Social Psychology

Straight From the Scientist's Mouth—Plain Language Summaries Promote Laypeople's Comprehension and Knowledge Acquisition When Reading About Individual Research Findings in Psychology

Martin Kerwer[1] [a], Anita Chasiotis[1], Johannes Stricker[2], Armin Günther[1], Tom Rosman[1]

[1] Leibniz Institute for Psychology (ZPID), Trier, Germany, [2] Psychology Department, University of Trier, Germany

Easily comprehensible summaries of scholarly articles that are provided alongside 'ordinary' scientific abstracts, so-called plain language summaries, can be a powerful tool for communicating research findings to a wider audience. Using an experimental within-person-design in a preregistered study (*N* = 166), we showed that the comprehensibility for laypeople was higher for plain language summaries compared to scientific abstracts in a psychological journal and also found that laypeople actually understood the corresponding information more correctly for plain language summaries. Moreover, in line with the easiness effect of science popularization, individuals perceived plain language summaries as more credible and were more confident about their ability to make a decision based on plain language summaries. If and under which circumstances this higher perceived credibility is justified, is discussed together with other practical implications and theoretical implications of our findings. In sum, our research further strengthens arguments for providing plain language summaries of psychological research findings by demonstrating that they actually work in practice.

Introduction

Easily comprehensible summaries of scholarly articles that are provided alongside 'ordinary' scientific abstracts, so-called plain language summaries, can be a powerful tool for communicating research findings to a wider audience (FitzGibbon et al., 2020; Kuehne & Olden, 2015). It has been argued, from a health communication perspective, that they may help individuals to overcome the language barrier that is imposed by the jargon of scholarly articles (Nunn & Pinfield, 2014). Thus, plain language summaries facilitate access to research outputs for the general public, which has been discussed as a promising avenue for sustaining trust in science (Grand et al., 2012; Pittinsky, 2015). We argue that sustaining trust in science is especially relevant in (social) psychology—a discipline which investigates topics of high relevance to the public, but which also struggles with a replicability crisis (Klein et al., 2018; Open Science Collaboration, 2015; Świątkowski & Dompnier, 2017) and has been stricken by several misconduct cases lately (Callaway, 2011; Stricker & Günther, 2019; Świątkowski & Dompnier, 2017). Without this trust, scientific findings are likely at risk of being marginalized, which may even lead to a proliferation of conspiracy theories. On a more fine-grained level, one may, however, question if high trust in science is desirable in all circumstances—especially if there is uncertainty with regard to the robustness of scientific findings in a certain discipline. Since plain language summaries are a way of achieving higher transparency of the research process (Barnes & Patrick, 2019; Kuehne & Olden, 2015), we argue that they may provide a way to earn trust for those studies that actually deserve it. A recent study corroborates this argument (Carvalho et al., 2019): scientific articles that provided plain language summaries had a higher methodological quality compared to articles that did not. In line with this reasoning, the American Psychological Association (APA) emphasized that "translating psychological science to the public" may help in addressing common misconceptions on the supposedly lacking rigor of psychological science (Kaslow, 2015).

However, although an APA Task Force published recommendations on this issue in 2014, knowledge on how effective plain language summaries are for communicating findings of individual psychological studies to broader audiences and (re)building trust in (psychological) research findings is limited. The overarching aim of this preregistered study is, therefore, to examine how different types of research summaries affect the information recipient's perception of presented information, as well as his or her knowledge acquisition. More specifically, this study employs openly accessible plain language summaries of peer-reviewed psychological research and a strong experimental

---

[a] mk@leibniz-psychology.org

design to empirically investigate (1) if plain language summaries are better suited for communicating research findings to lay audiences than ordinary scientific abstracts and (2) whether the structure of plain language summaries (subheadings vs. no subheadings) influences their comprehensibility.

Besides this rather straightforward test of the basic notion underlying plain language summaries, the present study examines whether the easiness effect of science popularization (i.e., a stronger reliance on information that is presented in an easily comprehensible manner, cf. Scharrer et al., 2019) can be replicated on a conceptual level for plain language summaries (compared to ordinary scientific abstracts). In other words, we test if providing individuals with plain language summaries instead of ordinary scientific abstracts leads to an increased trust in research findings and consequently a higher reliance (or even an overreliance) on these findings. Moreover, this study explores emotional and behavioral consequences of confronting individuals with plain language summaries (instead of ordinary scientific abstracts). Finally, we also investigate to what extent individual differences (in the effects of plain language summaries) emerge and whether beliefs on the justification of knowledge and English proficiency are able to predict these individual differences.

Plain Language Summaries

In this study, we use the term *plain language summary* to refer to all types of summaries of scientific articles which aim to communicate scientific findings to a broader audience. Typically, these plain language summaries are about the same length as ordinary scientific abstracts, and are written by the authors themselves (cf. FitzGibbon et al., 2020). In contrast to ordinary scientific abstracts, technical language and scientific jargon are avoided and more attention is paid to the background of the presented research and the practical significance of its findings for a lay audience (e.g., Cochrane Methods, 2013; Hauck, 2019). However, who exactly is considered to be a layperson (e.g., practitioners, patients, journalists, policy makers, the public in general) varies greatly between providing agencies, e.g. journals or scientific organizations (Shailes, 2017). Moreover, the term plain language summary is subject to a certain ambiguity for two reasons. First, it is often not used in the strict sense outlined above, but rather to refer to other formats (e.g., blog posts, research digests; see Shailes, 2017). Second, some providers refer to their plain language summaries by other terms such as lay abstracts, lay summaries, translational abstracts, author summaries, or non-technical summaries (cf. FitzGibbon et al., 2020).

