms, participants were asked about the assumed extent of these reforms (response options were "some reforms," "many reforms," or "don't know"). To control for *prior knowledge*, we asked participants whether they had heard about the "replication debate" in psychology before, and, if yes, how much they knew about it on a 6-point Likert scale ranging from 1="not much" to 6="a lot." Further, we measured participants' general Public Engagement with Science (PES) using two scales that had been used in previous research (BBVA Foundation, 2011): a 5-item scale measuring engagement *PES frequency* (e.g., "How often do you read news about science?", 5-point Likert scale ranging from 0="never" to 5="almost daily," Cronbach's α=.71) and a multiple choice scale measuring 15 potential *PES experiences* during the last 12 months (e.g., "I know someone who does scientific research", "I visited a science museum"). Finally, demographics (age, gender, occupation, academic discipline) and a "use-me" item ("Should we use your data for our analyses?", yes/no) were assessed. Participants had the opportunity to participate in a lottery and sign up for more information and were debriefed.

Sample. Participants were recruited via mailing lists (e.g., by the university, by the research unit) and social networks (e.g., Facebook, science blogs). Inclusion criteria were very good German language skills and a minimum age of 16. Five-hundred twenty-one participants completed the study. Applying pre-registered

exclusion criteria (see https://osf.io/qja78), 184 participants had to be excluded from the data set: 45 participants stated not to use their data; 34 participants spent less than 60 second viewing the manipulation; 105 participants failed the main manipulation comprehension checks.[15] The final sample consisted of *N*=337 participants (68.0% female, 32.0% male); ages ranged between 16 and 74 years (*M*=43.33; *SD*=14.73). Most participants were currently employed (59.6%; students: 20.2%; unemployed: 20.2%). Participants who were currently studying at a university or already had a university degree (61.4%) came from a variety of disciplines (law, economics, and social sciences: 33.2%; humanities: 10.7%; mathematics and natural sciences: 7.1%; engineering: 6.5%; medicine and life sciences: 3.3%). Although *N*=337 is lower than the determined sample size specified in our pre-registration, the power is still large enough (i.e., 90%) to detect a small-to-medium interaction effect in our 2×3 ANOVA (Φ^2=.195) on a 5% significance level (Faul et al., 2007).

Results

Supporting the effectiveness of our randomization, neither general public engagement with science (PES frequency: *p*=.27; PES experiences: *p*=.61) nor prior knowledge about the replication debate (*p*=.12) differed between the six cells of our design (mean difference tests via one-way ANOVAs). Across all conditions, 19% of our participants had heard about the replication debate before; on

[15] Further 83 participants could not correctly remember the extent of the reform intention but, in line with the preregistration, were not excluded. Excluding them did not change the results on our dependent variables. Deviating from the preregistration, we did not exclude 56 participants with same answers on every single item on a questionnaire page or choosing only extreme response options (regarding epistemic trustworthiness and credibility). Excluding them did not change our results on our dependent variables. Furthermore, one could argue that these are, in fact, plausible answers. Because of this and additional power concerns, we decided not to exclude these participants.

average, they judged their knowledge about the replication debate (M=4.03, SD=1.52) and QRPs (M=3.34, SD=1.61) to be moderate. Table 6.1 summarizes all means, standard deviations and correlations.

Next, we conducted a 2×3 MANOVA to test the effects of self-criticism and reform intentions on the two epistemic trustworthiness dimensions (expertise and integrity/benevolence), credibility, and willingness to further engage with the research.[16] Both self-criticism, F(4, 328)=4.38, p<.01, Pillai-V=0.05, η_p^2=.05 (CI_{95}=.01; .09), and reform intentions, F(8, 658)=20.53, p<.001, Pillai-V=0.40, η_p^2=.20 (CI_{95}=.14; .24), had multivariate main effects, while the interaction effect was not significant, F(8, 658)=0.96, p=.46, Pillai-V=0.02, η_p^2=.01 (CI_{95}=.00; .02). We followed up with univariate analyses. Means and standard deviations, broken down by conditions, are reported in Table 6.2.

[16] Because of the very high intercorrelations, we deviated from the preregistered analyses by adding all DVs to the MANOVA. The results did not differ significantly from the originally planned analyses.

Table 6.1

Means, Standard Deviations, and Correlations Between Measured Variables

Variable		*M*	*SD*	Correlations								
				1	2	3	4	5	6	7	8	9
1	Expertise	4.79	0.95	*.94*								
2	Integrity/Benevolence	4.47	1.00	.78**	*.95*							
3	Credibility	3.79	1.14	.73**	.75**	*.83*						
4	Willingness to Engage	2.53	1.64	.29**	.27**	.36**	-					
5	Public Funding Support	4.00	1.33	.59**	.66**	.68**	.35**	-				
6	Likeability	3.69	1.31	.66**	.73**	.71**	.32**	.68**	-			
7	Replication Debate Knowledge	4.03	1.52	-.13	-.01	-.13	-.04	-.09	-.08	-		
8	QRPs Knowledge	3.34	1.61	-.18	-.08	-.12	.02	-.16	-.16	.76**	-	
9	PES Frequency	3.16	0.63	-.03	-.06	-.09	.16**	.01	-.07	-.09	-.07	*.71*
10	PES Experiences	5.09	2.96	-.15**	-.15**	-.13	.09	-.05	-.07	.24	.25*	.51**

Notes. *N*=337; for variables 7 and 8: *N*=64. PES = Public Engagement with Science. * $p<.05$; ** $p<.01$. Cronbach's α for each scale are reported in the diagonal (in italics).

Table 6.2

Means and Standard Deviations, Broken Down by Conditions

Variable	Self-Criticism	Reform Intentions		
		none	minor	major
Expertise	no	4.00 (1.07) [a]	5.01 (0.84) [b]	5.14 (0.79) [b]
	yes	4.30 (0.84) [a]	5.05 (0.72) [b]	5.22 (0.70) [b]
Integrity/	no	3.44 (0.94) [a]	4.65 (0.71) [b]	4.84 (0.83) [b]
Benevolence	yes	3.83 (0.85) [a]	4.95 (0.76) [b]	5.06 (0.67) [b]
Credibility	no	2.83 (1.15) [a]	4.26 (0.85) [b]	4.22 (0.87) [b]
	yes	2.92 (0.96) [a]	4.18 (0.82) [b]	4.32 (0.94) [b]
Willingness to	no	2.12 (1.59)	2.91 (1.68)	2.67 (1.72)
Engage	yes	2.09 (1.53) [a]	2.35 (1.60)	3.05 (1.57) [b]

Notes. *N*=337. Means (standard deviations in brackets). In each line, different letters in the superscript indicate significant pairwise differences (i.e., $p<.05$; Tukey HSD test).

Epistemic Trustworthiness. Univariate analyses show a significant main effect of self-criticism on integrity/benevolence, $F(1, 331)=11.97$, $p<.001$, $\eta_p^2=.04$ ($CI_{95}=.01; .08$), but not on expertise, $F(1, 331)=2.39$, $p=.12$, $\eta_p^2=.01$ ($CI_{95}=.00; .04$), as well as significant main effects of reform intentions on both integrity/benevolence, $F(2, 331)=91.45$, $p<.001$, $\eta_p^2=.36$ ($CI_{95}=.27; .42$), and expertise, $F(2, 331)=50.62$, $p<.001$, $\eta_p^2=.23$ ($CI_{95}=.16; .30$). The interaction effects were non-significant on either DV ($p=.44$ for expertise; $p=.71$ for integrity/benevolence). Looking at the reform intentions factor, follow-up pairwise comparisons (i.e., Tukey HSD tests) suggest that both reform intention conditions differed from the no reform condition (all $p<.001$); yet, the extent of these reforms

(i.e., minor vs. major) did not affect the DVs (p=.35 for expertise; p=.37 for integrity/benevolence).

Credibility. On credibility, the main effect of self-criticism was not significant, $F(1, 331)=0.15$, p=.70, η_p^2<.001 (CI_{95}=.00; .02), while the main effect of reform intentions was significant, $F(2, 331)=78.72$, p<.001, η_p^2=.32 (CI_{95}=.24; .39). Again, there was no interaction effect, $F(2, 331)=0.28$, p=.76, η_p^2<.01 (CI_{95}=.00; .02). As before, the reform conditions significantly differed from the no reform condition (both p<.001); but the extent of reforms did not make a difference (p=.87).

Willingness to Engage. On participants' willingness to further engage with the research, we found no significant main effect of self-criticism, $F(1, 331)=0.14$, p=.71, η_p^2<.001 (CI_{95}=.00; .02), but a significant main effect of reform intentions, $F(2, 331)=6.85$, p<.01, η_p^2=.04 (CI_{95}=.01; .08). Once again, there was no significant interaction effect, $F(2, 331)=2.31$, p=.10, η_p^2=.01 (CI_{95}=.00; .04). Follow-up analyses reveal a slightly more complex pattern than before (see Table 6.2). Post-hoc tests only revealed a significant difference between the no reform and the major reforms conditions on this DV (p<.001); all other comparisons were non-significant (no vs. minor: p=.07; minor vs. major: p=.35).

We also conducted explorative analyses to investigate the effects of self-criticism and reform intentions on public funding support as well as on likeability ratings. The results closely mirror the findings for credibility. These findings are not the focus of our present paper and are provided in a supplementary file (https://osf.io/yhsbp/). To scrutinize whether the effects reported here merely reflect an unspecific "halo" effect, we re-ran our analyses with likeability as a covariate (see supplementary file). Importantly, the main effect of self-criticism on

integrity/benevolence remained significant, and the main effect of reform intentions on expertise, integrity/benevolence, and credibility remained significant, too. The main effect of reform intentions on willingness to engage, however, became non-significant. We will come back to this in the General Discussion.

Discussion

These findings suggest that there are no detrimental effects of self-criticism and reform intentions on laypeople's trustworthiness ascriptions and credibility perceptions. Expressing self-criticism (vs. no self-criticism) led to higher benevolence and integrity perceptions—the affective dimensions of trustworthiness. Announcing reform intentions (compared to no reform intentions) had positive effects on epistemic trustworthiness and credibility, as well as on participants' willingness to engage further with the expert's research.

This pattern is in line with previous findings (Hendriks et al., 2016). Contradictory, however, we found that self-criticism did not negatively affect expertise (Hendriks et al., 2016). This might be due to how self-criticism was particularly framed in our study: The self-critical researcher's statement implied a more advanced methodological knowledge by the time he was interviewed compared to when the study was conducted ("*according to my current methodological knowledge*"), which actually suggests increased expertise that participants might have picked up upon. In Hendriks et al.'s study (2016), however, the researcher revised his previous overgeneralizing statement in a blog entry which might not have been perceived as improved knowledge but rather as correcting a careless mistake that could happen again.

Self-criticism only impacted affective dimensions of epistemic trustworthiness, but had no effect on expertise or on credibility. One reason for

this could be that self-criticism regarding a particular study conducted in the past does not tell us much about the credibility of future research. A second reason could be that, in our study, reform intentions were mentioned at length at the end of the alleged interview, immediately before the dependent variables were measured. This might have overshadowed any effect of our self-criticism manipulation, which was mentioned earlier in the interview.

Announcing reforms had consistent and large effects on all of our dependent variables. Contrary to previous studies (Anvari & Lakens, 2018; Wingen et al., 2020), expressing even minor reform intentions led to higher ratings of epistemic trustworthiness and credibility compared to a no-reform intention control condition. Interestingly, the extent of such reforms did not have any effects. This suggests that it might be sufficient to signal at least some willingness to improve one's scientific practices.

It should be noted that, in Study 1, self-criticism and reform intention statements were contrasted with conditions in which the researcher explicitly expressed *no* self-criticism regarding prior findings and/or *refused* implementing reforms. Thus, it is unclear whether our results display an increase in trust and credibility due to expressing self-criticism and/or reform intentions, a decrease of the same due to being overconfident, or both.

Our design allowed us to investigate whether self-criticism and reform intentions interact with each other. Regarding such an interaction, both amplifying effects (e.g., self-criticism makes reform intentions seem more reasonable) as well as alleviating effects (e.g., self-criticism alleviates the effect of reform intentions as they seem imperative) would have been plausible. However, we did not find any interaction effects on any of our measured variables, which suggests that self-

criticism and reform intentions are independent of (and not contingent on) each other.

Study 2

Study 2 aimed to replicate and clarify the effects of self-criticism and reform intentions. For this purpose, we made some changes compared to Study 1. First, we manipulated each of the two independent variables, separately, because (a) we did not find evidence of an interaction between the two, and (b) doing so reduced the danger of any artificial "overshadowing" effects between the two manipulations. Second, to scrutinize which condition might drive the effects obtained in Study 1, we added neutral ("control") conditions for both self-criticism and reform intentions (see Method). Third, we refined the operationalization of self-criticism as expression of a self-corrective attitude towards science: The researcher's self-criticism was based on viewing his prior findings as preliminary and fragile (indicating high self-criticism) vs. as fixed and definite (indicating lacking self-criticism). Fourth, we no longer differentiate between minor and major reform intentions, since Study 1 suggested no differences between these two conditions.

Method

Experimental Manipulation. In Study 2, we again used a full between-subject design. However, we split the study in two parts (presented in randomized order): 2A) self-criticism (yes vs. no vs. no information), and 2B) reform intentions (yes vs. no vs. undecided). Participants read two alleged interviews with researchers, Dr. Kugler and Dr. Ecker, in which they talked about their own research on group processes (full materials are provided here: https://osf.io/yhsbp/).

In interview A, we manipulated self-criticism, defined as the expression of a self-critical attitude towards prior findings. When asked about a prior study, the researcher, Dr. Kugler, described the results and then expressed either no self-criticism ("*Still today, I actually do not have doubts about these findings. I see no reason why the results from back then shouldn't also apply today and would view them as definite*"), self-criticism ("*However, today, I do have some doubts about these findings. Viewed scientifically, there might be reasons why the results from back then do not apply today and, thus, I would view them as preliminary*"), or nothing of that kind (no information control condition).

In interview B, we manipulated reform intentions: After describing some general reform ideas in psychology and the ongoing debate, the interviewed researcher, Dr. Ecker, either stated that he would not implement any reforms in his research (no reform intentions), that he would implement such reforms (reform intentions), or that he was still undecided whether to implement reforms (undecided).

Dependent Variables. After each interview, we measured the same variables as in Study 1: First, they answered two respective attention checks (2A: "What was the optimal group size in Dr. Kugler's earlier study?", "Which topic would Dr. Kugler like to research next?"; 2B: "What is the goal of the discussed reforms according to Dr. Ecker?", "Which topic would Dr. Ecker like to research next?"). Next, trustworthiness and credibility (as well as likeability and support for public funding) were measured exactly as in Study 1. Additionally, we measured participants' willingness to engage further with the researcher and his findings (one item used in Study 1 plus three additional items; e.g., "I would like to learn more about Dr. Kugler's/ Dr. Ecker's research"; assessed on a 6-point Likert scale

ranging from 1="strongly disagree" to 6="strongly agree").

Other Measures. After each interview, we applied a comprehension check similar to Study 1 (2A: "Did Dr. Kugler express doubts about his prior findings?", 2B: "What did Dr. Ecker say about his own intentions to implement reforms?"). Additionally, after the interview concerning reforms in psychology (2B), we again asked whether participants had heard about the replication debate and, if yes, how extensive they judged their knowledge about the replication debate and questionable research practice to be. Then, and similar to Study 1, we assessed participants' general engagement with science (*PES frequency* and *PES experiences*; BBVA Foundation, 2011), demographics (age, gender, occupation, academic discipline), and a "use-me" item ("Do you think we should use your data for our analyses in this study?" yes/no). Finally, participants were fully debriefed and informed about their reimbursement.

Sample. As in Study 1, participants were recruited via mailing lists and social networks and could participate when they were older than 16 years, had very good German language skills, and had not previously participated in Study 1. We collected data from 400 participants as prescribed by our preregistered *a-priori* power analyses for each study part based on our findings in Study 1 (see https://osf.io/9szde/). Applying our preregistered exclusion criteria, 35 participants had to be generally excluded:[17] eleven participants denied the "use-me" question; 5 participants fell below the minimum threshold of 30 seconds for viewing the manipulation texts; 4 participants completed the questionnaire in less than 5

[17] We describe the sample with a general data set including participants who fulfilled the criteria for at least one part of the study. For specific analyses in each part of the study, we only used data from participants who fulfilled the criteria for the respective part of the study (N_{2A}=288; N_{2B}=322).

minutes; 15 participants failed the comprehension check in both parts. The final sample consisted of *N*=365 participants; ages ranged between 16 and 77 years (*M*=30.60; *SD*=13.95; 81.10% female, 18.63% male, 0.27% other). A majority was currently enrolled in a higher education program (64.1%; employed: 26.8%; unemployed: 9.0%). Participants who were studying at a university or already had a university degree (90.1%) came from a variety of disciplines (law, economics, and social sciences: 58.2%; humanities: 9.2%; mathematics and natural sciences: 9.7%; engineering: 3.3%; medicine and life sciences: 6.7%). Thus, compared to Study 1, the sample in the present study was younger, more female, and more highly educated.

Results

Supporting the effectiveness of our randomization, neither general public engagement with science (PES frequency: p_a=.58 and p_b=.63; PES experiences: p_a=.94 and p_b=.70) nor prior knowledge about the replication debate (p_a=.81 and p_b=.76; overall, 34% had heard of it before) differed between experimental conditions in any of the two study parts (mean differences tested via one-way ANOVAs). Descriptive statistics and correlations between measured variables are reported in Tables 6.3 (for Study 2A) and 6.4 (for Study 2B).

Self-Criticism (Study 2A). Using a MANOVA, we tested the effects of self-criticism on the two epistemic trustworthiness dimensions (expertise and integrity/benevolence), credibility, and willingness to engage: The multivariate main effect of self-criticism was, again, significant, and slightly larger than in Study 1, $F(8, 566)=6.38$, $p<.001$, Pillai-V=0.17, η_p^2=.08 (CI_{95}=.03; .12). We followed up with univariate analyses. Self-criticism had a significant effect on all our dependent variables: integrity/benevolence, $F(2, 285)=24.96$, $p<.001$, η_p^2=.15

(CI_{95}=.08; .22), expertise, F(2, 285)=15.24, p<.001, η_p^2=.10 (CI_{95}=.04; .16), credibility F(2, 285)=14.76, p<.001, η_p^2=.09 (CI_{95}=.04; .16), and willingness to engage, F(2, 285)=6.04, p<.01, η_p^2=.04 (CI_{95}=.01; .09). Follow-up pairwise comparisons (i.e., Tukey HSD tests) show that expressing no self-criticism compared to self-criticism or no information led to significantly lower mean values on all DVs. There were no significant differences, however, between expressing self-criticism vs. giving no information (i.e., the control condition). Means and standard deviations, broken down by conditions, and results for follow-up tests are reported in Table 6.5 (upper part).

Reform Intentions (Study 2B). Again, using a MANOVA, we tested the effects of reform intentions on epistemic trustworthiness (expertise and integrity/benevolence), credibility, and willingness to engage. The multivariate main effect of reform intentions was, again, significant, yet slightly smaller than in Study 1, F(8, 634)=9.87, p<.001, Pillai-V=0.22, η_p^2=.11 (CI_{95}=.06; .15). We followed up with univariate analyses. Reform intentions had a significant effect on all our dependent variables: integrity/benevolence, F(2, 319)=33.42, p<.001, η_p^2=.17 (CI_{95}=.10; .24), expertise, F(2, 319)=7.90, p<.001, η_p^2=.05 (CI_{95}=.01; .10), credibility, F(2, 319)=25.05, p<.001, η_p^2=.14 (CI_{95}=.07; .20), and willingness to engage, F(2, 319)=4.85, p<.01, η_p^2=.03 (CI_{95}=.00; .07). Follow-up tests showed that for all DVs, reform intentions compared to no reform intentions led to significantly higher mean values. Reform intentions compared to being undecided led to significantly higher mean values on all variables except for willingness to engage. And being undecided compared to no reform intentions led to significantly higher mean values on all variables except for expertise and, again, willingness to engage. Means and standard deviations, broken down by conditions, and results

for follow-up tests are reported in Table 6.5 (lower part).

Again, explorative analyses of the effects of self-criticism and reform intentions on public funding support as well as likeability indicate a very similar pattern of results as described above for our dependent variables (see https://osf.io/yhsbp/). When controlling for likeability in our analyses to scrutinize a possible halo-effect, self-criticism only had a significant effect on integrity/benevolence, and reform intentions only had significant effects on integrity/benevolence and credibility. We will come back to this in the General Discussion.

Table 6.3

Means, Standard Deviations, and Correlations Between Measured Variables for Study 2A (self-criticism).