Whereas plain language summaries are not well-established in psychology, comparably 'long' traditions of translating scientific results of (mainly) systematic reviews to lay audiences exist in the areas of medicine and public health. For example, Cochrane aims to enable laypersons to make informed health decisions and, for this purpose, has already provided plain language summaries for almost two decades (see Glenton et al., 2010; Santesso et al., 2015). Additionally, Cochrane developed the most comprehensive framework for writing plain language summaries for their authors

(Cochrane Methods, 2013) with guidance on text length (400 – 700 words), use of statistics, or reporting the quality of evidence. Whereas studies on Cochrane's plain language summaries found that they were perceived to be more comprehensible than more traditional research summary types (e.g., Buljan et al., 2018; Santesso et al., 2015), results on information recipients' ability to draw conclusions based on these plain language summaries were mixed (Alderdice et al., 2016; Maguire & Clarke, 2014).

For Cochrane and similar initiatives in other fields (e.g., the German 'Clearing House Unterricht' project for evidence-based teaching methods, see Seidel et al., 2017), meta-analytic or, broadly speaking, systematically reviewed and synthesized findings, commonly serve as a foundation for the provided information. Moving beyond this meta-analytic or systematic review level, lay summaries are now mandatory for clinical trials under a new EU regulation (European Medicines Association, 2019) and commendable initiatives for providing plain language summaries exist in some disciplines, such as the geosciences (Hauck, 2019) or biomedical research (FitzGibbon et al., 2020). Many empirical studies on this type of plain language summary, however, are only based on the technical evaluation of text properties (e.g., Rakedzon et al., 2017). Experimental research on the effectiveness of plain language summaries for communicating the results of individual studies is scarce.

Especially in the field of psychology, plain language summaries could, in the long run, improve the accessibility of research to a broader public and offer a scientific basis for informed decisions. In fact, psychological research questions (e.g., "How does playing violent video games affect children's emotional and cognitive development?") are of particular interest to the broader public. Still, many people lack the skills to understand the scientific abstracts of psychological studies. Furthermore, one may view plain language summaries as a logical next step in the open science movement which is currently gaining momentum in various disciplines, including psychology (cf. Hesse, 2018). In fact, making research outputs openly accessible (i.e., in open access journals free of charges) may be irrelevant to many target groups outside the scientific profession if this information is not provided in a manner that allows wider (lay) audiences to understand it (cf. Nunn & Pinfield, 2014). Potential target groups of plain language summaries in psychology and the social sciences in general include interested laypersons, practicing psychologists, (science) journalists, and students. A recent study based on readability indices (i.e., indices that quantify readability based on text characteristics such as word difficulty or sentence length) revealed better readability of plain language summaries compared to scientific abstracts of psychological journal articles (Stricker et al., 2020). However, this study assessed neither laypersons' actual perceived text comprehensibility nor their actual knowledge acquisition. Thus, to date, it is largely unclear whether plain language summaries have any advantages over scientific abstracts for their actual target audience.

Moreover, our knowledge of how plain language summaries need to be structured in order to efficiently communicate individual research findings is still limited. For example, the corresponding guidance of the American Psy-

chological Association (n.d.) lacks detailed information on this issue. In contrast, guidance (see above) and a template on how to write and structure plain language summaries of *systematic reviews* in the areas of medicine and public health have already been established (Cochrane Norway, 2019). According to this checklist, plain language summaries should be structured based on mandatory and preset subheadings. Subheadings include a plain review title, the aim of the review, key messages, what was studied, what the main results are, and how up-to-date the review is. Drawing on the Cochrane guidelines as well as on recommendations of the expert group on "Summaries of Clinical Trial Results for Laypersons" (Cochrane Norway, 2019; Expert group on clinical trials for the implementation of Regulation (EU) No 536/2014, n.d.), the inclusion of subheadings is a means to make scientific findings more accessible to laypersons. This approach is in line with findings from a randomized controlled trial, where Cochrane plain language summaries of systematic reviews structured by subheadings were better and easier to understand than plain language summaries without subheadings (Santesso et al., 2015). In other words, one might not only examine how plain language summaries perform—when it comes to the information recipient's perception of the presented information and the information recipient's knowledge acquisition—compared to ordinary scientific abstracts but also compared to other types of plain language summaries (e.g., plain language summaries with subheadings vs. plain language summaries without subheadings). Such comparisons have already been systematically carried out (see, for example, Buljan et al., 2018). Albeit, to our knowledge, this mainly applies to plain language summaries of systematic reviews or meta-analyses that, in addition, almost exclusively focus on medicine and public health.

Based on the general assumptions underlying the idea of plain language summaries (i.e., making research more comprehensible and facilitating knowledge acquisition) as well as on the research presented earlier, we therefore derived and preregistered the following hypotheses[1]:

Hypothesis 1. Perceived comprehensibility is higher for plain language summaries with subheadings compared to plain language summaries without subheadings (H1a1) and ordinary scientific abstracts (H1a2). Additionally, perceived comprehensibility is higher for plain language summaries without subheadings compared to ordinary scientific abstracts (H1b).

Hypothesis 2. Knowledge acquisition is higher for plain language summaries with subheadings compared to plain language summaries without subheadings (H2a1) and ordinary scientific abstracts (H2a2). Additionally, knowledge acquisition is higher for plain language summaries without subheadings compared to ordinary scientific abstracts (H2b).

It is essential to consider in this context that plain language summaries without subheadings can directly be de-

rived from plain language summaries with subheadings by excluding existing subheadings (see Method). This means that these conditions do indeed only differ in this one specific aspect (i.e., the presence of subheadings). When it comes to comparisons between scientific abstracts and both types of plain language summaries, the types of presented research summaries differ fundamentally with regard to many aspects, not just subheadings or use of technical terms. These aspects (e.g., usage of statistics or provision of background information) are thus not experimentally varied or controlled. As a consequence, (only) the *type* of presented research summary (of the same scientific publication) is varied here and not one specific aspect of the research summary.