Variable		*M*	*SD*	Correlations								
				1	2	3	4	5	6	7	8	9
1	Expertise	4.93	0.87	*.93*								
2	Integrity/Benevolence	4.68	0.88	.73**	*.95*							
3	Credibility	4.05	0.98	.64**	.63**	*.77*						
4	Willingness to Engage	3.74	1.14	.51**	.48**	.47**	*.77*					
5	Public Funding Support	4.44	1.11	.59**	.61**	.56**	.51**	-				
6	Likeability	4.24	1.24	.66**	.77**	.58**	.54**	.64**	-			
7	Replication Debate Knowledge	3.61	1.50	.11	.12	.07	.04	.03	.07	-		
8	QRPs Knowledge	3.06	1.66	.04	.08	.07	-.01	.04	.01	.83**	-	
9	PES Frequency	3.10	0.68	-.01	-.03	-.04	.05	.07	-.06	.15	.17	*.73*
10	PES Experiences	6.15	2.75	-.06	-.09	-.10	.01	.03	-.10	.12	.09	.45**

Notes. *N*=288; for variables 7 and 8: *N*=193. PES = Public Engagement with Science. * $p<.05$; ** $p<.01$. Cronbach's α for each scale are reported in the diagonal (in italics).

Table 6.4

Means, Standard Deviations, and Correlations Between Measured Variables for Study 2B.

Variable		*M*	*SD*	Correlations								
				1	2	3	4	5	6	7	8	9
1	Expertise	4.98	0.82	*.93*								
2	Integrity/Benevolence	4.47	0.96	.71**	*.95*							
3	Credibility	3.85	0.97	.54**	.69**	*.76*						
4	Willingness to Engage	3.51	1.18	.31**	.37**	.41**	*.81*					
5	Public Funding Support	4.23	1.21	.58**	.66**	.65**	.40**	-				
6	Likeability	3.94	1.23	.56**	.67**	.60**	.52**	.69**	-			
7	Replication Debate Knowledge	3.69	1.45	-.10	-.23*	-.23*	-.12	-.20*	-.20*	-		
8	QRPs Knowledge	3.07	1.66	-.16	-.26**	-.27**	-.09	-.21*	-.21*	.80**	-	
9	PES Frequency	3.15	0.67	.00	.02	-.07	.05	-.02	-.06	.14	.11	*.73*
10	PES Experiences	6.37	2.75	.01	-.02	-.09	.06	.02	-.06	.15	.10	.46**

Notes. *N*=322; for variables 7 and 8: *N*=203. PES = Public Engagement with Science. * $p<.05$; ** $p<.01$. Cronbach's α for each scale are reported in the diagonal (in italics).

Table 6.5

Means and Standard Deviations, Broken Down by Conditions for Study A (Self-Criticism) and Study B (Reform Intentions)

Variable	Self-Criticism (Study 2A)		
	no self-criticism	no information	self-criticism
Expertise	4.56 (0.96) [a]	5.10 (0.72) [b]	5.14 (0.77) [b]
Integrity/Benevolence	4.23 (0.98) [a]	4.79 (0.75) [b]	5.00 (0.69) [b]
Credibility	3.63 (1.05) [a]	4.24 (0.78) [b]	4.28 (0.93) [b]
Willingness to Engage	3.42 (1.09) [a]	3.90 (1.17) [b]	3.92 (1.11) [b]
	Reform Intentions (Study 2B)		
	no intentions	undecided	reform intentions
Expertise	4.83 (0.94) [a]	4.86 (0.77) [a]	5.22 (0.68) [b]
Integrity/Benevolence	4.03 (1.04) [a]	4.38 (0.84) [b]	4.98 (0.71) [c]
Credibility	3.46 (1.04) [a]	3.76 (0.93) [b]	4.31 (0.73) [c]
Willingness to Engage	3.24 (1.23) [a]	3.61 (1.13)	3.70 (1.14) [b]

Notes. Study 2A: *N*=288; Study 2B: *N*=322. Means (standard deviations in brackets). In each line, different letters in the superscript indicate significant pairwise differences (i.e., $p<.05$; Tukey HSD test).

Discussion

Study 2 replicates and further qualifies the results from Study 1. Again, our findings indicate no harm in being self-corrective. Actually, Study 2 suggests that there is harm in *not* being self-corrective: While researchers expression of a self-critical attitude towards their previous findings (compared to no such expression) did not affect trustworthiness (expertise and benevolence/integrity), perceived credibility, or willingness to further engage with this research, researchers who expressed no self-criticism and presented their findings as fixed and definite were perceived as less trustworthy and less credible, and participants were less willing

to engage with this research (compared to a neutral control condition). Whereas in Study 1, the researcher in the self-criticism condition expressed doubts about his prior finding as a "second thought" regarding his methodology, Study 2 operationalized self-criticism as being mindful of the fragility and preliminary nature of (his) research in general. Researchers who deny this fragility and tentativeness of science might not only be perceived as less benevolent and integer, but also as less competent (as this attitude contradicts the basic idea of science as being self-corrective), which, in turn, questions the quality of their future research and makes them seem less of exemplary researchers.

Regarding reform intentions, our findings indicate a benefit of being self-corrective: Announcing reform intentions (compared to being undecided and/or intending no reforms) increased perceived trustworthiness (expertise and benevolence/integrity) and credibility of their future research, and led participants to report a higher willingness to further engage with this research. Nevertheless, we also find that dismissing reforms can harm the public's trust in science: When the researcher announced *not* to implement reforms, integrity/benevolence and credibility ratings were considerably lower. Although the effects of reform intentions in Study 2 were smaller than in Study 1, our findings support the same conclusion: Researchers' expressed positions on reform intentions have the potential of enhancing (as well as impairing) the public's trust and interest to engage with science.

General Discussion

In two studies, we demonstrate the effects of two ways of being self-corrective in science: expressing self-criticism and intending to implement

reforms. Our findings suggest that there is no harm in expressing criticism towards one's own research or in announcing to implement reforms. In fact, such self-corrective behavior was superior to non-corrective behavior in terms of laypeople's perceptions of researchers' epistemic trustworthiness (expertise and integrity/benevolence) and the perceived credibility of their research, and led to a higher willingness to further engage with researchers and their findings.

Researchers' self-criticism (i.e., reflecting critically upon prior studies and regarding scientific findings as preliminary and fragile) did not have negative effects. However, explicitly expressing *not* having doubts impaired trustworthiness (especially integrity and benevolence), credibility, and even participants' willingness to further engage with such research. Thus, in the eyes of the public, self-criticism does not harm, while a *lack* of self-criticism does. This has important implications for researchers communicating their findings: Openly expressing uncertainties and acknowledging the inherent preliminary nature of new scientific findings seems unproblematic for researchers' reputation. However, quite contrary, appearing overconfident might have considerable reputational costs.

Across both studies, researchers' reform intentions (i.e., planning to implement currently discussed reforms in future research) consistently led to more trust in and willingness to engage with science. Interestingly, these effects were driven both by the intention to implement reforms and, in the reverse direction, by an explicit dismissal of such reforms. Additionally, Study 1 suggests that the *extent* to which researchers are willing to implement reforms (i.e., minor vs. major reforms) does not play a decisive role for their public perception. These findings cast a new light on psychological research on the public perception of reforms in

psychology: Contrary to previous studies (Anvari & Lakens, 2018; Wingen et al., 2020), participants reacted quite positively to the idea of implementing reforms in science. Our studies extend this prior research in two important ways: First, previous studies focused on transparency as the main aspect of science reforms, but did not explain in detail how this connects to more reliable and credible research. One could argue that the connection is obvious; yet, explaining the link between transparency and higher reliability in more detail (and simpler words) to participants may have contributed to the positive effects of reform intentions on trustworthiness and credibility that we found in our studies. This also has important implications for communicating science reforms to the public: focusing on the superordinate goal these reforms aim to achieve (instead of merely portraying these reforms to be good for their own sake) might help lay audiences understand what they are about and why they are relevant. That said, it should be noted that we did not really measure participants' understanding of the consequences of science reforms for science as a whole; instead, we focused on participants' epistemic trustworthiness in one particular scientist. Thus, future research should look at how such individualized trustworthiness perceptions may generalize onto trust in science and the perceived credibility of science as a whole.

As scientific progress is not a solitary endeavor but a collaborative effort, researchers might also worry about their colleagues' perceptions of them. Future research should investigate how other researchers perceive their self-corrective peers. In fact, first evidence suggests that researchers receive wrongness admission of their colleagues positively (Fetterman & Sassenberg, 2015) and that, following (self-)correction, they indeed update their scientific beliefs in light of

such new evidence (yet, not as much as they should; McDiarmid et al., 2021).

Our findings suggest that researchers are perceived as more trustworthy and their research as more credible when they express self-criticism and reform intentions. One might argue that this pattern reflects nothing more than a positive acknowledgement of other people's humility (e.g., Chancellor & Lyubomirsky, 2013; Powers & Zuroff, 1988). However, it should be noted that many of our findings persisted after controlling for general likeability of the target person (i.e., the researcher in Studies 1 and 2), even though likeability ratings were highly correlated with trustworthiness (see Tables 6.1, 6.3, and 6.4): integrity/benevolence ascriptions at least can, thus, not be explained by such a "halo" effect. However, likeability strongly predicted participants' willingness to engage with the research and also suppressed the significant main effects of self-criticism and reform intentions on this DV in both studies. This suggests that the extent to which laypeople are motivated to learn more about science is contingent on their overall impression of a scientist. Importantly however, laypeople's ascriptions of integrity and benevolence—the "affective" dimensions of epistemic trustworthiness (Hendriks et al., 2015; McAllister, 1995)—are specifically affected by expressions of self-criticism and reform intentions, irrespective of more general likeability ratings.

Conclusion

Our findings suggest that researchers' hesitation toward self-correction (e.g., Fetterman & Sassenberg, 2015; Frewer et al., 2003; Rohrer et al., in press; van der Bles et al., 2020) seems unwarranted: there is no harm in openly admitting doubts and regarding one's findings as preliminary or in intending to reform one's

work routines. Quite contrary, researchers who portray their findings as fixed and definite and who are unwilling to implement reforms are perceived as less trustworthy and less credible by laypeople. In this regard, the current discussion of self-criticism and reforms (e.g., the open science movement) might prove to be an attention-drawing door opener for greater lay engagement with science; a chance for science to improve not only its methodological rigor, but also its relationship with the public.

References

Anvari, F., & Lakens, D. (2018). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology*, *3*(3), 266–286. https://doi.org/10.1080/23743603.2019.1684822

BBVA Foundation. (2011). *International Study on Scientific Culture. Understanding of Science.* BBVA.

Bishop, D. V. M. (2018). Fallibility in science: Responding to errors in the work of oneself and others. *Advances in Methods and Practices in Psychological Science*, *1*(3), 432–438. https://doi.org/10.1177/2515245918776632

Chancellor, J., & Lyubomirsky, S. (2013). Humble beginnings: Current trends, state perspectives, and hallmarks of humility. *Social and Personality Psychology Compass, 7*(11), 819-833. https://doi.org/10.1111/spc3.12069

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fetterman, A. K., & Sassenberg, K. (2015). The reputational consequences of failed replications and wrongness admission among scientists. *PLoS ONE*, *10*(12). https://doi.org/10.1371/journal.pone.0143723

Fiske, S. T., & Dupree, C. (2014). Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences*, *111*(Supplement_4), 13593–13597. https://doi.org/10.1073/pnas.1317505111

Frewer, L., Hunt, S., Brennan, M., Kuznesof, S., Ness, M., & Ritson, C. (2003). The views of scientific experts on how the public conceptualize uncertainty. *Journal of Risk Research*, *6*(1), 75–85. https://doi.org/10.1080/1366987032000047815

Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring laypeople's trust in experts in a digital age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOS ONE*, *10*(10), e0139309. https://doi.org/10.1371/journal.pone.0139309

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Disclose your flaws! Admission positively affects the perceived trustworthiness of an expert science blogger. *Studies in Communication Sciences*, *16*(2), 124–131. https://doi.org/10.1016/j.scoms.2016.10.003

Hendriks, F., Kienhues, D., & Bromme, R. (2020). Replication crisis = trust crisis? The effect of successful vs failed replications on laypeople's trust in researchers and research. *Public Understanding of Science*, *29*(3), 270–288. https://doi.org/10.1177/0963662520902383

Jensen, J. D. (2008). Scientific uncertainty in news coverage of cancer research: Effects of hedging on scientists' and journalists' credibility. *Human Communication Research*, *34*(3), 347–369. https://doi.org/10.1111/j.1468-2958.2008.00324.x

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *The Academy of Management Review*, *20*(3), 709. https://doi.org/10.2307/258792

McAllister, D. J. (1995). Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations. *Academy of Management Journal*, *38*(1), 24–59. https://doi.org/10.5465/256727

McDiarmid, A., Tullett, A. M., Whitt, C. M., Vazire, S., Smaldino, P. E., & Stephens, E. E. (2021). *Self-Correction in Psychological Science: How do Psychologists Update their Beliefs in Response to Replications?* PsyArXiv. https://doi.org/10.31234/osf.io/hjcm4

Neal, T., Guadagno, R., Eno, C., & Brodsky, S. (2012). Warmth and competence on the witness stand: Implications for the credibility of male and female expert

witnesses. *The journal of the American Academy of Psychiatry and the Law*, *40*, 488–497.

Powers, T. A., & Zuroff, D. C. (1988). Interpersonal consequences of overt self-criticism: A comparison with neutral and self-enhancing presentations of self. *Journal of Personality and Social Psychology, 54*(6), 1054–1062. https://doi.org/10.1037/0022-3514.54.6.1054

Retzbach, A., & Maier, M. (2015). Communicating scientific uncertainty: Media effects on public engagement with science. *Communication Research*, *42*(3), 429–456. https://doi.org/10.1177/0093650214534967

Rohrer, J. M., Tierney, W., Uhlmann, E. L., DeBruine, L. M., Heyman, T., Jones, B. C., Schmukle, S. C., Silberzahn, R., Willén, R. M., Carlsson, R., Lucas, R. E., Strand, J., Vazire, S., Witt, J. K., Zentall, T. R., Chabris, C. F., & Yarkoni, T. (2021). Putting the self in self-correction: Findings from the Loss-of-Confidence project. *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691620964106

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 word solution* (SSRN Scholarly Paper ID 2160588). Social Science Research Network. https://doi.org/10.2139/ssrn.2160588

van der Bles, A. M., Linden, S. van der, Freeman, A. L. J., & Spiegelhalter, D. J. (2020). The effects of communicating uncertainty on public trust in facts and numbers. *Proceedings of the National Academy of Sciences*, *117*(14), 7672–7683. https://doi.org/10.1073/pnas.1913678117

Wingen, T., Berkessel, J. B., & Englich, B. (2020). No replication, no trust? How low replicability influences trust in psychology. *Social Psychological and Personality Science*, *11*(4), 454–463. https://doi.org/10.1177/1948550619877412

7 Manuscript 3: When research is me-search

This manuscript has been published:

Altenmüller, M. S., Lange, L. L., & Gollwitzer, M. (2021). When research is me-search: How researchers' motivation to pursue a topic affects laypeople's trust in science. *PLOS ONE*, *16*(7), e0253911. https://doi.org/10.1371/journal.pone.0253911

[manuscript printed as accepted]

Abstract

Research is often fueled by researchers' scientific, but also their personal interests: Sometimes, researchers decide to pursue a specific research question because the answer to that question is idiosyncratically relevant for themselves: Such "me-search" may not only affect the quality of research, but also how it is perceived by the general public. In two studies (*N*=621), we investigate the circumstances under which learning about a researcher's "me-search" increases or decreases laypeople's ascriptions of trustworthiness and credibility to the respective researcher. Results suggest that participants' own preexisting attitudes towards the research topic moderate the effects of "me-search" substantially: When participants hold favorable attitudes towards the research topic (i.e., LGBTQ or veganism), "me-searchers" were perceived as more trustworthy and their research was perceived as more credible. This pattern was reversed when participants held unfavorable attitudes towards the research topic. Study 2 furthermore shows that trustworthiness and credibility perceptions generalize to evaluations of the entire field of research. Implications for future research and practice are discussed.

[163 words]

Key words: me-search, trust, credibility, attitudes, science reception.

When Research is Me-Search: How Researchers' Motivation to Pursue a Topic Affects Laypeople's Trust in Science

"Being a scientist is, at the most fundamental level, about being able to study what's exciting to you", says Jeremy Yoder, a gay man studying experiences of queer individuals in science [1]. Like Yoder, many researchers are passionate about their research and dedicated to their field. After all, they are free to choose research questions they deem important and are interested in. Freedom of science and research secures the independence of the academic from the political and other spheres. In return, researchers are expected to be neutral and objective and make their research process transparent to guarantee that this freedom is not exploited for personal gains.

Just as people differ in what they are interested in in their personal lives, researchers differ in what they find more or less fascinating and worth studying. Such fascination can have multiple causes and is often rooted in a perceived personal connection to a topic. For instance, Sir Isaac Newton allegedly became interested in gravity after an apple fell on his head [2]. A specific type of personal connection exists when researchers study a phenomenon because they are directly (negatively) affected by that phenomenon. In 1996, Harvard alumni and neuroanatomist Jill Bolte Taylor suffered a rare form of stroke that made her undergo major brain surgery, affected her personal and academic life tremendously, and eventually awakened her interest in studying the plasticity of the brain [3]. In 2006, she published an award-winning book covering her research and her personal story that led her to pursue this path. The Jill Bolte Taylor case is, thus, a prototypical example for such "me-search": researchers studying a phenomenon out of a particular personal affection by (or connection to) this

phenomenon. "Me-search" thus means pursuing a scientific question when the answer to that question is idiosyncratically relevant for the individual researcher (as opposed to when the answer is relevant per se).

Being directly affected by a phenomenon provides researchers studying it with a high degree of expertise and motivation: Jill Bolte Taylor, for instance, claims to bring a deep personal understanding and compassion to her research and work with patients [4,5]. That said, being personally affected may also come at the cost of losing one's scientific impartiality and neutrality for the subject: Jill Bolte Taylor was criticized for being overly simplistic in her scientific claims and mixing them with esoteric ideas, and for pushing her own agenda (i.e., selling her story) by dramatizing her own experiences [4–7].

While some criticized Jill Bolte Taylor heavily, the general public does not seem to have a problem with her research as "me-search". Her book is currently translated into 30 languages, and thousands of people visit her talks and keynote addresses [4–6]. Does that suggest that the general public tends to turn a blind eye on conflicts of interest that may arise from a researchers' personal affection by their research object? While the Jill Bolte Taylor case seems to suggest so, research on science communication and public understanding of science has shown that people are highly sensitive to potential conflicts of interest arising from researchers' personal involvement: perceiving researchers as pursuing an "agenda" for personal reasons is a major factor predicting people's loss of trust in researchers and science [8–11]. On the other hand, people may see personal ("autoethnographic") experiences of researchers personally affected by their topic as valuable and laudable – it may imply that "they know what they're talking about" [12–14]. Similarly, revealing a personal interest or even passion for a

particular research topic (e.g., due to being personally affected) could also overcome the stereotypical perception of scientists as distant "nerds in the ivory tower" [15,16]: researchers who openly disclose the idiosyncratic relevance of their research topic may appear more approachable, more likeable, and more trustworthy [17–19].

Thus, the public's reaction to "me-search" seems to be ambivalent and contingent on certain boundary conditions. Thus, the question we are going to address in this article is whether and when – that is, under which circumstances – a researcher's personal affection by a research topic ("me-search") positively vs. negatively impacts public perceptions regarding the *trustworthiness* of the respective researcher (and the entire research area in general) and the extent to which this researcher's findings are perceived as *credible*.

Perceivers' Motivated Stance as a Moderating Variable

This potentially ambivalent perception of "research as me-search" can be understood from a *motivated reasoning* [20] perspective: Laypeople receive and process information in a manner biased towards their own beliefs, expectations, or hopes. This also applies to the reception of scientific information [21,22]: For example, laypeople are more likely to dismiss scientific evidence if it is inconsistent with their beliefs [23,24] or if it threatens important (moral) values [25,26] or their social identity, respectively [27–29].

However, identity-related and attitudinal motivated science reception might differ in their underlying mechanisms. For identity-related motivated science reception, biased perception of information, which is relevant to a social identity, is driven by a defense motivation to protect this positive social identity [30]. Thus, identity-threating scientific information is countered by identity-protection efforts,

such as discrediting the findings and the source. These efforts will be more pronounced among strongly identified individuals [27–29]. For attitudinal motivated science reception, however, the mechanism might function as a broader perception filter. When confronted with new scientific information about the respective attitude object, the perceptual focus will be directed at clues helping to uphold prior attitudes (i.e., *confirmation bias* [31]): Potentially attitude-inconsistent information is attenuated, while potentially attitude-consistent information is accentuated. The ambivalent nature of "me-search" might allow to be easily bend in such a motivated manner and, thus, lead to biased perceptions of a researcher either way: when the findings are in line with one's prior beliefs, being personally affected may be considered an asset – the respective researcher is perceived as more trustworthy and his/her findings as more credible (compared to no idiosyncratic relevance). However, when the findings are inconsistent with one's prior beliefs, idiosyncratic relevance may be considered a flaw – the respective research is perceived as biased, untrustworthy, and less competent, and his/her findings are likely perceived as less credible than when idiosyncratic relevance is absent.

Prior research on motivated science reception mainly focused on laypeople's reactions towards specific scientific findings: *after* learning about the outcome of a particular study, participants dismiss the research (and devalue the researcher) if these outcomes are consistent vs. inconsistent with their prior beliefs [23–25,27–29]. However, people might be prone to motivated science reception even *before* results are known, judging researchers proverbially just by their cover (e.g., by biographical data, personal and scientific interests and motivations). People who hold positive attitudes towards a certain research topic

might perceive "me-searchers" as more trustworthy and anticipate their results to be more credible (before knowing the specific outcomes). By contrast, people who hold negative attitudes towards a certain research topic might trust "me-searchers" less and expect their findings to be less credible.

Additionally, motivated reception processes can be extended over and above the specific information under scrutiny and lead to questioning the scientific method in itself – a phenomenon termed the "scientific impotence excuse" [32]. In line with that, critical evaluations of specific researchers and their findings are sometimes generalized to the entire field of research [27]. Thus, the fact that a researcher engages in "me-search" might be interpreted in a way that fits best to one's preexisting convictions and may generalize to the entire field of research.

The Present Research

In two studies, laypeople read alleged research proposals concerning potentially polarizing research topics (i.e., LGBTQ issues and veganism) which were submitted by researchers who disclosed being either personally affected or not affected by the respective topic. We investigated whether (Study 1) and when (i.e., moderated by preexisting positive attitudes towards the respective research topic, Studies 1 and 2) such "me-search" information increased or decreased laypeople's perceptions regarding these researchers' epistemic trustworthiness and the anticipated credibility of their future scientific findings. Of note, we use the term "credibility" to differentiate evidence-related trust/credibility from person-related trust/credibility (i.e. "trustworthiness"). Further, we test whether one researcher's "me-search" impacts the evaluation of the entire respective field (Study 2).

For both studies in this paper, we report how we determined our sample

size, all data exclusions (if any), all manipulations, and all measures [33]. All materials, the anonymized data, and analyses are available online at the Open Science Framework (OSF; see https://osf.io/phfq3/). Before starting the respective study, informed consent was obtained. Participants read a GDPR-consistent data protection and privacy disclosure declaration specifically designed for the present study. Only participants who gave their consent could start the respective survey. According to German laws and ethical regulations for psychological research [34], gathering IRB approval is not necessary if (i) the data are fully anonymized, (ii) the study does not involve deception, (iii) participants' rights (e.g., voluntary participation, the right to withdraw their data, etc.) are fully preserved, and (iv) participating in the study is unlikely to cause harm, stress, or negative affect. The present studies met all of these criteria; therefore, no IRB approval had to be obtained.

Study 1

In our first study, we conducted an online experiment investigating the main effect of a researcher's disclosure of being personally affected vs. not affected by their research on their trustworthiness and the credibility of their future research. Further, we tested whether laypeople's preexisting attitudes towards the research topic moderate this effect.

Method

Sample. Four-hundred and eleven German participants were recruited via mailing lists and social networks. Ninety-seven participants had to be excluded due to pre-specified criteria: Sixty-seven participants failed the manipulation check; 25 participants failed the pre-specified time criteria (viewing the manipulation stimulus less than 30sec, taking less than 3min or more than 20min

for participation); 5 participants had apparently implausible response patterns (e.g., "straight-lining;" identical responses on every single item on more than one questionnaire page in a row). Eighty-five further participants failed the attention check. Excluding them did not change the overall results, so, for the sake of statistical power, we did not exclude these 85 cases. The final sample consisted of N=314 participants. We conducted sensitivity analyses using G*Power [35] for determining which effect sizes can detected with this sample in a moderated (multiple) regression analysis: At α=0.05 and with a power of 80%, small-to-medium effects ($f^2 \geq 0.03$) can be detected with this sample. Participants were mostly female (74% female, 25% male, 2% other) and their age ranged between 16 and 68 years (M=26.79; SD=10.18). Most participants were currently studying at a university (71%; working: 21%; unemployed or other: 8%). Participants who were currently studying or already had a university degree (93%) came from a variety of disciplines (law, economics, and social sciences: 49%; humanities: 16%; mathematics and natural sciences: 14%; medicine and life sciences: 11%; engineering: 4%).

Materials and Procedure. After obtaining informed consent, we asked participants to imagine they were browsing the website of a research institute and came across a short proposal for a new research project by a researcher named Dr. Lohr (no gender was indicated for greater generalizability and avoiding possible gender confounds). Next, participants read the beginning of the alleged proposal of a planned research project for which Dr. Lohr was allegedly applying for external funding. The text briefly introduced the planned project (i.e., investigating social reactions to queer employees at the workplace) and a statement of Dr. Lohr explaining why they were interested in conducting this

project. Participants were randomly allocated to two groups. In the "not personally affected" condition, Dr. Lohr wrote:

> "*I am interested in investigating this research topic in more detail not only out of scientific reasons but also because I – as someone who does not identify as homosexual and is not affected by my own research – really think we need more evidence-based knowledge about queer topics which we can implement in everyday life.*"

In the "personally affected" condition, Dr. Lohr wrote:

> "*I am interested in investigating this research topic in more detail not only out of scientific reasons but also because I – as someone who identifies as homosexual and is affected by my own research – really think we need more evidence-based knowledge about queer topics which we can implement in everyday life.*"

We added a definition for the word "queer" below the proposal: "*Queer is a term used as self-description by people who do not identify as heterosexual and/or who do not identify with the gender assigned at birth. The term is often used as umbrella term for LGBTQ (lesbian, gay, bisexual, trans and queer) and describes all people who identify as queer.*" After completing an attention check question (see pre-registration), we measured participants' trust in Dr. Lohr with the *Muenster Epistemic Trustworthiness Inventory* (METI; [36]), which was constructed for measuring trust in experts encountered online. It consists of 14 opposite adjective pairs measuring an overall trustworthiness score (Cronbach's α=.95) as well as the sub-dimensions expertise (e.g., competent – incompetent, Cronbach's α=.92) and integrity/benevolence (e.g., honest – dishonest, Cronbach's α=.93) on 6-point bipolar Likert scales. Factor analyses (see Appendix A in the

supplementary materials, https://osf.io/phfq3/) suggest that a two-factor model (with expertise and integrity/benevolence) fit the data better than a three-factor model (as suggested by [36]), corroborating the idea of a cognitive-rational dimension and an affective dimension of trustworthiness [37]. Next, participants rated the extent to which they found Dr. Lohr's research credible on a 6-point Likert scale ranging from 1="not at all" to 6="very much" (6 items, e.g., "I think Dr. Lohr's future findings will be credible;" "I will be critical of Dr. Lohr's research results" (reverse-coded); Cronbach's α=.84).

Next, we measured participants' own positive attitudes towards LGBTQ issues—the moderator variable in our design—with eleven statements developed from research on sympathy, group attitudes, and allyship [38,39] rated on a 6-point Likert scale ranging from "not at all" to "very much" (e.g., "I think that LGBTQ-related topics receive more attention than necessary" (reverse-coded); "I am open to learning more about concerns raised by LGBTQ people;" Cronbach's α=.93). Next, we conducted a manipulation check by asking participants to indicate whether Dr. Lohr disclosed being personally affected by their research ("Dr. Lohr stated being personally affected;" "Dr. Lohr stated not being personally affected;" "Dr. Lohr did not say anything about being affected or not").

Finally, we measured demographic variables (age, gender, occupation, academic discipline) and control variables: general perceptions of researchers' neutrality (self-developed 6-point bipolar scale with 4 adjective-pairs, e.g. subjective – objective, and 6 distractor pairs, e.g. introverted – extraverted, Cronbach's α=.81) and *Public Engagement with Science* (PES) with two measures adapted from a survey by the BBVA Foundation [40]: a 5-item scale measuring *PES frequency* (e.g., "How often do you read news about science?" 5-point Likert

scale ranging from 0="never" to 5="almost daily," Cronbach's α=.78) and a multiple choice question measuring 15 potential *PES experiences* during the last 12 months (e.g., "I know someone who does scientific research;" "I visited a science museum"). Participants had the opportunity to participate in a lottery and sign up for more information and were debriefed.

Results

Our randomized groups did not differ in regard to general perception of neutrality in science (p=.924) or PES (PES frequency, p=.709; PES experiences, p=.533). Table 7.1 summarizes all means, standard deviations, correlations and internal consistencies of the measured variables.

Main Effect of Being Personally Affected. First, we tested the main effect of the researcher's disclosure of being personally affected on epistemic trustworthiness and credibility of future findings. Laypeople trusted Dr. Lohr significantly more in the "personally affected" condition (M=4.92, SD=0.75) than in the "not personally affected" condition (M=4.66, SD=0.81), t(312)=2.93, p=.004, d=0.33, 95% CI_d [0.11; 0.56]. For credibility, the difference between the "personally affected" condition (M=4.15, SD=0.96) and the "not personally affected" condition (M=4.04, SD=0.86) was not significant, t(312)=1.02, p=.306, d=0.12, 95% CI_d [-0.11; 0.34]. Further exploring the two dimensions of epistemic trustworthiness, Dr. Lohr was perceived as higher on integrity/benevolence, t(312)=3.19, p=.002, d=0.36, 95% CI_d [0.14; 0.59], and on expertise, t(312)=2.17, p=.030, d=0.25, 95% CI_d [0.02; 0.47] when disclosing being personally affected.

Moderation by Pre-existing Attitudes. Second, we tested whether the effect of being personally affected by the research topic on trustworthiness was moderated by participants' pre-existing attitudes towards LGBTQ issues. Using

standardized linear regression, we again found a main effect of condition on trustworthiness, *beta*=0.15, *p*=.004, 95% CI_{beta} [0.05, 0.26]. There was a significant main effect of participants' pre-existing attitudes, *beta*=0.30, *p*<.001, 95% CI_{beta} [0.20, 0.40] and the condition × attitudes interaction effect was significant, *beta*=0.19, *p*<.001, 95% CI_{beta} [0.08, 0.29], increasing the amount of explained variance in trustworthiness by 3% to R^2_{adj}=.14. Table 7.2 summarizes the results. Figure 1a displays the interaction effect and standardized simple slopes analysis further qualifies it: Participants with more positive attitudes towards LGBTQ issues (+1 *SD* above sample mean) trusted Dr. Lohr more when the researcher was personally affected vs. not affected, *beta*=0.34, *p*<.001, 95% CI_{beta} [0.20, 0.49]. For participants with less positive attitudes towards LGBTQ issues (-1 *SD* below sample mean), this effect appears to be reversed, yet the simple slope was not significant, *beta*=-0.03, *p*=.646, 95% CI_{B} [-0.18, 0.11]. The same pattern of interaction effects emerged for both, integrity/benevolence (*p*=.009, total R^2_{adj}=.14) and expertise (*p*<.001, total R^2_{adj}=.10); full analyses are reported in Appendix B (see https://osf.io/phfq3/).

Regarding our second dependent variable, credibility, we found no main effect of condition, *beta*=0.04, *p*=.456, 95% CI_{beta} [-0.06, 0.13]. However, there was a significant main effect of participants' pre-existing attitudes, *beta*=0.48, *p*<.001 95% CI_{beta} [0.39, 0.58]: Participants with more positive attitudes anticipated a higher credibility of future research findings in this condition than participants with less positive attitudes. Similar to epistemic trustworthiness, there was a significant condition × attitudes interaction effect, *beta*=0.21, *p*<.001, 95% CI_{beta} [0.12, 0.31], increasing the amount of explained variance in credibility by 4% to R^2_{adj}=.26. Table 7.2 summarizes the results. Figure 1b displays this interaction

effect: Again, participants with more positive attitudes towards LGBTQ issues (+1 *SD* above sample mean) anticipated Dr. Lohr's future research findings to be more credible when the researcher was personally affected vs. not affected, *beta*=0.25, *p*<.001, 95% CI_{beta} [0.12, 0.38]. However, for participants with more negative attitudes (-1 *SD* below sample mean) this effect was significantly reversed: They rated the future research as less credible when the researcher was personally affected vs. not affected, *beta*=-0.18, *p*=.009, 95% CI_{B} [-0.31, -0.04].

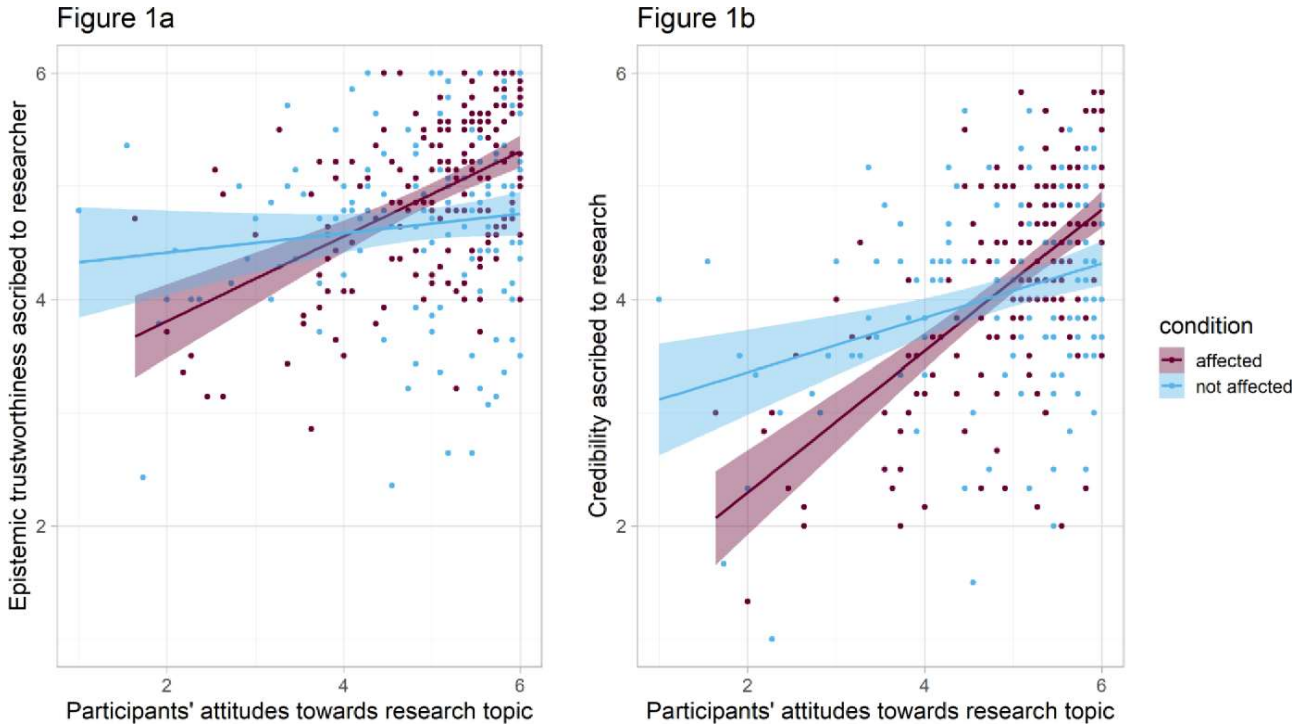


Fig 7.1. Linear regression plots for the interaction effect of attitudes × condition on epistemic trustworthiness (Figure 7.1a) and credibility (Figure 7.1b) with 95% confidence intervals: Participants' attitudes towards the research topic moderated how a researcher's disclosure of being personally affected (vs. being not personally affected) by one's own research was perceived.

Table 7.1

Means, standard deviations, correlations and internal consistencies

Variable	*M*	*SD*	α	1	2	3	4	5	6	7
1. Expertise	4.61	0.91	.92							
2. Integrity/benevolence	4.94	0.79	.93	.72***						
3. Epistemic trustworthiness	4.80	0.79	.95	.92***	.94***					
4. Credibility	4.10	0.92	.84	.68***	.68***	.73***				
5. Attitudes towards LGBTQ	4.93	1.02	.93	.23***	.32***	.30***	.47***			
6. Neutrality expectation	4.12	0.88	.81	.17**	.17**	.18**	.10	.08		
7. PES frequency	3.25	0.71	.78	-.08	-.09	-.09	-.08	-.07	.10	
8. PES experiences	6.92	2.98	-	-.11*	-.05	-.09	-.02	-.03	.05	.56***

Note. N=314. * indicates *p*<.05. ** indicates *p*<.01. *** indicates *p*<.001. α represents internal consistencies (Cronbach's α). Variables 1-6 ranged from 1 to 6, variable 7 ranged from 1-5 and variable 8 ranged from 0–15.

Table 7.2

Standardized regression results and semi-partial correlations, Study 1

Predictor	*beta*	*beta* 95% CI	sr^2	sr^2 95% CI
Epistemic trustworthiness				
condition	0.15**	[0.05, 0.26]	.02	[-.01, .05]
attitudes	0.30***	[0.20, 0.40]	.09	[.03, .15]
condition × attitudes	0.19***	[0.08, 0.29]	.04	[-.00, .07]
Credibility				
condition	0.04	[-0.06, 0.13]	.00	[-.01, .01]
attitudes	0.48***	[0.39, 0.58]	.24	[.16, .32]
condition × attitudes	0.21***	[0.12, 0.31]	.05	[.01, .09]

Note. N=314. ** indicates $p<.01$. *** indicates $p<.001$. Condition is standardized by effect coding (-1=not personally affected, 1= personally affected). *beta* represents standardized regression weights. sr^2 represents the semi-partial correlation squared.

Discussion

Results from Study 1 suggest that LGBTQ researchers are perceived as more trustworthy and their future findings as more credible when they disclose being personally affected by their research topic (i.e., being queer themselves), but only if perceivers hold positive attitudes towards LGBTQ issues. By contrast, holding less favorable attitudes towards LGBTQ issues lead to more skeptical reactions towards personally affected vs. unaffected researchers. This finding shows that learning about a researcher's personal affection by their research can, indeed, go both ways, as suggested by our theoretical reasoning. On a more general level, our research suggests that public reactions towards "me-search" is a matter of pre-existing attitudes, and, thus, a case of motivated science reception [21,22].

There are some limitations to this first study: As most people in our sample held rather positive attitudes towards the LGBTQ community (*M*=4.93, *SD*=1.02; on a scale from 1 to 6), predicted values on trustworthiness and credibility at the lower end of the attitude spectrum are probably less reliable. Also, we did not control for participants' own identification as belonging to the LGBTQ community. Thus, we cannot differentiate clearly between attitudinal and identity-related effects, which is important because attitudes and identity concerns have a psychologically distinguishable impact on motivated science reception [27,28]. Additionally, replicating our results in a different domain is necessary to be able to generalize our findings. Another question of generalizability that is left unanswered is how such individual experiences with one personally affected researcher might impact laypeople's perception of the entire field. This calls for more research on the double-edged nature of the moderating effect of preexisting attitudes.

Study 2

In our preregistered second study (see https://osf.io/c9r4e), we aimed to replicate our findings in a more diverse sample and with a different research topic that has the potential of polarizing participants even more strongly. We used the same design as in Study 1, but changed the proposed research topic to perceptions of vegans and introduced a vegan vs. non-vegan researcher. Again, we hypothesized that laypeople's attitudes towards veganism moderate the effects on trustworthiness as well as credibility of future research. Additionally, we tested whether the effect of one researcher being personally affected by their own research generalizes to the broader perception of their entire field. Furthermore, we also explored whether the moderation by attitudes towards veganism prevailed

when controlling for self-identification as being vegan (not included in preregistration).

Method

Sample. We conducted an a-priori power analysis using G*Power [35] for detecting the hypothesized interaction effect in a moderated multiple regression analysis (f^2=0.04, based on Study 1, with 1-β=0.90 and α=0.05, which resulted in a total sample of N=265. Anticipating exclusions (see specified criteria) of comparable size as in the previous study, we aimed for a sample of at least 350 participants.

We collected data from 364 participants via mailing lists and social media. Fifty-seven participants had to be excluded due to our preregistered criteria (see https://osf.io/c9r4e): one participant was younger than 16 years, 31 failed the manipulation check, 10 took less than 20sec viewing the proposal, 12 took less than 3min or more than 20min for participation, 3 had apparently implausible patters of response (i.e., "straight-lining;" identical responses on every single item on more than one questionnaire page in a row). The final sample consisted of N=307 participants (76% female, 23% male, 1 other) who were between 18 and 79 years old (M=33.55, SD=13.92). Approximately half of the sample (50%) was currently studying at a university, further 40% were working and 10% not working, one person was currently in training. Eighty-five percent were currently studying or already held a university degree (social sciences: 49%, humanities: 17%, natural sciences: 14%, life sciences: 8%, engineering: 6% and other 6%). Most participants did not consider themselves as vegans (89%).