Easiness Effect of Science Popularization

According to the easiness effect of science popularization, individuals rate information to be more trustworthy—and tend to agree more often with corresponding knowledge claims—if it is presented in an easily comprehensible manner (cf. Scharrer et al., 2019). This is possibly due to the fact that—having read easily comprehensible information on scientific subjects—laypeople may "consider that the underlying scientific subject matter is equally easy and uncomplex" (Scharrer et al., 2017, p. 1006). Another potential explanation for this effect—which is closely aligned to findings on the effects of information processing fluency on trust (e.g., Hansen et al., 2008)—would be that laypeople experience information processing as more positive for easily understandable popularized texts which in turn might result in a more positive evaluation of associated knowledge claims (Scharrer et al., 2017).

Empirical evidence in support of this effect emerged across various studies (Scharrer et al., 2012, 2013, 2014, 2017, 2019). For example, Scharrer et al. (2019) found that individuals more strongly agreed to knowledge claims on health-related issues (on a specific health problem) that were presented in an easily comprehensible manner—even if they stemmed from a less trustworthy source (an employee of a pharmaceutical company that produces a drug for this problem). Additionally, other studies revealed that individuals more confidently relied on their own judgments when confronted with more easily comprehensible information (e.g., Scharrer et al., 2014). Since presenting scientific findings in a comprehensible manner is exactly what plain language summaries are supposed to do, we also expect the easiness effect to impact perceptions of plain language summaries. Thus, drawing on the easiness effects of science popularization, we introduce the following (preregistered) hypotheses[2]:

Hypothesis 3. Perceived credibility is higher for plain language summaries with subheadings compared to ordinary scientific abstracts (H3a) and for plain language summaries without subheadings compared to ordinary scientific

---

1 These hypotheses were preregistered at PsychArchives (http://dx.doi.org/10.23668/psycharchives.2772).

2 Once more, these hypotheses were preregistered at PsychArchives (http://dx.doi.org/10.23668/psycharchives.2772).

abstracts (H3b).

Hypothesis 4. Perceived confidence in one's ability to evaluate the study is higher for plain language summaries with subheadings compared to ordinary scientific abstracts (H4a) and for plain language summaries without subheadings compared to ordinary scientific abstracts (H4b).

Hypothesis 5. Perceived ability to make a decision (without consulting an expert) is higher for plain language summaries with subheadings compared to ordinary scientific abstracts (H5a) and for plain language summaries without subheadings compared to ordinary scientific abstracts (H5b).

We did not specify any confirmatory hypotheses on differences between plain language summaries with/without subheadings since the easiness effect of science popularization has been most frequently shown in studies that consider 'text easiness' in terms of translating technical terms into familiar words (e.g., Scharrer et al., 2014). These previous studies strictly controlled for other characteristics of the text structure (e.g., its layout) by keeping them constant across conditions (e.g., Scharrer et al., 2012). Moreover, even if the original layout of popularized and scientific articles was retained (e.g., Scharrer et al., 2017), effects of specific text characteristics, such as the inclusion of subheadings, were not examined. Thus, little is known about how specific aspects of the text structure (in our case subheadings) might mitigate or amplify the easiness effect of science popularization. To shed more light on this issue, we will, therefore, also explore differences between plain language summaries with/without subheadings in the perceived study credibility, in one's perceived ability to evaluate the study, and in one's perceived ability to make a decision.

Exploratory Research Questions

Epistemic Emotions

When trying to address an information need, individuals will (in most circumstances) strive to understand the information that is conveyed within research summaries. Such aims or goals that are related to acquiring knowledge are often referred to as *epistemic aims* (Chinn et al., 2011). Educational research has shown that pursuing the epistemic aim of understanding something may have emotional consequences (cf. Muis et al., 2018). Muis et al. (2018) referred to the specific type of emotion considered here as *epistemic emotions*—emotions "that occur in epistemically related contexts" (p. 169). Reviewing current literature, they argued that whether positive or negative emotions arise depends on an individual's success (or failure) in achieving their epistemic aims (e.g., understanding information in research summaries). Since individuals should be more likely to achieve the aim of understanding research summaries for plain language summaries compared to ordinary scientific abstracts, we suggest that individuals experience more positive emotions (especially curiosity) and less negative emotions (such as confusion, boredom and frustration) when reading plain language summaries compared to ordinary scientific abstracts.

Full Text Access

For reasons of ecological validity, we strived to extend the scope of our measurement by not only including self-reports and a knowledge test, but also investigating behavioral consequences of reading different types of plain language summaries compared to ordinary scientific abstracts. The most obvious behavioral consequence of reading plain language summaries is whether individuals subsequently opt to access the corresponding full text or not (i.e., whether they intend to seek more information on the issue at hand by reading the corresponding article). For instance, individuals might find the study, due to a better understanding, more interesting and relevant when they read a plain language summary (compared to an ordinary scientific abstract), which is why they might choose to read its full text. They might, however, also realize after reading the plain language summary that the corresponding study is in fact irrelevant to them, while they are unable to draw this conclusion after reading the less comprehensible ordinary scientific abstract. Consequently, the opposite may also be true—individuals might, in this case, be more likely to opt *against* reading the full text after reading the plain language summary *but not* after reading the ordinary scientific abstract. As prior research failed to address this research question despite its considerable practical relevance—we aim to examine this type of behavioral consequence by means of exploratory analyses.

The Role of Justification Beliefs and English Proficiency

To gain more insights into the role of specific reader characteristics and their interaction with the type of presented research summary, we aim to examine, in exploratory analyses, (1) whether individual differences exist in the perception of and knowledge acquired through research summaries (regardless of the type of presented summary), (2) whether individual differences exist in the effects of plain language summaries compared to ordinary scientific abstracts, and (3) whether theoretically relevant predictors (i.e., reader characteristics) are able to explain these individual differences.

More specifically, we will explore effects of epistemic justification beliefs (i.e., beliefs on criteria for 'accepting' knowledge claims) as a potential predictor of individual differences in the perception of research summaries. To date, there exists a rather strong body of evidence suggesting that the way in which individuals choose between, evaluate and comprehend digital and non-digital sources depends, among others, on their epistemic beliefs (see Barzilai & Strømsø, 2018 and Strømsø & Kammerer, 2016 for a review). In particular, epistemic beliefs about the justification of knowledge have been shown to influence how individuals act in tasks which are related to multiple source use and multiple document comprehension (e.g., Bråten et al., 2013). Drawing on this line of research, one might, for example, argue that individuals who believe that knowledge is verified by authority (i.e., the expertise ascribed to the source of the knowledge claim) are likely to perceive a plain language summary to be less trustworthy if it no longer in-

cludes typical cues pertaining to the expertise of the source (e.g., the 'scientificness' of the writing; cf. Thomm & Bromme, 2012). Apart from justification by authority, other frequently distinguished positions on the justification of knowledge are personal justification (knowledge claims are perceived to be strongly based on personal opinions or experiences) or justification by multiple sources (validation of knowledge claims by means of cross-validation)—both of which will also be subjected to exploratory analyses.

Finally, in the context of communicating findings to wider audiences, it is often neglected that, worldwide, most people are not native English speakers. However, already in 2010, Clayman et al. showed that health information seeking behaviors differ depending on an individual's language skills. More specifically, Hispanics with low self-reported English proficiency rated health information from English media to be less trustworthy and reported lower access rates to these media in comparison to Hispanics who were more comfortable speaking English. To cross this barrier, Cochrane provides plain language summaries in different languages (e.g., Spanish, French, Portuguese) to reach those 75 % of the world's population who do not speak English at all (Behmen et al., 2019; Cochrane Collaboration, 2019). We therefore argue that studies investigating English plain language summaries in non-native speakers should control for English language proficiency. Yet, plain language summaries per se might also compensate for reduced English skills (at least to some degree) as they are written in an easily comprehensible manner and thus potentially more accessible to individuals with limited language skills. Therefore, we will also explore the relation of English proficiency with the predefined outcome variables.

Method

Design

We used a within-person experimental design with one factor (plain language summaries with subheadings, plain language summaries without subheadings, ordinary scientific abstracts). The design of our study as well as its procedures, hypotheses and statistical analyses were preregistered at PsychArchives (http://dx.doi.org/10.23668/psycharchives.2772).

Materials

All research summaries were extracted from the *Journal of Social and Political Psychology* (JSPP, https://jspp.psychopen.eu, see Table 1 for a list of included studies). JSPP is an applied journal at the intersection of social and political psychology which aims to publish research that is relevant to education and practice—without any restrictions regarding methodological and theoretical approaches. To reach out to lay audiences, JSPP allows its authors to provide plain language summaries (labelled as non-technical summaries). Providing plain language summaries is, however, not mandatory. The journal is part of the Web of Science Core Collection since 2020. We chose this journal because (1) it includes plain language summaries, (2) its articles and their abstracts are openly available under a CC-BY license, which permits adaptation (for creating different types of plain language summaries), reproduction (for creating intervention materials) and distribution (for providing full text access), and (3) its scope (i.e., research on social problems and social justice) is relevant for a wider audience. For our study, we selected 12 out of 32 available articles with plain language summaries (as of September, 2019) based on their research question's relevance to a broader German audience (e.g., studies on Brexit, the refugee crisis in Europe, group formation in xenophobic web forums, see Table 1). Moreover, a secondary criterion was that it was possible to build two knowledge test items which were identical for the plain language summaries and the ordinary scientific abstracts of the articles (see below). Since JSPP includes subheadings for its plain language summaries (1. Background, 2. Why was this study done, 3. What did the researchers do and find, 4. What do these findings mean), we additionally created plain language summaries without subheadings by removing these subheadings. DOIs and some basic properties of research summaries, such as readability scores, are listed in Table 1. Since articles and abstracts are openly accessible via DOIs at the journal homepage, we do not reproduce them in the present paper.

Participants and Procedures

Participants were recruited at Trier University via mailing lists, Facebook groups, and flyers. In the advertisements for recruiting participants, we stated that our study would examine the extent to which non-scientists find psychological research comprehensible. Moreover, participants were informed that they would have to read summaries of English research articles on social and political psychology in our study. Accordingly, we applied the following eligibility criteria: Participants had to be students at Trier University, had to be aged 18 to 70 years, had to have German language skills at native speaker level and had to consider their English language reading skills as sufficient to comprehend English research summaries. Data collection started on December 9, 2019, and ended on February 11, 2020. The actual sample included 166 students (71.08 % female) with a mean age of $M = 24.03$ ($SD = 4.04$, ranging from 18 to 48) years. Participants studied various subjects (e.g., educational sciences, economics, history), whereby psychology students were most strongly represented (38.55 %). All data collection procedures took place on a single measurement occasion for groups of up to 15 participants (minimum size two participants, median size 11 participants) at a computer lab using the survey software Unipark[3]. During the study, all instructions and questionnaires were administered in German language, whereas the research summaries themselves were not translated (i.e., presented in English language).

Data collection started with covariate measurements (i.e., justification beliefs, demographics, English language

---

3 https://www.unipark.com/en/

Table 1. List of included studies and descriptive statistic for research summaries.