Materials and Procedure. We used the same materials and procedure as in Study 1 (see OSF for full materials: https://osf.io/phfq3/). However, we changed the

research topic to "perceptions of vegans". Participants were randomly assigned to two conditions. In the "not personally affected" condition, the researcher Dr. Lohr wrote:

> "*I was interested in investigating this research questions not only out of scientific reasons but because, as someone who is not living as a vegan and, thus, not personally affected by my own research, I think we have a need for more evidence-based knowledge regarding the social embedding of vegan lifestyles, which we can acknowledge in everyday life.*"

In the "personally affected" condition, Dr. Lohr wrote:

> "*… because, as someone who is living as a vegan and, thus, personally affected by my own research, I think we have a need for more evidence-based knowledge regarding the social embedding of vegan lifestyles, which we can acknowledge in everyday life.*"

As dependent variables, we again used the 14-item METI [36] to measure epistemic trustworthiness, but we expanded the measure for credibility of future research by adding one more item ("I would express skepticism towards Dr. Lohr's future findings") to better capture the behavioral aspects of credibility (now: 7 items; Cronbach's α=.86). We also added a measure of participants' evaluation of the entire field (not the specific researcher) as a third dependent variable. This 12-item scale was adapted from a related study [28] (e.g., "I think researchers who do research on that topic sometimes lack competence," "I think it is difficult to apply results from this line of research to reality;" 6-point Likert scale ranging from 1= "not at all" to 6 = "very much;" Cronbach's α=.85). Next, participants' attitudes towards veganism (i.e., the moderator variable) were measured with a 14-item scale adapted from the attitude measure in Study 1 by changing and adding items

(e.g., "I think veganism is exaggerated" (reverse-coded) and "I can imagine being a vegan myself;" 6-point Likert scale ranging from 1= "not at all" to 6 = "very much;" Cronbach's α=.95).

To reduce exclusions after data collection, participants could proceed only if they answered all attention checks correctly (4 items; multiple choice). We added self-identification as vegan as a control variable ("Do you presently consider yourself a vegan?" yes/no); and an open-ended question about participants' opinion regarding the researcher being personally affected to explore how laypeople rationalize their opinion. These responses were later coded for valence (positive, negative, mixed, or neutral) and content (deductive and inductive coding) by two raters blind to the specific research question (see Appendix C in the supplementary materials, https://osf.io/phfq3/; interrater reliability for valence, Cohen's κ=.86, p<.01; and for content, Cohen's κ=.74, p<.01). Again, the questionnaire closed with a sign-up for a lottery and more information as well as a debriefing.

Results

Our randomized groups did not differ in regard to PES (PES frequency, p=.147; PES experiences, p=.101). However, they did differ significantly in regard to the general perception of neutrality in science (p=.049). Possible implications are addressed in the Discussion. Table 7.3 summarizes all means, standard deviations, correlations and internal consistencies. In the following, we report our findings for all three dependent variables (trustworthiness, credibility, evaluation of the entire field), consecutively.

Trustworthiness. First, we ran the standardized regression model for epistemic trustworthiness. There was neither a significant main effect of condition

on epistemic trustworthiness, *beta*=0.04, *p*=.482, 95% CI_{beta} [-0.07, 0.15] nor a significant main effect of attitudes towards veganism, *beta*=0.07, *p*=.205, 95% CI_{beta} [-0.04, 0.18]. However, the hypothesized condition × attitudes interaction effect was significant, *beta*=0.22, *p*<.001, 95% CI_{beta} [0.11, 0.34], increasing the amount of explained variance in trustworthiness by 4% to R^2_{adj}=.05. Table 7.4 summarizes the results. Figure 7.2a and standardized simple slopes analyses show that participants with more positive attitudes towards veganism (+1 *SD* above sample mean) trusted Dr. Lohr more when personally affected vs. not affected, *beta*=0.26, *p*=.001, 95% CI_{beta} [0.11, 0.42]. The interaction effect remained significant when controlling for participants' self-identification as being vegan (*p*<.001, total R^2_{adj}=.06). In secondary analyses, we explored the effects on the two facets of epistemic trustworthiness, separately. The same pattern of interaction effects emerged for both integrity/benevolence (*p*<.001, total R^2_{adj}=.08) and expertise (*p*=.005, total R^2_{adj}=.02); full analyses are reported in Appendix D in the supplementary materials (see https://osf.io/phfq3/).

Credibility. On credibility, there was no significant main effect of condition, *beta*=-.07, *p*=.146, 95% CI_{beta} [-0.17, 0.03] but a significant main effect of attitudes towards veganism, *beta*=.35, *p*<.001, 95% CI_{beta} [0.25, 0.45]. As predicted, the condition × attitudes interaction effect was also significant for credibility, *beta*=0.25, *p*<.001, 95% CI_{beta} [0.15, 0.35], increasing the amount of explained variance in credibility by 6% to R^2_{adj}=.21. Table 7.4 summarizes these results. Figure 7.2b and standardized simple slope analyses qualify the interaction effect: In line with the results for trustworthiness, participants with more positive attitudes (+1 *SD* above sample mean) anticipated Dr. Lohr's future findings to be more credible when personally affected vs. not affected, *beta*=0.18, *p*=.016, 95%

CI_{beta} [0.03, 0.32], while the conditional effect for participants with more negative attitudes (-1 *SD* below sample mean) changed its sign, *beta*=-0.32, *SE*(*B*)=0.14, *p*<.001, 95% CI_{beta} [-0.47, -0.18]. As before, the interaction effect remained significant when controlling for self-identification as being vegan (*p*<.001, total R^2_{adj}=.21).

Evaluation of the Field. Third, we investigated whether this moderation effect generalizes to the evaluation of the entire field of veganism research. There was no significant main effect of condition, *beta*=-.00, *p*=.989, 95% CI_{beta} [-0.10, 0.10] but a significant main effect of attitudes, *beta*=-.41, *p*<.001, 95% CI_{beta} [-0.51, -0.31]. Again, we found the hypothesized condition × attitude interaction effect, *beta*=-.27, *p*<.001, 95% CI_{beta} [-0.37, -0.18], increasing the amount of explained variance in critical evaluation by 7% to R^2_{adj}=.27. Again, Table 7.4 summarizes these results and Figure 7.2c and standardized simple slopes analyses further qualify the interaction effect: Participants with more positive attitudes towards veganism (+1 *SD* above sample mean) were less critical of research on veganism when Dr. Lohr was personally affected vs. not affected, *beta*=-0.28, *p*<.001, 95% CI_{beta} [-0.41, -0.14]. By contrast, this conditional effect was reversed for participants with more negative attitudes towards veganism (-1 *SD* below sample mean), *beta*=0.27, *p*<.001, 95% CI_{beta} [0.14, 0.41]. This interaction effect also remained significant when controlling for self-identification as being vegan (*p*<.001, total R^2_{adj}=.28).

Table 7.3

Means, standard deviations, correlations, and internal consistencies

Variable	*M*	*SD*	α	1	2	3	4	5	6	7	8
1. Expertise	4.68	0.91	.92								
2. Integrity/ benevolence	4.75	0.75	.90	.71***							
3. Epistemic trustworthiness	4.72	0.76	.94	.92***	.93***						
4. Credibility	3.97	0.96	.86	.66***	.70***	.73***					
5. Evaluation of field	2.95	0.78	.85	-.40***	-.43***	-.45***	-.64***				
6. Attitudes towards veganism	4.26	1.23	.95	.04	.15**	.11	.39***	-.46***			
7. Neutrality expectation	4.18	0.91	.86	.05	.06	.05	.08	-.26***	.02		
8. PES frequency	3.30	0.68	.75	-.13*	-.06	-.10	-.03	-.03	.11	.08	
9. PES experiences	6.47	3.25	-	-.14*	-.09	-.12*	-.09	.03	.07	-.02	.51***

Note. N=307. * indicates p<.05. ** indicates p<.01. *** indicates p<.001. α represents internal consistencies (Cronbach's α). Variables 1-7 ranged from 1 to 6, variable 8 ranged from 1-5 and variable 9 ranged from 0–15.

Table 7.4

Standardized regression results, Study 2

Predictor	*beta*	*beta* 95% CI	sr^2	sr^2 95% CI
Epistemic trustworthiness				
condition	0.04	[-0.07, 0.15]	.00	[-.01, .01]
attitudes	0.07	[-0.04, 0.18]	.01	[-.01, .02]
condition × attitudes	0.22***	[0.11, 0.34]	.05	[.00, .10]
Credibility				
condition	-0.07	[-0.17, 0.03]	.01	[-.01, .02]
attitudes	0.35***	[0.25, 0.45]	.12	[.05, .18]
condition × attitudes	0.25***	[0.15, 0.35]	.06	[.01, .11]
Critical evaluation of field				
condition	-0.00	[-0.10, 0.10]	.00	[-.00, .00]
attitudes	-0.41***	[-0.51, -0.31]	.17	[.09, .24]
condition × attitudes	-0.27***	[-0.37, -0.18]	.07	[.02, .12]

Note. N=307. *** indicates $p<.001$. Condition is standardized by effect coding (-1 = not personally affected, 1 = personally affected). *beta* represents standardized regression weights. sr^2 represents the semi-partial correlation squared.

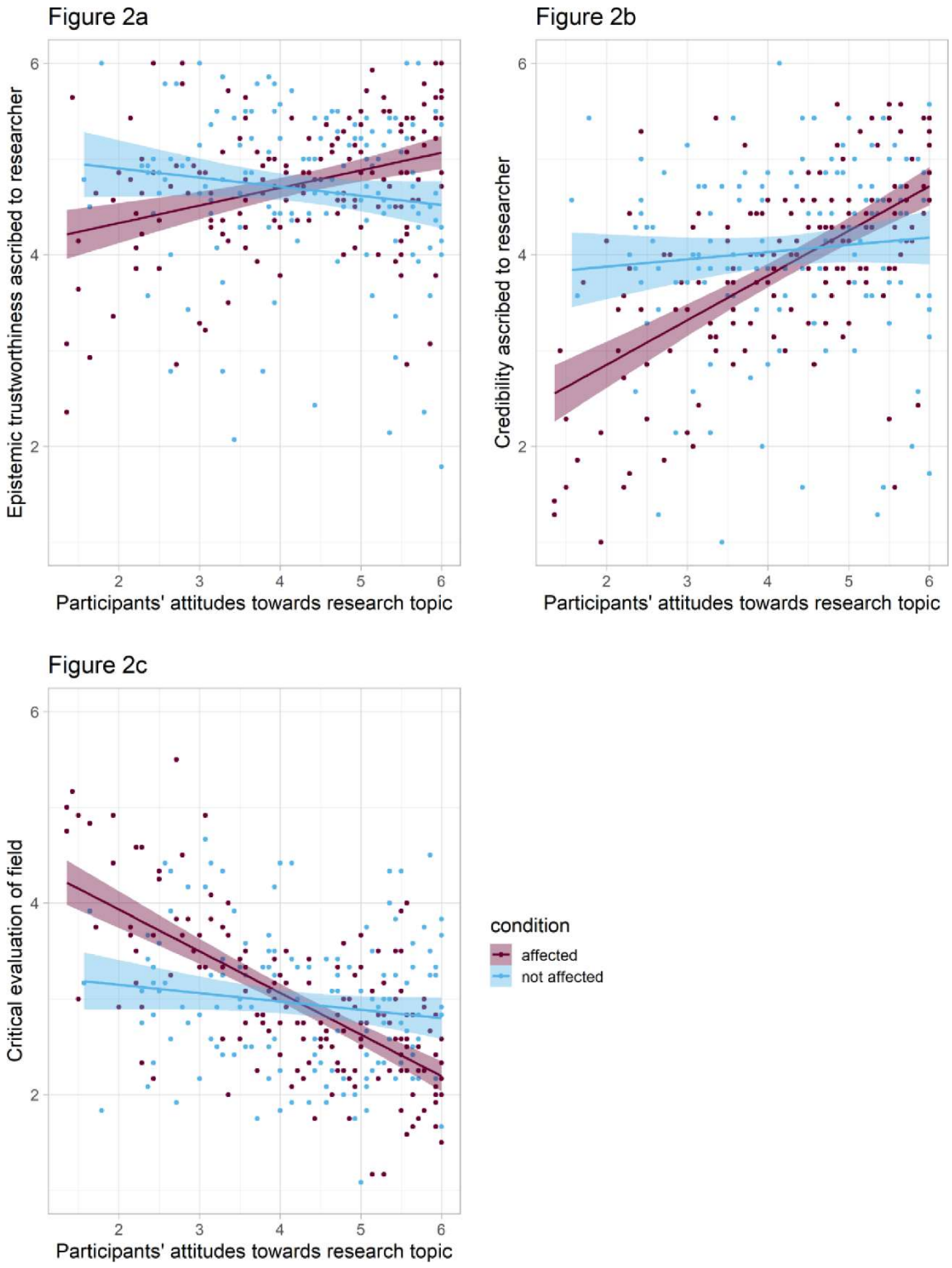


Fig 7.2. Linear regression plots for the interaction effect of attitudes × condition on epistemic trustworthiness (Figure 7.2a), credibility (Figure 7.2b) and critical evaluation of the entire field (Figure 7.2c) with 95% confidence intervals: Participants' attitudes towards the research topic moderated how a researcher's disclosure of being personally affected (vs. being not personally affected) by one's own research was perceived.

Participants' Opinion. Overall, participants who responded to the open-ended question expressed mostly negative opinions about the researcher being personally affected by his own research (negative: 48%, neutral: 21%, positive: 17%, and mixed: 14%). The most frequently mentioned (negative) remark was that a "me-searcher" might be biased towards their research (60%; e.g., "*By introducing himself as being affected, I fear he cannot evaluate the results of his research objectively*"). The second most frequently mentioned remark was that such idiosyncratic relevance is irrelevant (24%; e.g., "*It wouldn't make a difference*"). Positive remarks were mentioned less frequently: Participants ascribed more motivation (11%; e.g. "*I think interest, also personal interest, is an important prerequisite for determined research*") or knowledge about the topic (8%; e.g. "*Very good, most likely, he thus is knowledgeable about the subject matter and can conduct the study in a more purposeful manner*") to the "me-searcher", or recognized the transparency (7%; e.g., "*The main thing is transparency. People are always biased, perhaps even unconsciously*"; for more details, see Appendix C in the supplementary materials: https://osf.io/phfq3/).

Discussion

In Study 2, we replicated the moderation effect of preexisting attitudes on the effect of a researcher disclosing being personally affected (vs. not affected) by their own research on participants' epistemic trustworthiness and credibility ascriptions regarding the research and researcher's future findings. Further, we showed that this effect generalizes to the evaluation of the entire research area. Here, positive attitudes towards veganism determined how learning about an openly vegan researcher impacted participants' perceptions of trustworthiness and credibility as well as the evaluation of the entire field of veganism research

compared to learning about a non-vegan (i.e., non-affected) researcher. Participants who held more positive attitudes towards veganism reported more trust, higher anticipated credibility of future findings, and a less critical evaluation of the field when confronted with a vegan researcher. Conversely, for participants with less positive attitudes this effect was reversed. The moderation by positive attitudes towards veganism persisted when controlling for participants' self-identification as vegans. Overall, the interaction effects observed in Study 2 explained similar amounts of variance as in Study 1 (epistemic trustworthiness: 3% vs. 4%, and credibility: 4% vs. 6%). Further, qualitative analyses revealed that most participants reported negative – or, at least, mixed – perceptions of a "me-searcher" (e.g., "me-searchers" may be biased, but also highly motivation and knowledgeable), which corroborated our theoretical prediction that "me-search" may be a double-edged sword. Interestingly, these qualitative findings seem somewhat contradictory to the quantitative findings, according to which there was no main effect of researchers' idiosyncratic affection by their research topic.

In Study 2, one caveat is that the groups differed significantly in regard to participants' general expectations of neutrality in science. Participants who read about the personally affected researcher had weaker expectations of neutrality; yet, when added to the regression model as a control, the pattern of results remained unchanged (see Appendix E in the supplementary materials, https://osf.io/phfq3/). Further, as a second caveat, we show that participants generalized their perceptions to the overall field of veganism research. However, this research area might be considered quite narrow and, thus, future research should investigate how far such generalization processes stretch out to perceptions of broader areas of research (e.g., health psychology).

General Discussion

In two studies, we show that laypeople's perception of researchers who disclose being personally affected by their own research can be positive as well as negative: The effect of such "me-search" was moderated by laypeople's preexisting attitudes. Queer or vegan researchers were perceived as more trustworthy and their future findings were anticipated to be more credible when participants had positive, sympathizing attitudes towards the related research object (i.e., LGBTQ community or veganism). When participants' attitudes were less positive, this pattern reversed. In Study 2, we extended our research from individualized perceptions of single researchers and their findings to evaluations of the entire field of research. Participants who were confronted with a personally affected researcher seemed to consider this person a representative example and generalized their judgment to their evaluation of the entire (though here quite narrow) research area.

We explored epistemic trustworthiness in more detail in both studies, namely the cognitive-rational facet of expertise and the affective facet of integrity/benevolence: Both were impacted by researchers' disclosure of being personally affected, although effect sizes for expertise were descriptively smaller than for integrity/benevolence. This points to "me-search" – when received positively – possibly adding to the perception of competence-related aspects like a deeper knowledge of a phenomenon (e.g., via anecdotal insights) [12–14] and, even more so, warmth-related aspects like seeming more sincere, benevolent, transparent and, thus, approachable [15,16,41]. Disclosing such personal interest in a scientific endeavor might be able to bridge the stereotypical perception of cold and distant "science nerds" by revealing passionate, human and, thus, more

relatable side of a researcher. When received negatively, however, "me-search" might be regarded as harboring vested interests, which casts doubts on a researcher's neutrality and objectivity [8–11,42].

In general, the main models tested here explained between 5% and 28% of variance which may not appear impressive at first glance. However, our studies posed a very strict test of the effects of "me-search" by only using a subtle manipulation sparse in information followed by measures of very specific perceptions which might have contributed to an understatement of the real-world impact.

"Me-search" neither automatically sparks trust nor mistrust in laypeople, even if their explicit opinions seem rather negative. In line with assumptions from motivated science reception [22,43], our findings suggests that the ambivalence of the fact that a researcher is personally affected can be seized as an opportunity to interpret the situation in a manner that best fits to preexisting attitudes: Researchers, their findings and even their entire field of research are evaluated – even before learning about specific findings – based on prior attitudes towards the research topic. We show in Study 2 that the moderation effect of participants' positive attitudes towards the respective research topic (i.e., veganism) prevails when controlling for self-identification with the topic (i.e., being a vegan). This suggests that, indeed, in motivated reasoning attitudinal and identity-related processes can be differentiated: Here, social identity protection could be ruled out as alternative explanation for the effects of pre-existing attitudes. Noteworthily, we demonstrate that motivated science reception already operates when the results are not (yet) known. This points towards a perceptual filter made up of pre-existing attitudes that is activated when confronted with scientific information and leads to

biased pre-judgments: Ambivalent cues (i.e., "me-search") are prematurely interpreted in line with prior attitudes without actually knowing whether the new scientific information will be attitude-consistent or inconsistent (when, later, results are reported).

Future research. Future research on the motivated reception of "me-search" should focus on three open questions. First, while we consider it a strength of our studies that the results of the proposed research project were not yet known, it might be interesting to see how being personally affected or not interacts with the perceived direction of the communicated scientific results (e.g. supporting vs. opposing a certain position): To what extent can the first, premature evaluation of a "me-searching" researcher be adapted if the actual results are inconsistent with this pre-judgment?

Second, the investigation of what specific characteristics of "me-search" are instrumentalized by benevolent or skeptical perceivers might not only provide practical tips on how to handle being personally affected (e.g., in science communication) but also important theoretical insights on the building blocks of trust in science and researchers (see discussion above regarding the effects on the facets of epistemic trustworthiness). As one example, knowing that a qualitative level of knowledge is highly valued could further research on the trust-benefit of enriching statistical evidence with anecdotal and narrative elements [44,45]. As second example, we argue that researchers' self-disclosure of being personally affected by their research might signal transparency and, thus, improve the perception of the trust facets integrity and benevolence. Yet, even the disclosure of *not* being personally affected could have such an effect on a researcher's reputation and, at the same time, it might be less ecologically valid (as,

presumably, it is rather unusual to explicitly state to *not* be affected by something). Introducing a control group without any information about a researcher's relation towards their research object might bring light to this.

Third, we demonstrated the moderation effect of preexisting attitudes for two research areas (i.e., LGBTQ and veganism) and in different populations. Yet, further research should investigate whether this effect will hold up for other areas, more diverse samples and different kinds of "me-search", as well. For example, in some research fields being personally affected by the research might be perceived as more morally charged than in others and, thus, having stronger polarizing effects [46]: While, in veganism-research, "me-search" might be grounded in an ideological choice (e.g., thinking its morally wrong to consume animal products and, thus, being vegan), having a stroke and, following, studying stroke-related brain plasticity is likely perceived as less ideological. Also, different scientific methods (typically) used in a field might impact the perceptions of "me-search" depending on how prone for subjectivity these methods are perceived to be (e.g., qualitative "me-search" like autoethnographic analyses might be perceived more critically than when using seemingly objective, quantitative methods like physiological measures). Further, researchers who are not *directly* personally affected by their research but "merely" interested in something for personal reasons (e.g., being highly empathetic towards queer concerns without identifying as queer) might not profit from disclosure of such personal motivations: Such researchers might be perceived as impostors [47] lacking the expertise stemming from directly firsthand experiences.

Practical implications. Finally, for the applied perspective on public engagement with science, it should again be noted, that motivated reasoning

processes are activated even before specific results are presented (e.g. before hearing a talk or reading about a study). This might be important, as judgments are quickly formed and remembered [48,49] and, therefore, the first impression of a researcher might set the tone for further interactions and, particularly, for the acceptance and implementation of their findings. This emphasizes the importance of researchers knowing their audience (and their attitudes) when engaging in science communication.