Study Title	DOI	Number of Words$_{PLS}$	Number of Words$_{OSA}$	SMOG$_{PLS}$	SMOG$_{OSA}$
The Politicized Motivations of Volunteers in the Refugee Crisis: Intergroup Helping as the Means to Achieve Social Change	10.5964/ jspp.v5i1.642	337	177	19.287	22.248
Collective Memory as Tool for Intergroup Conflict: The Case of 9/11 Commemoration	10.5964/ jspp.v5i2.713	341	167	17.564	23.885
From I to We: Group Formation and Linguistic Adaption in an Online Xenophobic Forum	10.5964/ jspp.v6i1.741	438	163	12.952	16.527
Collective Memory of a Dissolved Country: Group-Based Nostalgia and Guilt Assignment as Predictors of Interethnic Relations Between Diaspora Groups From Former Yugoslavia	10.5964/ jspp.v5i2.733	657	154	17.067	20.457
One World in Diversity – A Social-Psychological Intervention to Foster International Collective Action Intention	10.5964/ jspp.v6i1.601	349	239	17.506	18.564
Self-Censorship Orientation: Scale Development, Correlates and Outcomes	10.5964/ jspp.v6i2.859	216	161	15.645	18.244
A Field Study Around a Racial Justice Protest on a College Campus: The Proximal Impact of Collective Action on the Social Change Attitudes of Uninvolved Bystanders	10.5964/ jspp.v7i1.1063	251	170	14.811	16.648
The Meaning of Being German: An Inductive Approach to National Identity	10.5964/ jspp.v7i1.557	241	150	16.404	18.458
Agentic and Communal Interaction Goals in Conflictual Intergroup Relations	10.5964/ jspp.v7i1.746	538	179	20.267	20.131
Contempt of Congress: Do Liberals and Conservatives Harbor Equivalent Negative Emotional Biases Towards Ideologically Congruent vs. Incongruent Politicians at the Level of Individual Emotions?	10.5964/ jspp.v7i1.822	569	193	20.176	17.946
Seen One, Seen 'Em All? Do Reports About Law Violations of a Single Politician Impair the Perceived Trustworthiness of Politicians in General and of the Political System?	10.5964/ jspp.v7i1.933	436	206	18.297	20.403
How Many Ways to Say Goodbye? The Latent Class Structure and Psychological Correlates of European Union Sentiment in a Large Sample of UK Adults	10.5964/ jspp.v7i1.981	434	193	15.836	16.828

Note: SMOG = Simple Measure of Gobbledygook (Mc Laughlin, 1969): lower scores indicate better readability; DOI = digital object identifier; indices specify summary type: PLS = plain language summary, OSA = ordinary scientific abstract. For plain language summaries with subheadings, there were additional 18 words which were, however, the same for all studies and are not included in the number of words reported here.

proficiency, etc.). Thereafter, twelve research summaries were presented in four blocks, whereby each block contained three research summaries (one for each condition: plain language summary with/without subheading, ordinary scientific abstract, see Figure 1 for an illustrative example). Hence, each participant read *one* research summary on *each* of twelve studies (i.e., all participants read research summaries on all studies that are presented in Table 1). Both the type of research summary for each study and the order of studies were randomized. We restricted each type of research summary to occur once in each of the four blocks so that each participant received each type of re-

search summary (e.g., plain language summaries without subheadings) four times. After each block (except the last one), there was a break of 90 seconds. All dependent variables were assessed after the corresponding summary (i.e., twelve times in total)—except the knowledge acquisition test, which was conducted at the end of each block (see Figure 2 for a graphical outline of the experimental design). After the data collection, each participant received a compensation of 20 Euros (around 22 US Dollars at the time of data collection). The median study duration was approximately 65 minutes.

How Many Ways to Say Goodbye? The Latent Class Structure and Psychological Correlates of European Union Sentiment in a Large Sample of UK Adults

Zusammenfassung

Most psychological research on Brexit categorises participants as either leave or remain supporters. In the current study we take a data-driven approach and identify different clusterings of attitudes towards the European Union (EU) using latent class analysis (LCA), as well as how these classes differ across a range of important social and psychological variables. This analysis revealed 10 distinct classes of voters in a large (N = 15860) adult sample of UK citizens using data from the British Election Study. These classes ranged from being quite uniformly pro- or anti-EU in sentiment, to more mixed groups with more complex patterns of attitudes. The classes that included majority-remain supporters were younger and better educated, and self-rated more highly on the measures of actively open-minded thinking, openness, political trust, and external locus of control. The classes that included majority-leave supporters were older and less well educated, and self-rated more highly on the measures of authoritarianism and conscientiousness. However, there were also notable demographic and psychological differences within the classes associated with leavers and remainers. A full consideration of these attitudinal nuances will be necessary to achieve a deeper understanding of why the UK decided to leave the EU.

Background
On June 23rd 2016 the United Kingdom (UK) – via an advisory referendum – voted to leave the European Union (EU). Perhaps unsurprisingly, there has been considerable interest concerning the demographic and psychological traits associated with 'leavers' vs 'remainers'. Much of the research and public discourse around Brexit has since adopted this dichotomized framework. However, this approach is almost certainly an oversimplification of the general public's views towards the EU.

Why was this study done?
We wanted to examine whether attitudes towards the EU were more nuanced than this simple dichotomy. For example, perhaps there are individuals who are generally supportive of the EU, but opposed to the EU on specific issues such as immigration or sovereignty. Or perhaps there are individuals who are generally opposed to the EU, but value features such as its ability to minimize war. We also wanted to see whether these different 'types' of EU supporters could be distinguished on a range of demographic (e.g. educational attainment) and psychological (e.g. personality traits) characteristics.

What did the researchers do and find?
We used a large dataset (>15k individuals) from the British Election Study and a statistical technique called latent class analysis to determine how many 'types' of EU supporter exist in the general public. We found that 10 types of EU supporter were present in the data, roughly split across leavers and remainers. Some of these types were staunchly pro- or anti-EU across a range of issues. Others were more nuanced. For example, one of our types characterised individuals who were pro-immigration, but who also believed that the EU created more red-tape and were dissatisfied with EU democracy. A number of demographic and personality variables were strongly associated with some of these different EU types. The classes that included majority-remain supporters were younger and better educated, and self-rated more highly on measures of actively open-minded thinking, openness, political trust, and external locus of control. The classes that included majority-leave supporters were older and less well educated, and self-rated more highly on measures of authoritarianism and conscientiousness. However, there were also notable demographic and psychological differences within the classes between remain classes, and between leave classes.

What do these findings mean?
These findings tell us that viewing Brexit through a simple leave vs remain lens misses much of what people think about the EU, as well as the demographic and psychological characteristics that are related to these EU attitudes. If we want to gain a richer, more complete understanding of the reasons why the UK voted to leave the EU we will need to understand not just what caused people to vote leave or remain, but understand the demographic and psychological factors associated with different clusters of attitudes towards the EU.

On June 23rd 2016 the United Kingdom (UK) – via an advisory referendum – voted to leave the European Union (EU). Perhaps unsurprisingly, there has been considerable interest concerning the demographic and psychological traits associated with 'leavers' vs 'remainers'. Much of the research and public discourse around Brexit has since adopted this dichotomized framework. However, this approach is almost certainly an oversimplification of the general public's views towards the EU. We wanted to examine whether attitudes towards the EU were more nuanced than this simple dichotomy. For example, perhaps there are individuals who are generally supportive of the EU, but opposed to the EU on specific issues such as immigration or sovereignty. Or perhaps there are individuals who are generally opposed to the EU, but value features such as its ability to minimize war. We also wanted to see whether these different 'types' of EU supporters could be distinguished on a range of demographic (e.g. educational attainment) and psychological (e.g. personality traits) characteristics. We used a large dataset (>15k individuals) from the British Election Study and a statistical technique called latent class analysis to determine how many 'types' of EU supporter exist in the general public. We found that 10 types of EU supporter were present in the data, roughly split across leavers and remainers. Some of these types were staunchly pro- or anti-EU across a range of issues. Others were more nuanced. For example, one of our types characterised individuals who were pro-immigration, but who also believed that the EU created more red-tape and were dissatisfied with EU democracy. A number of demographic and personality variables were strongly associated with some of these different EU types. The classes that included majority-remain supporters were younger and better educated, and self-rated more highly on measures of actively open-minded thinking, openness, political trust, and external locus of control. The classes that included majority-leave supporters were older and less well educated, and self-rated more highly on measures of authoritarianism and conscientiousness. However, there were also notable demographic and psychological differences within the classes between remain classes, and between leave classes. These findings tell us that viewing Brexit through a simple leave vs remain lens misses much of what people think about the EU, as well as the demographic and psychological characteristics that are related to these EU attitudes. If we want to gain a richer, more complete understanding of the reasons why the UK voted to leave the EU we will need to understand not just what caused people to vote leave or remain, but understand the demographic and psychological factors associated with different clusters of attitudes towards the EU.

Figure 1. Illustrative example of intervention materials. For each study, only one out of these three options was shown. Left: Ordinary scientific abstract. Center: Plain language summary with subheadings. Right: Plain language summary without subheadings. The heading "Zusammenfassung" is German and means "summary".
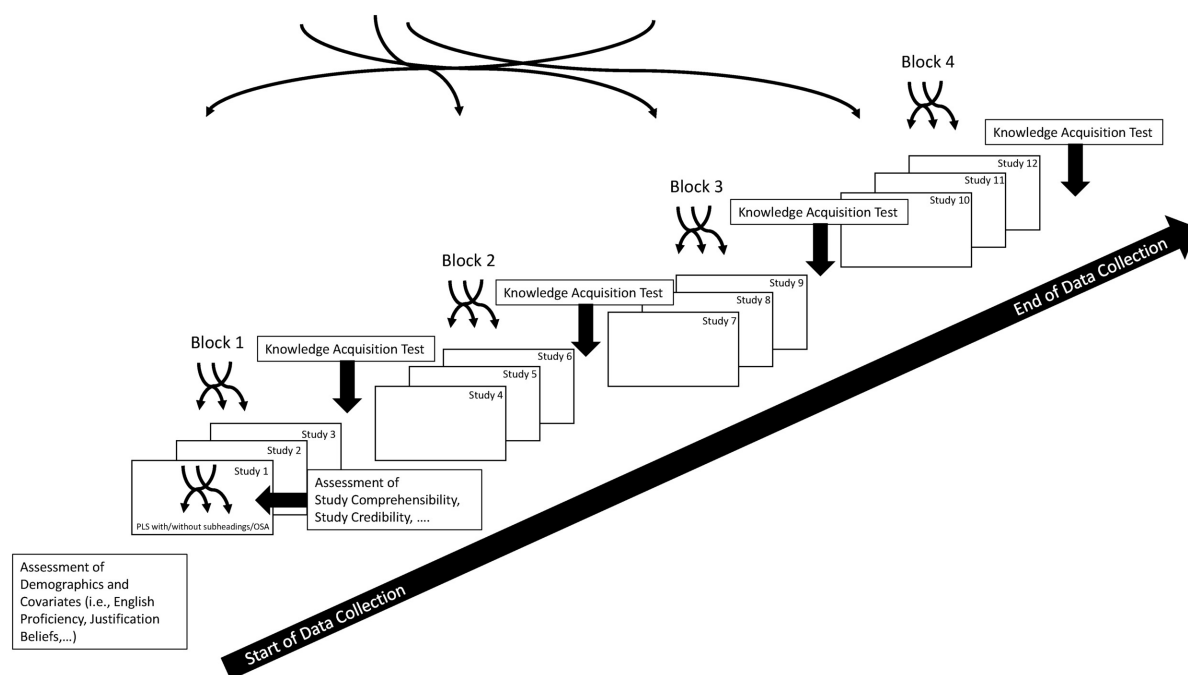
Figure 2. Graphical outline of the experimental design. Overlapping curved arrows indicate that randomization took place for the order of blocks (curved arrows above blocks), the order of studies within blocks (curved arrows above studies) and the type of presented research summary (exemplary curved arrows within Study 1). PLS = plain language summary, OSA = ordinary scientific abstract.