Of course, there are also ethical considerations concerning "me-search": Researchers should always declare any conflict of interests when conducting research [50,51]. Failing to disclose being personally affected by one's own research might backfire severely on researchers' reputation – especially concerning their trustworthiness and the credibility of their findings – and in particular, when this information is disclosed by someone else and not themselves. At least for achieving positive reputational effects, it seems researchers need to freely initiate the disclosure of limitations and problems themselves [41,52]. A possible solution for reaping all the benefits and protecting against the potential harms of engaging in "me-search" might be to actively seek out mixed research teams. Including affected as well as non-affected individuals in research projects might be worth considering from the stance of the public's trust in science: It enables deep, even personal insights to the studied phenomenon, while still securing balanced perspectives and impartiality.

Conclusion

Neuroanatomist Jill Bolte Taylor became famous for turning her "stroke of fate" into productive and well-selling "me-search". Yet, she was praised as well as heavily criticized for mixing her personal and scientific motivations: When

research is also “me-search”, it can be perceived positively as well as negatively depending on laypeople’s preexisting attitudes towards the research object. Researchers who disclose being personally affected by their own research can benefit from this disclosure in terms of trustworthiness and credibility when it is perceived by laypeople with positive attitudes; however, for audiences with more negative attitudes this effect is reversed and disclosure can be harmful. One experience with a personally affected researcher might be enough to impact the evaluation of the whole field. Thus, openly acknowledging “me-search” in one’s research is an ambivalent matter and its communicative framing as well as the targeted audience should be well considered.

References

1. Hu JC. Studying the science of science. Science | AAAS [Internet]. 2016 Mar 23 [cited 2020 Sep 7]; Available from: https://www.sciencemag.org/careers/2016/03/studying-science-science
2. Keesing RG. The history of Newton's apple tree. Contemporary Physics. 1998 Sep;39(5):377–91.
3. Taylor JB. My Stroke of Insight. Hachette UK; 2009. 140 p.
4. Kaufman L. A Superhighway to Bliss. The New York Times [Internet]. 2008 May 25 [cited 2020 Oct 28]; Available from: https://www.nytimes.com/2008/05/25/fashion/25brain.html
5. Klitzman R. Myths About Our Right and Left Brains. Psychology Today [Internet]. 2013 Jan 7 [cited 2020 Oct 28]; Available from: http://www.psychologytoday.com/blog/am-i-my-genes/201301/myths-about-our-right-and-left-brains
6. Blech J. ESOTERIK : Mit einem Schlag ins Nirwana. Der Spiegel [Internet]. 2008 Jul 21 [cited 2020 Oct 28]; Available from: https://www.spiegel.de/spiegel/print/d-58302618.html
7. Vieth E. Can someone really know what it's like to have a stroke? A skeptic considers Jill Bolte Taylor's "Stroke of Insight." – Dangerous Intersection [Internet]. 2008 [cited 2020 Oct 28]. Available from: https://dangerousintersection.org/2008/04/15/can-someone-really-know-what-its-like-to-have-a-stroke-comments-regarding-stroke-of-insight/
8. Critchley CR. Public opinion and trust in scientists: the role of the research context, and the perceived motivation of stem cell researchers. Public Underst Sci. 2008 Jul;17(3):309–27.
9. Critchley CR, Nicol D. Understanding the impact of commercialization on public support for scientific research: Is it about the funding source or the organization conducting the research? Public Underst Sci. 2011 May 1;20(3):347–66.

10. Gierth L, Bromme R. Beware of vested interests: Epistemic vigilance improves reasoning about scientific evidence (for some people). Perales JC, editor. PLoS ONE. 2020 Apr 15;15(4):e0231387.

11. Rabinovich A, Morton TA, Birney ME. Communicating climate science: The role of perceived communicator's motives. Journal of Environmental Psychology. 2012 Mar 1;32(1):11–8.

12. Gardner SK, Hart J, Ng J, Ropers-Huilman R, Ward K, Wolf-Wendel L. "Me-search": Challenges and opportunities regarding subjectivity in knowledge construction. Studies in Grad and Post Doc Ed. 2017 Nov 13;8(2):88–108.

13. Ngunjiri FW, Hernandez K-AC, Chang H. Living Autoethnography: Connecting Life and Research. 1. 2010 Aug 23;6(1):E1–E1.

14. Story MD. Personal and professional perspectives on social nudism: Should you be personally involved in your research? The Journal of Sex Research. 1993 May 1;30(2):111–4.

15. Fiske ST, Dupree C. Gaining trust as well as respect in communicating to motivated audiences about science topics. Proceedings of the National Academy of Sciences. 2014 Sep 16;111(Supplement_4):13593–7.

16. Rutjens BT, Heine SJ. The Immoral Landscape? Scientists Are Associated with Violations of Morality. PLOS ONE. 2016 Apr 5;11(4):e0152798.

17. Collins NL, Miller LC, Collins NL, Miller LC. Self-disclosure and liking: A meta-analytic review. Psychological Bulletin. 1994;

18. Sprecher S, Treger S, Wondra JD. Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. Journal of Social and Personal Relationships. 2013 Jun 1;30(4):497–514.

19. Wheeless LR, Grotz J. The measurement of trust and its relationship to self-disclosure. Human Communication Research. 1977 Mar;3(3):250–7.

20. Kunda Z. The case for motivated reasoning. Psychological Bulletin. 1990;108(3):480–98.

21. Hornsey MJ. Why Facts Are Not Enough: Understanding and Managing the Motivated Rejection of Science. Curr Dir Psychol Sci. 2020 Nov 12;0963721420969364.

22. Lewandowsky S, Oberauer K. Motivated Rejection of Science. Current Directions in Psychological Science. 2016 Aug;25(4):217–22.

23. Greitemeyer T. I Am Right, You Are Wrong: How Biased Assimilation Increases the Perceived Gap between Believers and Skeptics of Violent Video Game Effects. PLOS ONE. 2014 Apr 10;9(4):e93440.

24. Nisbet EC, Cooper KE, Garrett RK. The Partisan Brain: How Dissonant Science Messages Lead Conservatives and Liberals to (Dis)Trust Science. The ANNALS of the American Academy of Political and Social Science. 2015 Mar 1;658(1):36–66.

25. Bender J, Rothmund T, Nauroth P, Gollwitzer M. How Moral Threat Shapes Laypersons' Engagement With Science. Pers Soc Psychol Bull. 2016 Dec 1;42(12):1723–35.

26. Rothmund T, Bender J, Nauroth P, Gollwitzer M. Public concerns about violent video games are moral concerns—How moral threat can make pacifists susceptible to scientific and political claims against violent video games. European Journal of Social Psychology. 2015;45(6):769–83.

27. Nauroth P, Gollwitzer M, Bender J, Rothmund T. Gamers against science: The case of the violent video games debate. European Journal of Social Psychology. 2014;44(2):104–16.

28. Nauroth P, Gollwitzer M, Bender J, Rothmund T. Social Identity Threat Motivates Science-Discrediting Online Comments. PLOS ONE. 2015 Feb 3;10(2):e0117476.

29. Nauroth P, Gollwitzer M, Kozuchowski H, Bender J, Rothmund T. The effects of social identity threat and social identity affirmation on laypersons' perception of scientists. Public Underst Sci. 2017 Oct 1;26(7):754–70.

30. de Hoog N. Processing of Social Identity Threats: A Defense Motivation Perspective. Social Psychology. 2013 Jan 1;44(6):361–72.

31. Nickerson RS. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology. 1998 Jun 1;2(2):175–220.

32. Munro GD. The Scientific Impotence Excuse: Discounting Belief-Threatening Scientific Abstracts. Journal of Applied Social Psychology. 2010;40(3):579–600.

33. Simmons JP, Nelson LD, Simonsohn U. A 21 Word Solution [Internet]. Rochester, NY: Social Science Research Network; 2012 Oct [cited 2020 Jul 23]. Report No.: ID 2160588. Available from: https://papers.ssrn.com/abstract=2160588

34. German Psychological Society, German Association of Psychologists. Berufsethische Richtlinien des Berufsverbandes Deutscher Psychologinnen und Psychologen e.V.und der Deutschen Gesellschaft für Psychologie e.V. 2016.

35. Faul F, Erdfelder E, Lang A-G, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods. 2007 May;39(2):175–91.

36. Hendriks F, Kienhues D, Bromme R. Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). Wicherts JM, editor. PLOS ONE. 2015 Oct 16;10(10):e0139309.

37. McAllister DJ. Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. AMJ. 1995 Feb 1;38(1):24–59.

38. Jones KN, Brewster ME, Jones JA. The creation and validation of the LGBT Ally Identity Measure. Psychology of Sexual Orientation and Gender Diversity. 2014 Jun;1(2):181–95.

39. Vossen HGM, Piotrowski JT, Valkenburg PM. The Longitudinal Relationship Between Media Violence and Empathy: Was It Sympathy All Along? Media Psychology. 2017 Apr 3;20(2):175–93.

40. BBVA Foundation. International Study on Scientific Culture. Understanding of Science. Bilbao, Spain: BBVA; 2011.

41. Hendriks F, Kienhues D, Bromme R. Disclose your flaws! Admission positively affects the perceived trustworthiness of an expert science blogger. Studies in Communication Sciences. 2016 Jan 1;16(2):124–31.

42. Sah S, Feiler D. Conflict of interest disclosure with high-quality advice: The disclosure penalty and the altruistic signal. Psychology, Public Policy, and Law. 2020;26(1):88–104.

43. Hertwig R, Engel C. Homo Ignorans: Deliberately Choosing Not to Know. Perspectives on Psychological Science. 2016 May;11(3):359–72.

44. Dahlstrom MF. Using narratives and storytelling to communicate science with nonexpert audiences. Proceedings of the National Academy of Sciences. 2014 Sep 16;111(Supplement_4):13614–20.

45. Moore A, Stilgoe J. Experts and Anecdotes: The Role of "'Anecdotal Evidence"' in Public Scientific Controversies. Science, Technology, & Human Values. 2009 Sep 1;34(5):654–77.

46. Kovacheff C, Schwartz S, Inbar Y, Feinberg M. The Problem with Morality: Impeding Progress and Increasing Divides. Social Issues and Policy Review. 2018;12(1):218–57.

47. Hornsey MJ, Jetten J. Not being what you claim to be: impostors as sources of group threat. European Journal of Social Psychology. 2003;33(5):639–57.

48. Ames DL, Fiske ST, Todorov AT. Impression formation: A focus on others' intents. In: Decety J, Cacioppo JT, editors. The Oxford handbook of social neuroscience. New York, NY, US: Oxford University Press; 2011. p. 419–33. (Oxford library of psychology).

49. Gilron R, Gutchess AH. Remembering first impressions: Effects of intentionality and diagnosticity on subsequent memory. Cogn Affect Behav Neurosci. 2012 Mar 1;12(1):85–98.

50. Gurney S, Sass J. Public trust requires disclosure of potential conflicts of interest. Nature. 2001 Oct;413(6856):565–565.

51. Resnik D. Disclosing Conflicts of Interest to Research Subjects: An Ethical and Legal Analysis. Accountability in Research. 2004 Apr 1;11(2):141–59.

52. Jensen JD. Scientific Uncertainty in News Coverage of Cancer Research: Effects of Hedging on Scientists' and Journalists' Credibility. Hum Commun Res. 2008 Jul 1;34(3):347–69.

8 Final Discussion

Stereotypical perceptions of researchers as highly competent yet only moderately warm (Fiske & Dupree, 2014; Imhoff, Koch, et al., 2018; Rosman & Merk, 2021) might challenge laypeople's trust in science. Trustworthiness is rooted in both warmth-related, socio-affective perceptions of benevolence and integrity as well as competence-related, cognitive-rational perceptions of expertise (Hendriks et al., 2015, 2016; Mayer et al., 1995; McAllister, 1995; Morrow et al., 2004). Thus, laypeople's perceptions of researchers as lacking warmth could impede their trust in science. To overcome these stereotypical perceptions, researchers could reveal personal information (e.g., details from their lives, self-critical thoughts about their work, or their personal involvement with their research topic) while communicating scientific findings to the lay public and, thus, present themselves as more approachable, warm, and likable to strengthen their perceived affective trustworthiness. However, such self-disclosure can also have detrimental effects, particularly regarding perceptions of researchers' expertise due to violating expectations of appropriateness, professionalism, and objectivity. Resulting changes in researchers' trustworthiness ascriptions could be extended to the perceived credibility of their scientific findings and even generalized to laypeople's higher-order trust in science.

In this dissertation, I presented three manuscripts including eleven experiments testing the effects of researchers' self-disclosure in science communication on laypeople's trust in science (i.e., their perceptions of researchers' epistemic trustworthiness and the credibility of their findings). The present findings demonstrate ambivalent effects of self-disclosure: First, basic, illustrative self-disclosure (i.e., framing information which is irrelevant to the

communicated science as either personal vs. neutral) produced small, volatile trade-off effects on researchers' perceived trustworthiness — slight increases in ascriptions of affective trustworthiness, yet also slight decreases in ascriptions of expertise — but not on the perceived credibility of their findings, and these effects are presumably of little practical relevance. Second, self-disclosure that conveyed researchers' self-critical stance towards science (i.e., expressing self-criticism and reform intentions) led to positive effects on laypeople's ascriptions of trustworthiness and credibility and even increased laypeople's willingness to further engage with science. Third, researchers' disclosure of being a "me-searcher" (i.e., being personally affected by their own research) had positive as well as negative effects on trustworthiness and credibility ascriptions depending on laypeople's preexisting attitudes. These individualized perceptions of one "me-searcher" were generalized to perceptions of the entire research area.

Evidently, self-disclosure *per se* is no effective remedy for stereotypical perceptions of researchers as lacking warmth as it has no uniform effects on laypeople's trust in science. Instead, findings in this dissertation suggest that the potential of researchers' self-disclosure to promote public trust in science crucially depends on the informative value of the specific content which is revealed. For example, self-disclosure can have positive effects if it implies "good" scientific practice (i.e., self-correction, Manuscript 2) or if laypeople can use it to upvalue science that is congruent with their prior attitudes (i.e., ascribing outstanding expertise and motivation to a personally affected researcher, Manuscript 3). However, this also entails that, other times, self-disclosure can have negative effects. For example, if it implies "bad" scientific practice (i.e., *lacking* self-correction, Manuscript 2) or if it can be used to devalue science that is

incongruent with laypeople's attitudes (i.e., assuming "me-searchers" to be biased, Manuscript 3).

8.1 Open Questions and Limitations

Of course, this research program left some questions unanswered and has important limitations. In the following, I will outline three issues that are common to all studies presented here.

8.1.1 Effect sizes and heterogeneity

Across all studies, we primarily detected small effects. While these might be common for social psychological research in general (e.g., Richard et al., 2003; T. Schäfer & Schwarz, 2019) and self-disclosure research in particular (e.g., Collins & Miller, 1994; Dindia, 2002; Henretty et al., 2014), it needs to be discussed how much this research program can contribute to theory and practice. Of note, most experimental manipulations applied in this program were rather subtle and information-sparse to avoid producing demand effects. It is thus possible that the present studies underestimate the "real-world" effect of self-disclosure in science communication. At the same time, however, most studies presented here used stimulus material set in very applied research contexts and, thus, the empirical effects could also be not far off from the real-world impact.

Theoretically, these small effect sizes point out that self-disclosure in science communication can only be one part of the attempt to counter stereotypical perceptions of scientists and to promote trust in science. It is worth discussing how, at least in Manuscript 1, replicability of effects across studies was low and effect sizes were quite heterogeneous. This might be explained by contextual factors (cf., Gollwitzer & Schwabe, 2021) that could further inform theory, in particular regarding the question of which effects moderate self-

disclosure in science communication. In Manuscript 1, we tested possible candidates (i.e., researchers' gender and discipline, opportunity for exchange, and role of discloser) and, later, I also conducted exploratory analyses with the data sets regarding possible effects of participant characteristics (e.g., age, gender, general engagement with science). None of these tested variables had a consistent, meaningful influence on the effects of self-disclosure on trust in science. Evidently, there are a lot of open questions regarding potential boundary conditions of the effects of self-disclosure on trust in science.

Practically, a cost-benefit assessment helps determine the usefulness of these effects: Which "costs" (e.g., compromising one's expertise, revealing personal information that might be used against oneself) can be tolerated for the presumably rather small benefits of engaging in self-disclosure? If the risk of revealing personal information is low (e.g., among neutral or positively inclined audiences), then disclosing some, appropriate (e.g., not too intimate) warmth-related information might be an easy, effortless way to reap some increase in perceived affective trustworthiness (especially in contexts where researchers' expertise is well-established).

8.1.2 Specificity of self-disclosure

One explanation for the detected heterogeneity of effects could be the content-specificity of self-disclosure. Commonly, self-disclosure is defined very broadly as intentionally revealing personal information to others (e.g., Cozby, 1973; Greene et al., 2006; Omarzu, 2000). However, this is problematic for specification: If self-disclosure refers to revealing *any* information, but the type of information is likely responsible for diverse effects, how then can we hope to formulate specific predictions for engaging in self-disclosure in general? In this

regard, previous research tried to refine theory on self-disclosure by, for example, considering the degree of intimacy or valence of the disclosed information (e.g., Caltabiano & Smithson, 2010; Cayanus & Martin, 2008; Collins & Miller, 1994; Henretty et al., 2014; Rains & Brunner, 2018) with mixed success. Findings from this dissertation, too, point to the importance of defining the type and degree of self-disclosure more precisely as effects varied considerably across (and within) manuscripts where we applied a broad definition of self-disclosure.

While the studies in Manuscript 1 clearly compared self-disclosure with *no* self-disclosure, manipulations in Manuscripts 2 and 3 were less precisely operationalized in this regard. These articles included studies that compared conditions in which researchers' communicative behavior might be considered self-disclosure either way: It might be just as much "self-disclosure" to communicate a self-corrective stance towards science versus communicating *not* being self-corrective (Manuscript 2) and to disclose being affected versus being *not* affected by one's research (Manuscript 3). Still, one could perceive openly engaging in self-correction (vs. being confident with the status quo) and revealing being personally affected by something (vs. not affected) to be considerably more intimate, and thus, as a more salient (and informative) form of disclosure. Future research on these topics might benefit from including measures of the degree of *perceived* self-disclosure similar to the manipulation check used in the studies in Manuscript 1, and thereby providing insights in the more specific effects of the extent of self-disclosure. And, importantly, it is evident that a more precise definition of self-disclosure is needed.

Furthermore, theory on self-disclosure would profit from empirical research aimed at specifying the mechanisms underlying the effects of self-disclosure on

person perception outcomes like trustworthiness. In addition to perceived closeness (see Manuscript 1), a promising direction might be to focus on perceptions of similarity conveyed by self-disclosure. For example, as discussed in the Introduction chapter, negative effects of self-disclosure could be explained by more revealed information leading to more opportunities to find *dissimilarities* to one's interaction partner (Norton et al., 2007). Contradicting this, we found significant positive correlations of the extent of perceived self-disclosure and similarity, as well as positive correlations between perceptions of similarity and the facets of trustworthiness and credibility (see Manuscript 1). It seems that perceived similarity towards a researcher might be an important pathway mediating the effects of self-disclosure on trust in science. Here, future research should further investigate, specify, and disentangle the underlying psychological mechanisms at play.

8.1.3 Limited generalizability

The findings of the presented research program are limited in their generalizability – especially in two regards: across samples and across scientific contexts. First, how generalizable are the findings to other samples? As in most studies in the social sciences, our participants came from convenience samples drawn from WEIRD societies (i.e., western, educated, industrialized, rich, and democratic; Henrich et al., 2010) and the samples mostly skewed female, young, and scientifically literate (e.g., holding, or currently studying for, a university degree). In general, I agree with the need to study more diverse, non-weird samples (e.g., Ghai, 2021; Henrich et al., 2010; Lewis, 2021). When studying trust in science, it might be particularly important to focus on recruiting diverse participants from the full attitude and trust spectrum, especially those who are

skeptical towards science. Manuscript 3 demonstrates the role of sample diversity quite powerfully: In Study 1, on average, participants displayed strong pro-LGBTQ attitudes, producing positive main effects for a researcher's disclosure of being homosexual (while proposing a study on LGBTQ perceptions). Study 2 of this manuscript puts this into perspective: Here, participants' attitudes towards veganism were a lot more diverse, making the main effect of a researcher's disclosure of being vegan (and proposing a study on veganism perceptions) disappear – only the interaction with their prior attitudes towards veganism now explained significant variance. Thus, only studying participants with positive attitudes towards science or the topic in question could produce unrealistically positive effects (e.g., increases in trust in science). This could mean that, when communicating science to the broad public, engaging in communicative strategies like self-disclosure might backfire, especially among science-skeptical samples.