Variables

Outcome Variables

For each research summary, perceived comprehensibility and perceived study credibility were assessed on 1 to 8 semantic differentials ranging from "not comprehensible/ credible at all" to "extremely comprehensible/credible". Moreover, to measure our participants' perceived confidence in their ability to evaluate the study and their perceived ability to make a decision without further information, they were asked to indicate their agreement on Likert scales ranging from 1 ("I do not agree at all") to 8 ("I totally agree") regarding the following statements: "Based on this summary, I am able to evaluate the veracity of the corresponding study./Based on this summary, I am able to make a decision without needing any further information (i.e., reading the full text or talking to an expert).".

Using the short version of the EES questionnaire (Pekrun et al., 2017), we assessed to what extent our participants experienced the following epistemic emotions while reading each research summary on 5-point Likert scales: curiosity (positive epistemic emotion), boredom, confusion, frustration (negative epistemic emotions). Additionally, participants could request the link to the corresponding full text (yes/no) and were informed that they would receive this link after finishing the study ("I want to receive the link to this study after today's data collection is finished.").

To assess knowledge acquisition, after each block (see Figure 2), participants had to indicate for six statements—two on each of the three studies that were present-ed in the block—whether they deemed these statements to be correct or incorrect. Importantly, all statements could, in principle, be correctly answered based on both plain language summaries and ordinary scientific abstracts. In total, we created 14 correct statements and ten distractors/incorrect statements. One distractor stated, for example, that a study revealed that anger resulted in both political activism and volunteerism. In contrast, both plain language summary and ordinary scientific abstract of the corresponding study stressed that a connection between anger and political activism but not volunteerism was found. Before data analysis, all data on the knowledge acquisition test measure were recoded from 'correct/incorrect' to 'right answer/ wrong answer' (i.e., correctly differentiating correct statements from distractors).

Covariates

Psychology-specific justification beliefs were measured using an adaption of Klopp and Stark's (2016) domain-general questionnaire, which, in turn, builds upon a measurement instrument by Bråten et al. (2013). Participants indicated their agreement to nine items on a 6-point Likert scale to assess the following three dimensions of justification beliefs: justification by authority (McDonald's omega = .69), personal justification (omega = .75) and justification by multiple sources (omega = .70). A sample item for justification by multiples sources is "To be able to trust knowledge claims in psychology, various knowledge sources have to be checked" (Klopp & Stark, 2016). English proficiency was measured by the E-PA ("Englischtest für die Person-

alauswahl" [English test for personnel selection]), a normed and validated German measurement instrument for assessing English proficiency of German adults (Liepmann et al., 2013). Moreover, we assessed other covariates (i.e., demographics, self-reported ability to evaluate knowledge claims of scientific studies [1 to 8 semantic differential] and self-reported familiarity with scientific studies [1 to 8 semantic differential]). Finally, perceived 'scientificness' of the summaries was measured on the level of the individual research summaries, using a 1 to 8 semantic differential, ranging from "not scientific at all" to "extremely scientific".

Statistical Analysis

Statistical Models

We employed mixed models to analyze our data in the statistical environment R (R Core Team, 2019) with the lme4 (Bates et al., 2015) and lmertest (Kuznetsova et al., 2017) packages. Random factors in our model were "study" (on which the research summary was based) and "subject" (i.e., participant). In other words, we accounted for systematic variation in the individual perception of research summaries at the participant-level (that participants perceived research summaries to be in general more [or less] comprehensible compared to other participants) and at the study-level (that research summaries on a specific study were consistently perceived to be more [or less] comprehensible compared to research summaries on other studies). Independent variables were dummy-coded research summary type variables (for ordinary scientific abstracts and plain language summaries without subheadings), where plain language summaries with subheadings were used as reference category, and an additional contrast comparing ordinary scientific abstracts and plain language summaries without subheadings was computed with the multcomp package (Hothorn et al., 2008). To facilitate the interpretation of our results, we computed standardized effect estimates. For this purpose, we divided effect estimates by the residual standard deviations at the text-level (i.e., the proportion of variance that cannot be explained by the fixed effects of research summary type or the random effects for studies and participants; see Westfall, 2016).

For dichotomous outcome variables (i.e., full text access, and knowledge acquisition test items) we applied logistic linear mixed-effects models for binary outcomes. Since two knowledge test items existed for each study, an additional nested random effect of items within study was included in the corresponding analyses.

For all outcome variables that were included in confirmatory analyses, we explored individual differences and whether these individual differences could be explained by justification beliefs and English language proficiency in two steps. To do this, we employed likelihood ratio (LR) tests.

LR tests compare two nested models (e.g., models that include justification beliefs/English language proficiency as predictors or not) based on the ratio of their likelihoods. More specifically, they can be used to determine if additional parameters (e.g., random slope variances) significantly improve the model fit. The anova function in R provides an approximately chi-square distributed test statistic for LR tests. The degrees of freedom of this test statistic's distribution are the number of parameters tested (i.e., in our case, the number of additional fixed effects or random effect [co]variances). In our first step, LR tests were computed to determine whether significant individual differences in the overall information perception existed (random intercepts on the participant level) or in the effects of plain language summaries (random slopes on the participant level). In a second step, we predicted the identified individual differences by means of linear effects of justification beliefs or English proficiency (for random intercepts) and interactions between justification beliefs or English proficiency and research summary type (for random slopes). Once more, LR tests were computed to test if (these sets of) linear or interaction effects were statistically significant.