Second, how generalizable are the findings in this dissertation to other scientific contexts? As argued in the Introduction chapter, science reception is likely domain-specific (e.g., Buehl et al., 2002; Muis et al., 2006; Scheitle & Guthrie, 2019). However, in this research program, the effects of self-disclosure in science communication were only studied regarding specific domains (i.e., social sciences, medicine, marine science) and in applied contexts. Additionally, the present theoretical rationale builds on laypeople's science stereotypes which might differ depending on the scientific domain (see Study 5, Manuscript 1) and could, thus, impact the influence of researchers' self-disclosure on laypeople's science reception. This raises the question of whether it is possible to generalize the present evidence to other fields of disciplines (e.g., humanities) or to the reception of basic research (e.g., astrophysics). Further, we mostly used fictitious findings

from scientific contexts that were unlikely to be strongly polarizing among our samples (except Manuscript 3). Additionally, considering the potential role of prior knowledge and scientific literacy (e.g., Allum et al., 2008; Bauer et al., 2007; Drummond & Fischhoff, 2017), it is unclear whether the findings presented here would generalize to trust in more familiar and broadly discussed scientific evidence (e.g., on climate change).

8.2 Conceptualization of trust in science

In this dissertation, I determine trust in science as laypeople's person-related perceptions of researchers' epistemic trustworthiness and their content-related perceptions of the credibility of scientific findings, and I further divide these concepts into different facets. However, these conceptualizations as well as their relation to each other are worthy of discussion. In the following, I reflect on the present empirical insights into the concept of trust in science and derive implications for future research and practice.

8.2.1 Conceptualization of researchers' trustworthiness

In line with theoretical considerations by Hendriks and colleagues (2015, 2016), I defined researchers' trustworthiness as consisting of expertise, integrity, and benevolence. However, I further determined the facets of integrity and benevolence as warmth-related facets of trustworthiness (collapsing them into the factor "affective trustworthiness") and expertise as the competence-related factor of trustworthiness. While there are strong theoretical arguments for such a two-dimensional perspective on researchers' trustworthiness (e.g., Fiske & Dupree, 2014; McAllister, 1995; Morrow et al., 2004), here, the decision to use only two dimensions of trustworthiness was predominantly data-driven and, thereby, dependent on the measurement instrument. Across all studies, we used the

Muenster Epistemic Trustworthiness Inventory (METI, Hendriks et al., 2015), which was originally designed as a three-factorial instrument measuring expertise, integrity, and benevolence. However, for data sets in all three manuscripts, correlations as well as factorial analyses suggested a better fit of a two-factorial solution with the factors *expertise* and *affective trustworthiness* (i.e., integrity + benevolence). Of note, even in the original article introducing the METI (Hendriks et al., 2015), integrity and benevolence are more closely linked to each other than to expertise. One explanation for the deviation from its originally intended three-faceted structure might be the lack of items referring to the theoretical core of benevolence (e.g., benevolent, prosocial, other-oriented; cf., Hendriks et al., 2016); thus, the instrument might be less content-valid for this facet.

Despite these shortcomings, I used the METI because of its consistently high reliability, practical applicability and due to the lack of a better-validated alternative. However, if the measurement of the benevolence facet is indeed imprecise, this could pose a serious problem for the test of the theoretical rationale proposed in this dissertation: Benevolence seems theoretically and empirically most closely related to the warmth dimension of person perception. Therefore, benevolence should be the most diagnostic target outcome when testing whether researchers' self-disclosure in science communication has the potential to overcome laypeople's views of researchers as rather cold and, thus, to increase (affective) trustworthiness ascriptions. Therefore, the lack of effects or small effect sizes detected in this research program might also be explained by inadequacies of the measurement instrument.

Beyond, while it is in line with theoretical as well as empirical considerations to conceptualize researchers' trustworthiness with two broad facets

of competence-related expertise and warmth-related affective trustworthiness, this does not imply that the originally proposed conceptualization of trustworthiness as three faceted (i.e., expertise, benevolence, and integrity) is void. For the present research, building on assumptions based on the two-dimensional Stereotype Content Model (Fiske et al., 2002; Fiske & Dupree, 2014), it was fitting to conceptualize trustworthiness in a parallel, two-dimensional fashion. However, theoretically and practically it might be useful to consider a more fine-grained picture. For example, as outlined in the Introduction chapter, the German Science Barometer 2021 clearly demonstrates that laypeople's reasons to trust or distrust science can be differentiated in regard to their perceptions of researchers' integrity, benevolence, and expertise. Further, building on the present rationale, interventions aimed at increasing trust in science could target perceptions of researchers' integrity (e.g., by making research routines more transparent) as well as their benevolence (e.g., by outlining researchers' prosocial intentions). Thus, instead of discarding the idea of three facets of trustworthiness, it might be necessary to refine measurement instruments to better capture the full psychological construct.

8.2.2 Conceptualization of credibility

In the Introduction chapter, I pointed out that, despite a wealth of research, there is still a lack of conceptual clarity regarding the structural aspects of credibility. What does credibility — psychologically — entail? Is it just the cognitive acceptance of a finding as "true" (cf., Appelman & Sundar, 2016) or are behavioral aspects part of it as well (e.g., as implied by Rieh & Danielson, 2007, and Wathen & Burkell, 2002)? An inclusive perspective on credibility (i.e., including cognitive as well as behavioral elements) considers that information

which is perceived as "credible" is not only integrated into one's cognitive knowledge system but also informs one's behavior. Thus, this captures a broader picture of credibility. In the studies included in this dissertation, I used such an inclusive definition of credibility with a cognitive as well as a behavioral facet. However, comparing this approach to other research on the credibility of science, this appears to be quite unusual. Commonly, credibility is measured only in regard to its cognitive facets (e.g., "How accurate, believable and authentic is the presented content?", Appelman & Sundar, 2016). Behavioral elements are, if at all, considered as downstream consequences of scientific credibility or of a broader trust in science (e.g., adhering to COVID-19 protective measures; Dohle et al., 2020; Pagliaro et al., 2021; Plohl & Musil, 2021). This is in line with psychological models of behavior such as the Theory of Planned Behavior (Ajzen, 1985, 1991), which considers behavioral intentions as a product of attitudes, norms, and behavioral control / self-efficacy. Thus, behavioral intentions informed by scientific findings are likely determined by more aspects than "just" their credibility. This might call for a narrower definition of credibility as used here, namely, by excluding behavioral intentions.

Findings in this dissertation also paint an ambivalent picture of the structure of credibility: Across all three manuscripts, factor analyses do not come to consistent conclusions whether the credibility items used in the respective studies represent one or two factors. In particular, item loadings on two-factor-solutions tended to be messy and not fully consistent with the assumed cognitive and behavioral facets. Nevertheless, even if two factors emerged, these always correlated highly. Thus, the measures of credibility applied in this research program might – despite being theoretically derived and assumedly highly content

valid (assuming a broad concept of credibility) – not be fully matured yet. Future research should aim for a more in-depth analysis of these issues and develop better-validated measures to capture the credibility of scientific findings.

While, as outlined above, there are arguments for using less inclusive conceptualizations (and, thus, measures) of scientific credibility, there could also be value in thinking about credibility in broader terms. One interesting idea to capture an inclusive concept of credibility is to consider credibility as an *attitude* towards scientific findings. Leaning on three-component models of attitude structure (e.g. Breckler, 1984; Rosenberg & Hovland, 1960; Zanna & Rempel, 1988), one could define not only *cognitive* and *conative* (i.e., behavioral) aspects of credibility but also *affective* elements (e.g., feelings towards the findings). This approach might be particularly informative regarding influences of motivated science reception: As information that contradicts one's world view elicits feelings of threat (e.g., Evans & Fetterman, 2021; Nauroth et al., 2015, 2017) which leads to motivated reception (e.g., ascribing less credibility to findings) to attenuate that threat, there are arguably affective processes involved in forming credibility judgments. Therefore, a three-component concept of credibility might be able to provide more specific insights into how laypeople form credibility judgments regarding scientific evidence.

8.2.3 Relationship of trustworthiness and credibility

Beyond their individual structures, what can we learn about the relationship of perceptions of researchers' trustworthiness and credibility from this research program? As outlined in the Introduction chapter, the two concepts summarized under the umbrella of "trust in science" are likely, though not necessarily, closely related. Here, their association is corroborated by correlations across all studies

included in this dissertation: Both affective trustworthiness (i.e., integrity and benevolence) and expertise correlated similarly strongly with credibility, but specific correlations somewhat varied. For example, in Manuscript 1, we find correlations between judgments of researchers' expertise and acceptance of their findings ranging from .30 to .62. This variability suggests that the size of the relationship of trustworthiness and credibility (e.g., how much laypeople rely on the source to judge message credibility) depends on other factors, as also pointed out in the Introduction. The lowest correlation in Manuscript 1 was found in the field study. As the scientific information was presented as part of an exhibition in a reputable and famous science museum, the trustworthiness of the audio speaker accompanying this exhibition was evidently less important for judging the credibility of the presented science. Additionally, correlations were, on average, larger in Manuscripts 2 and 3, where credibility was not ascribed to actual research findings; rather, participants judged the *anticipated* credibility of future findings. In these situations, first-hand credibility judgments of scientific findings were not possible and necessarily relied more on the presently available information: the impressions of researchers' trustworthiness. Furthermore, it should be noted that the statistical association of trustworthiness and credibility might also be (partly) spurious as third variables (e.g., general epistemic beliefs) might inflate the observed statistical correlation.

From a more practical standpoint, the strong association of trustworthiness and credibility is encouraging (if it is not merely spurious) because it suggests that, even though the effects of self-disclosure in science communication on trustworthiness were not particularly powerful, researchers possibly do hold sway over laypeople's science reception to some degree. Researchers who present

themselves in a way that increases perceptions of their trustworthiness (e.g., by appearing more approachable, warm, similar) could, thereby, increase credibility ascriptions of their scientific claims (assuming that this correlation represents, at least in large parts, a causal effect of trustworthiness on credibility; cf., Siegrist, 1999; Terwel et al., 2009). Thus, optimistically, researchers appearing trustworthy likely plays a key role for laypeople's acceptance of science and their behavioral intentions to act in line with scientific findings (e.g., getting vaccinated when trustworthy researchers communicate its effectiveness). Going beyond that, the speculation that third variables like general epistemic beliefs are a driving factor for trust in science (i.e., contributing to potentially spurious correlations of trustworthiness and credibility) also makes the case for the importance of not only focusing on being trustworthy science communicators but also building a solid basis of scientific knowledge and epistemic beliefs among the lay public. Note, however, that this is speculative and was not tested in this dissertation. Still, next to communication-oriented approaches to trust in science, learning-oriented approaches should not be disregarded (cf., Bromme & Goldman, 2014). For example, science education that improves scientific literacy is another important pillar of laypeople's willingness to adhere to scientific evidence.

Beyond the general relationship of trustworthiness and credibility, it is interesting to analyze how their facets relate to each other: Which facet of epistemic trustworthiness is more important for forming credibility judgments, or, more specifically, for affecting cognitions and behavior (i.e., the cognitive and behavioral facet of credibility)? Surveys like the Science Barometer seem to suggest that expertise perceptions might be the most important driver for laypeople's overall trust in science (Wissenschaft im Dialog/Kantar, 2021);

however, researchers' expertise might just be what comes most readily to mind when thinking about what makes science credible. Interestingly, laypeople seem to react more intensely to cues indicating researchers' lack of integrity or benevolence than to cues indicating a lack of expertise, highlighting the relative importance of these facets (e.g., Hendriks et al., 2015). This is in line with past research demonstrating the primacy of warmth-related judgments over competence-related judgments in evaluating distant others (e.g., Wojciszke & Abele, 2008). However, other scholars reviewing the literature concluded that the differential weights of the trustworthiness facets are unclear (Pornpitakpan, 2004). In fact, in the present studies, expertise and affective trustworthiness (i.e., integrity and benevolence) are very similarly related to overall credibility and its facets, acceptance and behavioral intentions. Taken together, this suggests that all epistemic trustworthiness facets are important for forming judgments about scientific credibility and that there is no strong dominance of one facet over the other(s).

8.3 Directions for future research

In the above, I outlined open questions, limitations, and issues regarding the conceptualization of trust in science. All of this deserves more attention in research on trust in science and science communication. In the following section, I highlight further directions for future research that are inspired by, but go beyond, the present research program, addressing very basic to more applied research questions: differentiating trust and distrust, considering trust in and within science, and the use of personalized narratives in science communication.

8.3.1 Trust versus distrust

Do laypeople who report low trust in science *distrust* science and, vice versa, is laypeople's distrust towards science the same as low *trust* in science? One might assume that lower scores on trust measures equal distrust, while higher scores on the same equal trust. However, such conclusions might be invalid and, thus, the present research program (and other research) should be interpreted carefully in this regard. The assumption that trust and distrust are the opposites of one dimension has been questioned repeatedly (e.g., Lewicki et al., 1998; Marsh & Dibben, 2005; Saunders et al., 2014; Van De Walle & Six, 2014). In an influential paper, Lewicki and colleagues (1998) argue that trust and distrust should not be viewed as the ends of one continuum but as two separate, yet linked dimensions. They draw parallels to research demonstrating that positive and negative valence can coexist regarding a broad range of psychological conditions (e.g., love and hate, expectations of benefit and harm) and propose a two-dimensional perspective crossing high and low levels of trust *and* distrust in a fourfold framework (Lewicki et al., 1998). For example, a combination of both high trust and high distrust could mean that the trust/distrust-relationship is characterized by confidence and hope as well as wariness and vigilance.

Considering such a view on trust *and* distrust in science raises many future research questions (and questions past conclusions). For example, learning about researchers' disclosure of past mistakes might spark laypeople's distrust, while not decreasing (or even increasing) their trust: Laypeople might be more skeptical of science and, thus, evaluate scientific findings more vigilantly – but they might still be willing to trust evidence that seems solid after such scrutiny. This actually fits very well with the notion that laypeople who report high trust in science do not

necessarily "blindly" trust it (e.g., Hendriks et al., 2016; Origgi, 2004; Sperber et al., 2010). Thus, trust and distrust in science might coexist. Moreover, the two dimensions might lead to separate and differing reception processes (Mayo, 2015). For example, I speculate that distrust is more relevant for sparking epistemic vigilance, while trust is more relevant for credibility perceptions, as exemplified above. There is much we do not know about the distinction of trust versus distrust in general and, even less so, regarding (dis)trust in science. Theoretically and practically, it seems very relevant to devote much more attention to this in future research.

8.3.2 Trust in and trust within science

This dissertation focuses on laypeople's trust in science. Their trust in science is not only relevant for themselves (e.g., for basing decisions on reliable information), but it is also crucial for scientists. Researchers rely on the public for support of their research activities (e.g., public funding, freedom of research) and feel motivated or even obligated to "give away" their science which requires a trusting public that is willing to receive this knowledge (i.e., the *third mission* of science; Contera, 2021; Lewis, 2021). However, trust in science does not necessarily only refer to the trust relationship between science and the public. Trust in science is also essential among scientists (i.e., trust within science; cf., Hendriks et al., 2016).

In recent years, there has been a notable shift of researchers' focus turning inwards to the processes, norms, and structures making up the system of "science" (e.g., the "Open Science Movement" or "Credibility Revolution;" e.g., Munafò et al., 2017; Nosek et al., 2015; Vazire, 2018). In this context, trust among scientists is of central relevance. For example, recently, we outlined how science

entails many social dilemmas that can only be solved by researchers making themselves vulnerable to other researchers (i.e., by trusting each other; Altenmüller & Gollwitzer, 2022). Looking farther back, the concept of cognitive labor division, and the need for trust it entails, has not only been considered for the relationship of science and the public but has been prominently applied to the functioning of science's social structure (i.e., researchers' collaboration; e.g., Kitcher, 1990; Muldoon, 2013): Cumulative science can only function effectively when researchers trust in each other's expertise. Thus, further investigating what fuels or stifles researchers' trust in other researchers will be an important research focus, inter alia, accompanying the ongoing reform movement in (psychological) science. Noteworthily, in this context, recent research has taken a special interest in the effects of self-correction in science, extending findings from Manuscript 2 in this dissertation (e.g., McDiarmid et al., 2021; Rohrer et al., 2021; Vazire & Holcombe, 2021).

Moreover, the interplay between science stereotypes and trust in science or, broader speaking, science reception is also important for the (future) social composition of science. Research on science stereotypes is often related to observations of lacking diversity in science; for example, in the context of science's gender-gap (i.e., overrepresentation of men in academia): Girls might be motivated to aspire to STEM careers by reducing stereotypes about researchers (e.g., Ferguson & Lezotte, 2020; Losh, 2010; Miller et al., 2018; Wyer et al., 2010). There are likely parallels to research on expectations of cultural fit between the self and an institution (e.g., perceived mismatch obstructing pursuit of higher education; Kristof, 1996; Kristof-Brown et al., 2005; Manstead, 2018; Stephens et al., 2019): Particularly those who perceive themselves as fitting to the expected

culture of academia (e.g., by identifying as someone who prioritizes competitiveness and competence over prosociality and warmth) might consider a career in science. This would effectively perpetuate and consolidate the reception of researchers as highly competent, yet only moderately warm (Fiske & Dupree, 2014; Imhoff, Koch, et al., 2018; Rosman & Merk, 2021) – within science as well as in the eyes of the public. This might, then, not only obstruct public trust in science but also trust among researchers (e.g., via researchers' self-stereotyping; Altenmüller & Gollwitzer, 2022). Communicating science in a warm and inclusive fashion could not only strengthen the public's trust in science but also trust among researchers and, as one downstream consequence, spark interest in science careers among a more diverse audience.

8.3.3 (Personalized) narratives in science communication

Lastly, this research program's perspective on researchers' self-disclosure could be extended to the investigation of narratives in science communication. While self-disclosure only refers to the revelation of some personal details, narratives are characterized by a cohesive storyline detailing character(s) (e.g., a personal perspective) but also causality and temporality of events (Dahlstrom, 2014). Engaging in self-disclosure can mean adding personal or even anecdotal elements to science communication and, thus, showing similarities to narrative communication. Instead of only presenting scientific evidence in a seemingly objective way by focusing on purely scientific arguments, self-disclosure might even facilitate the use of narrative elements: It could function as a "catalyst" for narrative communication by providing personal examples of how the communicated science "looks like" in everyday life. For example, a researcher might tell a personal story of when they encountered a phenomenon they now

study. Dahlstrom (2014) reviewed research on using narratives and storytelling to communicate science. He points out that narrative elements are easier to understand, more engaging, and more persuasive due to different processing of narrative compared to "logical-scientific" communication. Further, he draws parallels to news media's successful use of "personification" when reporting about societal issues (Dahlstrom, 2014), that is, focusing on a protagonist's story to exemplify a problem. Similarly, researchers' self-disclosure adds personal elements to the communication about *scientific* issues. Therefore, it might be interesting to extend the scope of the presented research program from limited self-disclosure elements to more cohesive narratives in science communication.

In health communication, which shares many commonalities with science communication, the use of narratives has been a focus of research in recent years (e.g., de Graaf et al., 2016; De Wit et al., 2008; Zebregs et al., 2015). For example, it seems effective to display desirable health behavior embedded in an emotional narrative from a first-person perspective (de Graaf et al., 2016). Interestingly, in a meta-analysis, Zebregs and colleagues (2015) demonstrate differential effects of evidence type (i.e., statistical vs. narrative) on beliefs and attitudes (comparable to the cognitive facets of credibility) and intentions (comparable to the behavioral facet of credibility): While statistical evidence (i.e., generalizable, quantitative information based on large numbers of observations) was relatively stronger related to beliefs and attitudes, narrative evidence (i.e., information based on a cohesive story of one or more protagonists) was comparably stronger related to behavioral intentions. The authors explain these effects by linking statistical evidence to cognitive responses and narrative elements to affective responses (Zebregs et al., 2015). In line with the theoretical rationale underlying the present

research program, this suggests an interesting application potential of narratives in science communication for countering laypeople's stereotypical views of scientists. The findings in this dissertation, however, do not support similar differential relations of a stronger link between affective responses (i.e., affective trustworthiness) and behavioral intentions as well as cognitive responses (i.e., expertise) and acceptance compared to vice versa combinations, as outlined above.

Taken together, self-disclosure in science communication might only be one step towards more effective science communication. Adding extended, coherent, and personalized narratives to scientific arguments could improve science communication even further. It might increase understanding, engagement, and persuasiveness of scientific findings, while also strengthening affective pathways of reception. However, there is uncertainty on what classifies as "narrative," what particular characteristics make narratives effective and why, and whether the use of narratives in science communication is actually received positively among a broad audience (Dahlstrom, 2014). Following insights of the present research program, negative effects demonstrated for self-disclosing science communication (e.g., decreases in researchers' perceived expertise) could also apply to the reception of narrative communication. Thus, it is unknown whether the positive effects of cohesive, personalized narratives might outweigh their potentially negative effects. In particular, there is hardly any research on how researchers' use of narratives might affect laypeople's trust in science.

9 Conclusion

In her recent book, historian of science Naomi Oreskes (2021) asked "Why trust science?" and made the case that not a specific "scientific" method but rather the social characteristics of the scientific system (e.g., peer review, debate, consensus) is what make science trustworthy. Likewise, this dissertation views science as a social endeavor but focuses not on interactions within science but on researchers' interactions with the public. I drew on theory about stereotypes, trust in science, and self-disclosure and applied it to the context of science communication. In three manuscripts, I investigated how laypeople perceive researchers' disclosure of basic personal information, self-critical thoughts about their own work, and personal involvement with their research topic. The findings across eleven studies paint an ambivalent picture of the effects of researchers' self-disclosure in science communication on public trust in science. Basic self-disclosure of illustrative personal information led to positive as well as negative effects but is likely of little practical relevance. Researchers expressing scientific self-criticism and intending to reform their work routines were received positively among the public. Finally, laypeople instrumentalized researchers' disclosure of engaging in "me-search" for adapting their science reception in a motivated manner to fit to their prior attitudes.

This dissertation contributes to research on self-disclosure and trust in science. Regarding self-disclosure, the present findings set boundaries on the wealth of literature suggesting positive effects on person perception and, particularly, trustworthiness. Regarding trust in science, the findings demonstrate the value of differentiating trust in science into researchers' trustworthiness and the credibility of scientific findings. Further, they inform theoretical considerations

regarding the underlying structures of these concepts and their relation to each other. Moreover, this research allows for interesting insights into how individualized trust relates to generalized trust in science and it adds to literature on motivated science reception.

Practically, this research program shows that self-disclosure on its own has likely little value for building public trust in science via science communication. Rather, it is important what exactly researchers reveal. First, researchers should consider carefully what they want to disclose to which audience and what possible costs might be. Then, there could be (small) benefits associated with disclosing certain information that makes them appear more approachable, benevolent, and integer; but, this information should also underline that they are "good" scientists, for example, by expressing a self-corrective attitude towards science and by highlighting their scientific integrity and other-oriented stance despite being personally affected by their own research topic.

On a final note, this dissertation raises at least as many questions as it answers. Thereby, it can be taken as an invitation to humbly reflect on our knowledge about theory and practice of self-disclosure, science communication, and trust in science. It should serve the purpose of sparking motivation to expand research on these issues to improve our scientific and practical knowledge about effective and trustful science communication. Fruitful directions for future research might point towards intensifying basic research on trust (and distrust), extending research on trust in science to trust *within* science, and studying the application of narratives in science communication.

We might be engrossed by a worldwide pandemic right now, but this will very likely not be the last or the biggest challenge the future holds for us. Science,

politics, and society depend on trusting each other. As psychologists, we can do a worthwhile part by studying, understanding, and explaining the basic and applied psychology of trust in science.

10 References[18]

Abele-Brehm, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314. https://doi.org/10.1037/rev0000262

Ajzen, I. (1985). From Intentions to Actions: A Theory of Planned Behavior. In J. Kuhl & J. Beckmann (Eds.), *Action Control: From Cognition to Behavior* (pp. 11–39). Springer. https://doi.org/10.1007/978-3-642-69746-3_2

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Allea. (2018). *Loss of Trust? Loss of Trustworthiness? Truth and Expertise Today*. ALLEA - All European Academies.

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Allum, N. (2011). What Makes Some People Think Astrology Is Scientific? *Science Communication*, *33*(3), 341–366. https://doi.org/10.1177/1075547010389819

Allum, N., Sturgis, P., Tabourazi, D., & Brunton-Smith, I. (2008). Science knowledge and attitudes across cultures: A meta-analysis. *Public Understanding of Science*, *17*(1), 35–54. https://doi.org/10.1177/0963662506070159

Altenmüller, M. S., & Gollwitzer, M. (2022). Prosociality in science. *Current Opinion in Psychology*, *43*, 284–288. https://doi.org/10.1016/j.copsyc.2021.08.011

Altenmüller, M. S., Lange, L. L., & Gollwitzer, M. (2021). When research is me-search: How researchers' motivation to pursue a topic affects laypeople's

[18] Here, references for chapters "Introduction" and "Final Discussion" are displayed. The manuscript chapters include their own reference lists.

trust in science. *PLOS ONE*, *16*(7), e0253911. https://doi.org/10.1371/journal.pone.0253911

Altenmüller, M. S., Nuding, S., & Gollwitzer, M. (2021). No harm in being self-corrective: Self-criticism and reform intentions increase researchers' epistemic trustworthiness and credibility in the eyes of the public. *Public Understanding of Science*, *30*(8), 09636625211022181. https://doi.org/10.1177/09636625211022181

Altman, I., & Taylor, D. A. (1973). *Social penetration: The development of interpersonal relationships*. Holt, Rinehart & Winston.

Anvari, F., & Lakens, D. (2018). The replicability crisis and public trust in psychological science. *Comprehensive Results in Social Psychology*, *3*(3), 266–286. https://doi.org/10.1080/23743603.2019.1684822

Appelman, A., & Sundar, S. S. (2016). Measuring Message Credibility: Construction and Validation of an Exclusive Scale. *Journalism & Mass Communication Quarterly*, *93*(1), 59–79. https://doi.org/10.1177/1077699015606057

Aron, A., Melinat, E., Aron, E. N., Vallone, R. D., & Bator, R. J. (1997). The Experimental Generation of Interpersonal Closeness: A Procedure and Some Preliminary Findings. *Personality and Social Psychology Bulletin*, *23*(4), 363–377. https://doi.org/10.1177/0146167297234003

Bauer, M. W., Allum, N., & Miller, S. (2007). What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public Understanding of Science*, *16*(1), 79–95. https://doi.org/10.1177/0963662506071287

Bazarova, N. N. (2012). Public Intimacy: Disclosure Interpretation and Social Judgments on Facebook: Public Intimacy. *Journal of Communication*, *62*(5), 815–832. https://doi.org/10.1111/j.1460-2466.2012.01664.x

Bender, J., Rothmund, T., Nauroth, P., & Gollwitzer, M. (2016). How Moral Threat Shapes Laypersons' Engagement With Science. *Personality and Social*

Psychology Bulletin, *42*(12), 1723–1735. https://doi.org/10.1177/0146167216671518

Breckler, S. J. (1984). Empirical validation of affect, behavior, and cognition as distinct components of attitude. *Journal of Personality and Social Psychology*, *47*(6), 1191–1205. https://doi.org/10.1037/0022-3514.47.6.1191

Brewer, M. B., & Mittelman, J. (1980). Effects of normative control of self-disclosure on reciprocity. *Journal of Personality*, *48*(1), 89–102. https://doi.org/10.1111/j.1467-6494.1980.tb00968.x

Bromme, R., & Goldman, S. R. (2014). The Public's Bounded Understanding of Science. *Educational Psychologist*, *49*(2), 59–69. https://doi.org/10.1080/00461520.2014.921572

Bromme, R., Mede, N. G., Thomm, E., Kremer, B., & Ziegler, R. (2022). An anchor in troubled times: Trust in science before and within the COVID-19 pandemic. *PLOS ONE*, *17*(2), e0262823. https://doi.org/10.1371/journal.pone.0262823

Bromme, R., & Thomm, E. (2016). Knowing Who Knows: Laypersons' Capabilities to Judge Experts' Pertinence for Science Topics. *Cognitive Science*, *40*(1), 241–252. https://doi.org/10.1111/cogs.12252

Bromme, R., Thomm, E., & Wolf, V. (2015). From Understanding to Deference: Laypersons' and Medical Students' Views on Conflicts Within Medicine. *International Journal of Science Education, Part B*, *5*(1), 68–91. https://doi.org/10.1080/21548455.2013.849017

Buehl, M. M., Alexander, P. A., & Murphy, P. K. (2002). Beliefs about Schooled Knowledge: Domain Specific or Domain General? *Contemporary Educational Psychology*, *27*(3), 415–449. https://doi.org/10.1006/ceps.2001.1103

Caltabiano, M. L., & Smithson, M. (2010). Variables Affecting the Perception of Self-Disclosure Appropriateness. *The Journal of Social Psychology*. https://www.tandfonline.com/doi/abs/10.1080/00224545.1983.9712017

Cao, J., & Banaji, M. R. (2016). The base rate principle and the fairness principle in social judgment. *Proceedings of the National Academy of Sciences*, *113*(27), 7475–7480. https://doi.org/10.1073/pnas.1524268113

Carpenter, A., & Greene, K. (2016). Social Penetration Theory. In C. R. Berger & M. E. Roloff (Eds.), *The International Encyclopedia of Interpersonal Communication* (Vol. 1). John Wiley & Sons. https://doi.org/10.1002/9781118540190.wbeic160

Cayanus, J., & Martin, M. (2008). Teacher Self-Disclosure: Amount, Relevance, and Negativity. *Communication Quarterly*, *56*, 325–341. https://doi.org/10.1080/01463370802241492

Chaikin, A. L., & Derlega, V. J. (1974). Variables affecting the appropriateness of self-disclosure. *Journal of Consulting and Clinical Psychology*, *42*(4), 588–593. https://doi.org/10.1037/h0036614

Collins, N. L., & Miller, L. C. (1994). Self-disclosure and liking: A meta-analytic review. *Psychological Bulletin*, *116*(3), 457–475. https://doi.org/10.1037/0033-2909.116.3.457

Contera, S. (2021). Communication is central to the mission of science. *Nature Reviews Materials*, 1–2. https://doi.org/10.1038/s41578-021-00316-w

Cozby, P. C. (1972). Self-Disclosure, Reciprocity and Liking. *Sociometry*, *35*(1), 151–160. https://doi.org/10.2307/2786555

Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin*, *79*(2), 73–91. https://doi.org/10.1037/h0033950

Dahlstrom, M. F. (2014). Using narratives and storytelling to communicate science with nonexpert audiences. *Proceedings of the National Academy of Sciences*, *111*(Supplement_4), 13614–13620. https://doi.org/10.1073/pnas.1320645111

de Graaf, A., Sanders, J., & Hoeken, H. (2016). Characteristics of Narrative Interventions and Health Effects: A Review of the Content, Form, and Context of Narratives in Health-related Narrative Persuasion Research.

Review of Communication Research, *4*, 88–131. https://doi.org/10.12840/issn.2255-4165.2016.04.01.011

De Wit, J. B. F., Das, E., & Vet, R. (2008). What works best: Objective statistics or a personal testimonial? An assessment of the persuasive effects of different types of message evidence on risk perception. *Health Psychology*, *27*(1), 110–115. https://doi.org/10.1037/0278-6133.27.1.110

Derlega, V. J., Lovell, R., & Chaikin, A. L. (1976). Effects of therapist disclosure and its perceived appropriateness on client self-disclosure. *Journal of Consulting and Clinical Psychology*, *44*(5), 866–866. https://doi.org/10.1037/0022-006X.44.5.866

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18. https://doi.org/10.1037/0022-3514.56.1.5

Dindia, K. (2002). Self-disclosure research: Knowledge through meta-analysis. In M. Allen, R. W. Preiss, B. M. Gayle, & N. Burrell (Eds.), *Interpersonal Communication Research: Advances Through Meta-analysis* (pp. 169–185). Routledge.

Dindia, K., & Allen, M. (1992). Sex differences in self-disclosure: A meta-analysis. *Psychological Bulletin*, *112*(1), 106–124. https://doi.org/10.1037/0033-2909.112.1.106

Dohle, S., Wingen, T., & Schreiber, M. (2020). Acceptance and Adoption of Protective Measures During the COVID-19 Pandemic: The Role of Trust in Politics and Trust in Science. *Social Psychological Bulletin*, *15*(4), 1–23. https://doi.org/10.32872/spb.4315

Drummond, C., & Fischhoff, B. (2017). Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. *Proceedings of the National Academy of Sciences*, *114*(36), 9587–9592. https://doi.org/10.1073/pnas.1704882114

Earle, T. C., & Siegrist, M. (2006). Morality Information, Performance Information, and the Distinction Between Trust and Confidence. *Journal of Applied Social*

Psychology, *36*(2), 383–416. https://doi.org/10.1111/j.0021-9029.2006.00012.x

Ensari, N., & Miller, N. (2002). The out-group must not be so bad after all: The effects of disclosure, typicality, and salience on intergroup bias. *Journal of Personality and Social Psychology*, *83*(2), 313–329. https://doi.org/10.1037/0022-3514.83.2.313

Evans, N. D., & Fetterman, A. K. (2021). It Doesn't Apply to Me, So It Isn't Real: People Are Likely to Deny Science if It Contradicts Their Personality. *Social Psychological and Personality Science*, 19485506211051464. https://doi.org/10.1177/19485506211051464

Ferguson, S. L., & Lezotte, S. M. (2020). Exploring the state of science stereotypes: Systematic review and meta-analysis of the Draw-A-Scientist Checklist. *School Science and Mathematics*, *120*(1), 55–65. https://doi.org/10.1111/ssm.12382

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Dupree, C. (2014). Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences*, *111*(Supplement_4), 13593–13597. https://doi.org/10.1073/pnas.1317505111

Freitag, M., & Traunmüller, R. (2009). Spheres of trust: An empirical analysis of the foundations of particularised and generalised trust. *European Journal of Political Research*, *48*(6), 782–803. https://doi.org/10.1111/j.1475-6765.2009.00849.x

Funk, C. (2017). Mixed Messages about Public Trust in Science. *Issues in Science and Technology*, *34*(1), 86–88.

Gelso, C. J., & Palma, B. (2011). Directions for research on self-disclosure and immediacy: Moderation, mediation, and the inverted U. *Psychotherapy*, *48*(4), 342–348. https://doi.org/10.1037/a0025909

Ghai, S. (2021). It's time to reimagine sample diversity and retire the WEIRD dichotomy. *Nature Human Behaviour*, *5*(8), 971–972. https://doi.org/10.1038/s41562-021-01175-9

Gollwitzer, M., & Schwabe, J. (2021). Context Dependency as a Predictor of Replicability. *Review of General Psychology*, 10892680211015636. https://doi.org/10.1177/10892680211015635

Greene, K., Derlega, V. J., & Mathews, A. (2006). Self-Disclosure in Personal Relationships. In A. L. Vangelisti & D. Perlman (Eds.), *The Cambridge handbook of personal relationships* (pp. 409–427). Cambridge University Press. https://doi.org/10.1017/CBO9780511606632.023

Greitemeyer, T. (2014). I Am Right, You Are Wrong: How Biased Assimilation Increases the Perceived Gap between Believers and Skeptics of Violent Video Game Effects. *PLOS ONE*, *9*(4), e93440. https://doi.org/10.1371/journal.pone.0093440

Hardwig, J. (1991). The Role of Trust in Knowledge. *The Journal of Philosophy*, *88*(12), 693–708. https://doi.org/10.2307/2027007

Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOS ONE*, *10*(10), e0139309. https://doi.org/10.1371/journal.pone.0139309

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Trust in Science and the Science of Trust. In B. Blöbaum (Ed.), *Trust and Communication in a Digitized World* (pp. 143–159). Springer International Publishing. https://doi.org/10.1007/978-3-319-28059-2_8

Hendriks, F., Kienhues, D., & Bromme, R. (2020). Replication crisis = trust crisis? The effect of successful vs failed replications on laypeople's trust in researchers and research. *Public Understanding of Science*, *29*(3), 270–288. https://doi.org/10.1177/0963662520902383

Henretty, J. R., Currier, J. M., Berman, J. S., & Levitt, H. M. (2014). The impact of counselor self-disclosure on clients: A meta-analytic review of experimental and quasi-experimental research. *Journal of Counseling Psychology*, *61*(2), 191–207. https://doi.org/10.1037/a0036189

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hornsey, M. J. (2020). Why Facts Are Not Enough: Understanding and Managing the Motivated Rejection of Science. *Current Directions in Psychological Science*, 0963721420969364. https://doi.org/10.1177/0963721420969364

Huang, L.-S. (2015). Trust in product review blogs: The influence of self-disclosure and popularity. *Behaviour & Information Technology*, *34*(1), 33–44. https://doi.org/10.1080/0144929X.2014.978378

Hutmacher, F., Reichardt, R., & Appel, M. (2022). The role of motivated science reception and numeracy in the context of the COVID-19 pandemic. *Public Understanding of Science*, *31*(1), 19–34. https://doi.org/10.1177/09636625211047974

Imhoff, R., Koch, A., & Flade, F. (2018). (Pre)occupations: A data-driven model of jobs and its consequences for categorization and evaluation. *Journal of Experimental Social Psychology*, *77*, 76–88. https://doi.org/10.1016/j.jesp.2018.04.001

Imhoff, R., Lamberty, P., & Klein, O. (2018). Using Power as a Negative Cue: How Conspiracy Mentality Affects Epistemic Trust in Sources of Historical Knowledge. *Personality and Social Psychology Bulletin*, *44*(9), 1364–1379. https://doi.org/10.1177/0146167218768779

Jarreau, P. B., Cancellare, I. A., Carmichael, B. J., Porter, L., Toker, D., & Yammine, S. Z. (2019). Using selfies to challenge public stereotypes of scientists. *PLOS ONE*, *14*(5), e0216625. https://doi.org/10.1371/journal.pone.0216625

Jensen, E. A., Kennedy, E. B., & Greenwood, E. (2021). Pandemic: Public feeling more positive about science. *Nature*, *591*(7848), 34–34. https://doi.org/10.1038/d41586-021-00542-w

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. https://doi.org/10.1037/0022-3514.89.6.899

Kadji, K., & Schmid Mast, M. (2021). The effect of physician self-disclosure on patient self-disclosure and patient perceptions of the physician. *Patient Education and Counseling*, *104*(9), 2224–2231. https://doi.org/10.1016/j.pec.2021.02.030

Keil, F. C. (2005). Doubt, deference, and deliberation: Understanding and using the division of cognitive labour. In T. Gendler & J. P. Hawthorne (Eds.), *Oxford Studies in Epistemology* (Vol. 1, pp. 143–166). Clarendon Press.

Keil, F. C., Stein, C., Webb, L., Billings, V. D., & Rozenblit, L. (2008). Discerning the Division of Cognitive Labor: An Emerging Understanding of How Knowledge Is Clustered in Other Minds. *Cognitive Science*, *32*(2), 259–300. https://doi.org/10.1080/03640210701863339

Kim, T.-Y., Lee, D.-R., & Wong, N. (2016). Supervisor Humor and Employee Outcomes: The Role of Social Distance and Affective Trust in Supervisor. *Journal of Business & Psychology*, *31*(1), 125–139. https://doi.org/10.1007/s10869-015-9406-9

Kitcher, P. (1990). The Division of Cognitive Labor. *The Journal of Philosophy*, *87*(1), 5–22. https://doi.org/10.2307/2026796

Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–

progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675–709. https://doi.org/10.1037/pspa0000046

Koch, A., Imhoff, R., Unkelbach, C., Nicolas, G., Fiske, S., Terache, J., Carrier, A., & Yzerbyt, V. (2020). Groups' warmth is a personal matter: Understanding consensus on stereotype dimensions reconciles adversarial models of social evaluation. *Journal of Experimental Social Psychology*, *89*, 103995. https://doi.org/10.1016/j.jesp.2020.103995

Koch, A., Yzerbyt, V., Abele-Brehm, A., Ellemers, N., & Fiske, S. T. (2021). Chapter One - Social evaluation: Comparing models across interpersonal, intragroup, intergroup, several-group, and many-group contexts. *Advances in Experimental Social Psychology*, *63*, 1–68. https://doi.org/10.1016/bs.aesp.2020.11.001

Kotzur, P. F., Schäfer, S. J., & Wagner, U. (2019). Meeting a nice asylum seeker: Intergroup contact changes stereotype content perceptions and associated emotional prejudices, and encourages solidarity-based collective action intentions. *British Journal of Social Psychology*, *58*(3), 668–690. https://doi.org/10.1111/bjso.12304

Krause, N. M., Brossard, D., Scheufele, D. A., Xenos, M. A., & Franke, K. (2019). Trends—Americans' Trust in Science and Scientists. *Public Opinion Quarterly*, *83*(4), 817–836. https://doi.org/10.1093/poq/nfz041

Kristof, A. L. (1996). Person-Organization Fit: An Integrative Review of Its Conceptualizations, Measurement, and Implications. *Personnel Psychology*, *49*(1), 1–49. https://doi.org/10.1111/j.1744-6570.1996.tb01790.x

Kristof-Brown, A. L., Zimmerman, R. D., & Johnson, E. C. (2005). Consequences of Individuals' Fit at Work: A Meta-Analysis of Person–Job, Person–Organization, Person–Group, and Person–Supervisor Fit. *Personnel Psychology*, *58*(2), 281–342. https://doi.org/10.1111/j.1744-6570.2005.00672.x

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480–498. https://doi.org/10.1037/0033-2909.108.3.480

Lewandowsky, S., & Oberauer, K. (2016). Motivated Rejection of Science. *Current Directions in Psychological Science*, *25*(4), 217–222. https://doi.org/10.1177/0963721416654436

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust And Distrust: New Relationships and Realities. *Academy of Management Review*, *23*(3), 438–458. https://doi.org/10.5465/amr.1998.926620

Lewis, N. A. (2021). What counts as good science? How the battle for methodological legitimacy affects public psychology. *American Psychologist*, *76*(8), 1323–1333. https://doi.org/10.1037/amp0000870

Losh, S. C. (2010). Stereotypes about scientists over time among US adults: 1983 and 2001. *Public Understanding of Science*, *19*(3), 372–382. https://doi.org/10.1177/0963662508098576

Manstead, A. S. R. (2018). The psychology of social class: How socioeconomic status impacts thought, feelings, and behaviour. *British Journal of Social Psychology*, *57*(2), 267–291. https://doi.org/10.1111/bjso.12251

Marsh, S., & Dibben, M. R. (2005). The role of trust in information science and technology. *Annual Review of Information Science and Technology*, *37*(1), 465–498. https://doi.org/10.1002/aris.1440370111

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709. https://doi.org/10.2307/258792

Mayo, R. (2015). Cognition is a matter of trust: Distrust tunes cognitive processes. *European Review of Social Psychology*, *26*(1), 283–327. https://doi.org/10.1080/10463283.2015.1117249

Mazer, J. P., Murphy, R. E., & Simonds, C. J. (2009). The effects of teacher self-disclosure via Facebook on teacher credibility. *Learning, Media and Technology*, *34*(2), 175–183. https://doi.org/10.1080/17439880902923655

McAllister, D. J. (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. *Academy of Management Journal*, *38*(1), 24–59. https://doi.org/10.5465/256727

McDiarmid, A. D., Tullett, A. M., Whitt, C. M., Vazire, S., Smaldino, P. E., & Stephens, J. E. (2021). Psychologists update their beliefs about effect sizes after replication studies. *Nature Human Behaviour*, *5*(12), 1663–1673. https://doi.org/10.1038/s41562-021-01220-7

Merk, S., Rosman, T., Muis, K. R., Kelava, A., & Bohl, T. (2018). Topic specific epistemic beliefs: Extending the Theory of Integrated Domains in Personal Epistemology. *Learning and Instruction*, *56*, 84–97. https://doi.org/10.1016/j.learninstruc.2018.04.008

Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & Mccann, R. M. (2003). Credibility for the 21st Century: Integrating Perspectives on Source, Message, and Media Credibility in the Contemporary Media Environment. *Annals of the International Communication Association*, *27*(1), 293–335. https://doi.org/10.1080/23808985.2003.11679029

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies. *Child Development*, *89*(6), 1943–1955. https://doi.org/10.1111/cdev.13039

Morrow, J. L., Hansen, M. H., & Pearson, A. W. (2004). The Cognitive and Affective Antecedents of General Trust Within Cooperative Organizations. *Journal of Managerial Issues*, *16*(1), 48–64.

Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-Generality and Domain-Specificity in Personal Epistemology Research: Philosophical and Empirical Reflections in the Development of a Theoretical Framework. *Educational Psychology Review*, *18*(1), 3–54. https://doi.org/10.1007/s10648-006-9003-6

Muldoon, R. (2013). Diversity and the Division of Cognitive Labor. *Philosophy Compass*, *8*(2), 117–125. https://doi.org/10.1111/phc3.12000

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Nauroth, P., Gollwitzer, M., Bender, J., & Rothmund, T. (2015). Social Identity Threat Motivates Science-Discrediting Online Comments. *PLOS ONE*, *10*(2), e0117476. https://doi.org/10.1371/journal.pone.0117476

Nauroth, P., Gollwitzer, M., Kozuchowski, H., Bender, J., & Rothmund, T. (2017). The effects of social identity threat and social identity affirmation on laypersons' perception of scientists. *Public Understanding of Science*, *26*(7), 754–770. https://doi.org/10.1177/0963662516631289

Nazione, S., Perrault, E. K., & Keating, D. M. (2019). Finding Common Ground: Can Provider-Patient Race Concordance and Self-disclosure Bolster Patient Trust, Perceptions, and Intentions? *Journal of Racial and Ethnic Health Disparities*, *6*(5), 962–972. https://doi.org/10.1007/s40615-019-00597-6

Nifadkar, S. S., Wu, W., & Gu, Q. (2019). Supervisors' work-related and nonwork information sharing: Integrating research on information sharing, information seeking, and trust using self-disclosure theory. *Personnel Psychology*, *72*(2), 241–269. https://doi.org/10.1111/peps.12305

Nisbet, E. C., Cooper, K. E., & Garrett, R. K. (2015). The Partisan Brain: How Dissonant Science Messages Lead Conservatives and Liberals to (Dis)Trust Science. *The ANNALS of the American Academy of Political and Social Science*, *658*(1), 36–66. https://doi.org/10.1177/0002716214555474

Norton, M. I., Frost, J. H., & Ariely, D. (2007). Less is more: The lure of ambiguity, or why familiarity breeds contempt. *Journal of Personality and Social Psychology*, *92*(1), 97–105. https://doi.org/10.1037/0022-3514.92.1.97

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open

research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Omarzu, J. (2000). A disclosure decision model: Determining how and when individuals will self-disclose. *Personality and Social Psychology Review*, *4*(2), 174–185. https://doi.org/10.1207/S15327957PSPR0402_05

Oreskes, N. (2021). Why Trust Science? In *Why Trust Science?* Princeton University Press. https://doi.org/10.1515/9780691222370

Oreskes, N., & Conway, E. M. (2010). *Merchants of doubt. How a handful of scientists obscured the truth on issues from tobacco smoke to global warming*. Bloomsbury Publishing USA.

Origgi, G. (2004). Is Trust an Epistemological Notion? *Episteme*, *1*(1), 61–72. https://doi.org/10.3366/epi.2004.1.1.61

Origgi, G. (2012). Epistemic Injustice and Epistemic Trust. *Social Epistemology*, *26*(2), 221–235. https://doi.org/10.1080/02691728.2011.652213

Pagliaro, S., Sacchi, S., Pacilli, M. G., Brambilla, M., Lionetti, F., Bettache, K., Bianchi, M., Biella, M., Bonnot, V., Boza, M., Butera, F., Ceylan-Batur, S., Chong, K., Chopova, T., Crimston, C. R., Álvarez, B., Cuadrado, I., Ellemers, N., Formanowicz, M., … Zubieta, E. (2021). Trust predicts COVID-19 prescribed and discretionary behavioral intentions in 23 countries. *PLOS ONE*, *16*(3), e0248334. https://doi.org/10.1371/journal.pone.0248334

Pettigrew, T. F. (1998). Intergroup Contact Theory. *Annual Review of Psychology*, *49*(1), 65–85. https://doi.org/10.1146/annurev.psych.49.1.65

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922–934. https://doi.org/10.1002/ejsp.504

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205. https://doi.org/10.1016/S0065-2601(08)60214-2

Pew Research Center. (2020). *Science and Scientists Held in High Esteem Across Global Publics* (p. 133).

Pew Research Center. (2022). *In the age of COVID-19, do Americans trust science?* https://www.youtube.com/watch?v=H9L37cPVSqo

Plohl, N., & Musil, B. (2021). Modeling compliance with COVID-19 prevention guidelines: The critical role of trust in science. *Psychology, Health & Medicine*, *26*(1), 1–12. https://doi.org/10.1080/13548506.2020.1772988

Pornpitakpan, C. (2004). The Persuasiveness of Source Credibility: A Critical Review of Five Decades' Evidence. *Journal of Applied Social Psychology*, *34*(2), 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x

Rains, S. A., & Brunner, S. R. (2018). The Outcomes of Broadcasting Self-Disclosure Using New Communication Technologies: Responses to Disclosure Vary Across One's Social Network. *Communication Research*, *45*(5), 659–687. https://doi.org/10.1177/0093650215598836

Resnik, D. B. (2011). Scientific Research and the Public Trust. *Science and Engineering Ethics*, *17*(3), 399–409. https://doi.org/10.1007/s11948-010-9210-x

Richard, F. D., Bond, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology*, *7*(4), 331–363. https://doi.org/10.1037/1089-2680.7.4.331

Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, *41*(1), 307–364. https://doi.org/10.1002/aris.2007.1440410114

Rohrer, J. M., Tierney, W., Uhlmann, E. L., DeBruine, L. M., Heyman, T., Jones, B., Schmukle, S. C., Silberzahn, R., Willén, R. M., Carlsson, R., Lucas, R. E., Strand, J., Vazire, S., Witt, J. K., Zentall, T. R., Chabris, C. F., & Yarkoni, T.

(2021). Putting the Self in Self-Correction: Findings From the Loss-of-Confidence Project. *Perspectives on Psychological Science*, 1745691620964106. https://doi.org/10.1177/1745691620964106

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). Yale University Press.

Rosman, T., & Merk, S. (2021). Teacher's Reasons for Trust and Distrust in Scientific Evidence: Reflecting a "Smart But Evil" Stereotype? *AERA Open*, *7*, 233285842110285. https://doi.org/10.1177/23328584211028599

Rothmund, T., Gollwitzer, M., Nauroth, P., & Bender, J. (2017). Motivierte Wissenschaftsrezeption. *Psychologische Rundschau*, *68*(3), 193–197. https://doi.org/10.1026/0033-3042/a000364

Rubinstein, R. S., Jussim, L., & Stevens, S. T. (2018). Reliance on individuating information and stereotypes in implicit and explicit person perception. *Journal of Experimental Social Psychology*, *75*, 54–70. https://doi.org/10.1016/j.jesp.2017.11.009

Rutjens, B. T., & Heine, S. J. (2016). The Immoral Landscape? Scientists Are Associated with Violations of Morality. *PLOS ONE*, *11*(4), e0152798. https://doi.org/10.1371/journal.pone.0152798

Rutjens, B. T., Sutton, R. M., & van der Lee, R. (2018). Not All Skepticism Is Equal: Exploring the Ideological Antecedents of Science Acceptance and Rejection. *Personality and Social Psychology Bulletin*, *44*(3), 384–405. https://doi.org/10.1177/0146167217741314

Saffran, L., Hu, S., Hinnant, A., Scherer, L. D., & Nagel, S. C. (2020). Constructing and influencing perceived authenticity in science communication: Experimenting with narrative. *PLOS ONE*, *15*(1), e0226711. https://doi.org/10.1371/journal.pone.0226711

Saunders, M. N., Dietz, G., & Thornhill, A. (2014). Trust and distrust: Polar opposites, or independent but co-existing? *Human Relations*, *67*(6), 639–665. https://doi.org/10.1177/0018726713500831

Schäfer, M. S., Füchslin, T., Casutt, G.-A., Suggs, S., Aberer, K., Burkard, P., Godinho, A., Hirschi, C., Jacobs, A., Jarren, O., Kaufmann, A., Knutti, R., Maier, M., Metag, J., Müller, T., Strasser, B., & Weichselbraun, A. (2021). Science in the Swiss Public. The State of Science Communication and Public Engagement with Science in Switzerland. In *Swiss Academies Reports* (Vol. 16, Issue 8) [Report]. ETH Zurich. https://doi.org/10.5281/ZENODO.4974311

Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. *Frontiers in Psychology*, *10*, 813. https://doi.org/10.3389/fpsyg.2019.00813

Scheitle, C. P., & Guthrie, S. K. (2019). Public responses to scientific research: Does disciplinary attribution matter? *Public Understanding of Science*, *28*(2), 234–245. https://doi.org/10.1177/0963662518814365

Schlosser, A. E. (2020). Self-disclosure versus self-presentation on social media. *Current Opinion in Psychology*, *31*, 1–6. https://doi.org/10.1016/j.copsyc.2019.06.025

Schwabe, J., Altenmüller, M.S., & Gollwitzer, M. (in prep.). When and Why Informality Fosters Closeness and Trustworthiness.

Schwarz, N., & Newman, E. (2017). How does the gut know truth? The psychology of "truthiness." *APA Science Brief*.

Siegrist, M. (1999). A Causal Model Explaining the Perception and Acceptance of Gene Technology1. *Journal of Applied Social Psychology*, *29*(10), 2093–2106. https://doi.org/10.1111/j.1559-1816.1999.tb02297.x

Sperber, D., Clément, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic Vigilance. *Mind & Language*, *25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Sprecher, S., Treger, S., & Wondra, J. D. (2013). Effects of self-disclosure role on liking, closeness, and other impressions in get-acquainted interactions. *Journal of Social and Personal Relationships*, *30*(4), 497–514. https://doi.org/10.1177/0265407512459033

Sprecher, S., Treger, S., Wondra, J. D., Hilaire, N., & Wallpe, K. (2013). Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology*, *49*(5), 860–866. https://doi.org/10.1016/j.jesp.2013.03.017

Stahl, E., & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, *17*(6), 773–785. https://doi.org/10.1016/j.learninstruc.2007.09.016

Stephens, N. M., Townsend, S. S. M., & Dittmann, A. G. (2019). Social-Class Disparities in Higher Education and Professional Workplaces: The Role of Cultural Mismatch. *Current Directions in Psychological Science*, *28*(1), 67–73. https://doi.org/10.1177/0963721418806506

Tajfel, H. (1982). Social Psychology of Intergroup Relations. *Annual Review of Psychology*, *33*(1), 1–39. https://doi.org/10.1146/annurev.ps.33.020182.000245

Tajfel, H., & Turner, J. C. (1979). An integrative theory of group conflict. In M. J. Austin & M. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole.

Tal-Or, N., & Hershman-Shitrit, M. (2015). Self-Disclosure and the Liking of Participants in Reality Tv. *Human Communication Research*, *41*(2), 245–267. https://doi.org/10.1111/hcre.12047

Tardy, C. H., & Smithson, J. (2018). Self-disclosure: Strategic revelation of information in personal and professional relationships 1. In O. Hargie (Ed.), *The Handbook of Communication Skills* (4th ed., pp. 217–258). Routledge. https://doi.org/10.4324/9781315436135-8

Terwel, B. W., Harinck, F., Ellemers, N., & Daamen, D. D. L. (2009). Competence-Based and Integrity-Based Trust as Predictors of Acceptance of Carbon

Dioxide Capture and Storage (CCS). *Risk Analysis*, *29*(8), 1129–1140. https://doi.org/10.1111/j.1539-6924.2009.01256.x

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.

Van De Walle, S., & Six, F. (2014). Trust and Distrust as Distinct Concepts: Why Studying Distrust in Institutions is Important. *Journal of Comparative Policy Analysis: Research and Practice*, *16*(2), 158–174. https://doi.org/10.1080/13876988.2013.785146

Vazire, S. (2018). Implications of the Credibility Revolution for Productivity, Creativity, and Progress: *Perspectives on Psychological Science*. https://doi.org/10.1177/1745691617751884

Vazire, S., & Holcombe, A. O. (2021). Where are the Self-Correcting Mechanisms in Science? *Review of General Psychology*, 10892680211033912. https://doi.org/10.1177/10892680211033912

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the Web. *Journal of the American Society for Information Science and Technology*, *53*(2), 134–144. https://doi.org/10.1002/asi.10016

Wheeless, L. R., & Grotz, J. (1977). The measurement of trust and its relationship to self-disclosure. *Human Communication Research*, *3*(3), 250–257. https://doi.org/10.1111/j.1468-2958.1977.tb00523.x

Wilholt, T. (2013). Epistemic Trust in Science. *The British Journal for the Philosophy of Science*, *64*(2), 233–253. https://doi.org/10.1093/bjps/axs007

Wingen, T., Berkessel, J. B., & Englich, B. (2020). No Replication, No Trust? How Low Replicability Influences Trust in Psychology. *Social Psychological and Personality Science*, *11*(4), 454–463. https://doi.org/10.1177/1948550619877412

Wissenschaft im Dialog/Kantar. (2021). *Wissenschaftsbarometer 2021*.

Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology*, *38*(7), 1139–1147. https://doi.org/10.1002/ejsp.549

Worthy, M., Gary, A. L., & Kahn, G. M. (1969). Self-disclosure as an exchange process. *Journal of Personality and Social Psychology*, *13*(1), 59–63. https://doi.org/10.1037/h0027990

Wyer, M., Schneider, J., Nassar-McMillan, S., & Oliver-Hoyo, M. (2010). Capturing Stereotypes: Developing a Scale to Explore U.S. College Students' Images of Science and Scientists. *International Journal of Gender, Science and Technology*, *2*(3). http://genderandset.open.ac.uk/index.php/genderandset/article/view/78

Yzerbyt, V. Y., Kervyn, N., & Judd, C. M. (2008). Compensation Versus Halo: The Unique Relations Between the Fundamental Dimensions of Social Judgment. *Personality and Social Psychology Bulletin*, *34*(8), 1110–1123. https://doi.org/10.1177/0146167208318602

Zanna, M. P., & Rempel, J. K. (1988). Attitudes: A new look at an old concept. In D. Bar-Tal & W. Kruglanski (Eds.), *The social psychology of knowledge* (pp. 315–334). Cambridge University Press.

Zebregs, S., van den Putte, B., Neijens, P., & de Graaf, A. (2015). The differential impact of statistical and narrative evidence on beliefs, attitude, and intention: A meta-analysis. *Health Communication*, *30*(3), 282–289. https://doi.org/10.1080/10410236.2013.842528

Zhu, L. (Lei), Aquino, K., & Vadera, A. K. (2016). What makes professors appear credible: The effect of demographic characteristics and ideological beliefs. *Journal of Applied Psychology*, *101*(6), 862–880. https://doi.org/10.1037/apl0000095

11 Danksagung

Das Abschließen der Dissertation ist für mich kein Endpunkt meiner „academic journey“ – selbstverständlich aber ein wichtiger Meilenstein. Ich möchte die Gelegenheit nutzen, auf dem Weg innezuhalten und meinen Begleiter:innen DANKE zu sagen.

Zunächst möchte ich meinem Betreuer Prof. Dr. Mario Gollwitzer danken. In Marburg hat er maßgeblich dazu beigetragen, meine Begeisterung für die Psychologie zu schüren (und das mit Methoden und Statistik!). In München hat er mir dann die Gelegenheit gegeben, mein Interesse für die Sozialpsychologie weiterzuverfolgen und zu vertiefen. Mit der richtigen Mischung aus Freiheit und kritischem Nachhaken („Aber was ist die THEORIE?!“) konnte ich unter seiner Begleitung beginnen, mein eigenes Forschungsprofil zu entwickeln und zu schärfen. Ich freue mich sehr auf die weitere Zusammenarbeit! Dank gebührt auch meinem Zweitgutachter Prof. Dr. Rainer Bromme, dessen Forschung ein wichtiger Grundstein für meine eigenen Arbeit darstellt und mein Denken über Vertrauen in Wissenschaft nachhaltig geprägt hat, sowie Prof. Dr. Hans-Bernd Brosius für seine Bereitschaft, der Prüfungskommission der Verteidigung beizuwohnen.

Von Herzen danke ich meinen Kolleg:innen: Dr. Mathias Twardawski, Lucas Köhler, Dr. Zoe Magraw-Mickelson, Moritz Fischer, Moritz Fedeneder, Dr. Johannes Schwabe und Franziska Brotzeller. Alles, was in dieser Arbeit steht, ist auch ein Widerschein unserer zahlreichen Gespräche und Diskussionen. Ob bei Brown Bags, Mensa, Kaffee oder zwischen Tür und Angel – alle hatten stets ein offenes Ohr und eine gute Antwort für mich parat. Ich nehme es nicht für selbstverständlich, Teil eines solchen Teams zu sein. Ich freue mich sehr darauf, weiter an unserem vermeintlichen „work hard, play hard“ Ruf zu arbeiten!

Außerdem gilt der Dank auch Dr. Tobias Wingen. Als ich mit dieser Arbeit begonnen habe – vor der Pandemie – fühlte es sich oft so an, als wäre ich mit meinem Interesse für Wissenschaftskommunikation und Vertrauen in Wissenschaft recht alleine, in Tobi habe ich früh einen Gleichgesinnten, Ansprechpartner und jetzt auch Kollaborator gefunden.

Ich danke meinem Vater Dr. Thomas Altenmüller und meinem Partner Emanuel Thaler, die mich mit ihrem „Blick von außen" immer wieder zum Nachdenken über meine Forschung angeregt und mir wertvolles Feedback gegeben haben. Außerdem sei gesagt, dass die LMU zwar natürlich eine tolle Uni ist, ich aber besonders dankbar bin, dass mich der Weg an die LMU auch zurück zu meiner Familie geführt hat. Eine Doktorarbeit schreibt sich leichter, wenn man ihre Unterstützung an seiner Seite weiß: Danke also auch insbesondere an meine Mutter Julia sowie meinem Bruder Laurenz und seiner Frau Lori Altenmüller.

Zuletzt möchte ich meinen Studierenden und unseren Hiwis danken. Hier seien ganz besonders Stephan Nuding und Leonie Lange erwähnt, deren Abschlussarbeiten Ausgangspunkte für zwei Herzstücken dieser Dissertation waren. Und abschließend möchte ich die Gelegenheit nutzen, auch allen Personen zu danken, die an meinen vielen Studien bereitwillig teilgenommen haben. Die hier durchgeführte Forschung ist auf die Teilnahme interessierter Freiwilliger angewiesen und ich halte es für sehr wichtig, deren Beitrag hoch wertzuschätzen.

Der Weg der Wissenschaft ist nicht immer ein Spaziergang. Ich widme diese Arbeit meinem Vater und meiner Mutter, denen ich die leidenschaftliche Neugierde und das nötige Selbstvertrauen zu verdanken haben, um alle Wegabschnitte bewältigen zu können.