Sample Size Calculation

Based on the introductory paper by Judd et al. (2017) and their tool on power analysis for experimental designs with more than one random factor[4], we performed a power analysis which indicated that a sample of 150 participants would be sufficient to achieve a power of .908 for a medium-sized effect (specification: d = 0.500, Residual VPC[5]: 0.500, Participant intercept VPC: 0.175, Target intercept VPC: 0.175, Participant-by-Target VPC: 0.050, Participant slope VPC: 0.050, Target slope VPC: 0.050, Total number of Targets: 12).

Data Cleaning

Two participants failed to complete the study in the available timeframe (approximately 100 minutes), so that data on four research summaries is missing for these participants. Moreover, we had to discard data on one (out of twelve) research summaries for 43 participants because of an error in the Unipark script.

Results

Table 2 provides descriptive statistics of covariates on participant-level. According to the test manual, the mean English language proficiency score corresponds to a stanine value of 8, indicating that English language proficiency was high in our sample. Table 3 provides text-level descriptive statistics on all outcome variables, while descriptive statistics of continuous confirmatory outcomes are also visualized in Figure 3. Since perceived 'scientificness' was strong-

---

4 https://jakewestfall.shinyapps.io/two_factor_power/

5 The acronym VPC is short for variance partitioning coefficient and denotes "the relative magnitude of the estimable variance components" (Judd et al., 2017, p. 18).

Table 2. Sample description.

Covariate	Measurement	M	SD
Semester	Number of semesters participants studied with regard to the primary degree[1] they currently pursued	4.158	3.362
English Language Proficiency	Score on the short scale of a standardized English as a second language test	51.253[2]	6.063
Personal Justification	Mean value of the personal justification scale. Assessed by 3 items on 1 to 6 Likert scales	2.299	0.914
Justification by Authority	Mean value of the justification by authority scale. Assessed by 3 items on 1 to 6 Likert scales	3.819	0.819
Justification by Multiple Sources	Mean value of the justification by multiple sources scale. Assessed by 3 items on 1 to 6 Likert scales	5.026	0.783

Note: M = arithmetic mean; *SD* = standard deviation; *N* is 166 for all variables but semester; for semester *N* is 165 since one participant provided no answer on the corresponding item.
[1] 102 students (61.45 %) were enrolled in Bachelor degree programs, 57 students (34.34 %) in master degree programs and 7 (4.22 %) in other study programs.
[2] This mean English language proficiency score corresponds to a stanine value of 8 according to the test manual

ly correlated with perceived credibility ($r > .77$) and was not part of any of our hypotheses, we opted against analyzing this variable.

Confirmatory Analyses

Hypothesis 1: Perceived Comprehensibility

As expected, perceived comprehensibility was higher when subjects read plain language summaries with subheadings compared to plain language summaries without subheadings and ordinary scientific abstracts, but also in plain language summaries without subheadings compared to ordinary scientific abstracts (see Table 3). Mixed model analyses revealed that differences between all conditions were significant (all ps < .001; see Table 4). As can be seen in Figure 3A, these differences were quite large (e.g., mean comprehensibility scores for plain language summaries with subheadings and scientific abstract differed by more than 0.600 residual standard deviations at the text-level, see Table 4). Thus, H1 was fully confirmed.

Hypothesis 2: Knowledge Acquisition

In total, 2,932 of 3,886 answers on knowledge items were correct (75.45 %), with item difficulty (i.e., proportion of correct responses) ranging from .58 to .95. In accordance with the hypothesized pattern of effects, the rate of correctly answered items was higher for plain language summaries with subheadings (78.68 %) compared to plain language summaries without subheadings (75.91 %) and ordinary scientific abstracts (71.98 %). Mixed model analyses for binary variables indicated that differences between plain language summaries without subheadings and plain language summaries with subheadings ($z = -2.08$, $p = .019$), as well as between ordinary scientific abstracts and plain lan-

guage summaries with subheadings ($z = -4.54$, $p < .001$) and ordinary scientific abstracts and plain language summaries without subheadings ($z = -2.51$, $p = .006$) were significant. Thus, H2 was also fully confirmed.

Hypothesis 3, 4 and 5: Perceived Credibility, Perceived Ability to Evaluate the Corresponding Study and Perceived Ability to Make a Decision Based on the Information Provided

As expected, perceived credibility, perceived ability to evaluate the corresponding study as well as perceived ability to make an informed decision were higher when subjects read plain language summaries with subheadings compared to ordinary scientific abstracts (see Table 3). Scores of all three outcomes were also higher in plain language summaries with subheadings compared to plain language summaries without subheadings (albeit no expectation on the difference between plain language summaries with/without subheadings was specified in our hypotheses). Mixed model analyses revealed that these differences were significant (see Table 4 and Figure 3). All effects on these measures were, however, considerably smaller (though still practically relevant) when compared to the corresponding effects that were obtained for perceived comprehensibility and ranged from 0.135 to 0.287 residual standard deviations (see Table 4). The expected difference between plain language summaries without subheadings and ordinary scientific abstracts did not emerge in our study (see Table 4 and Figure 3). Corresponding effects were very small and not practically relevant (less than 0.040 residual standard deviations). As a consequence, H3a, H4a and H5a were fully confirmed, but H3b, H4b, and H5b were not.

Table 3. Means, standard deviations and intra-class-correlation coefficients of outcome variables.

Outcome	Scale	ICC$_{Participants}$	ICC$_{Study}$	PLS with subheading		PLS without subheading		ordinary scientific abstract	
				M	SD	M	SD	M	SD
Comprehensibility	1 to 8, semantic differential	.178	.182	6.203	1.651	5.964	1.791	5.443	1.980
Credibility	1 to 8, semantic differential	.339	.042	5.368	1.584	5.199	1.646	5.193	1.611
Ability to evaluate	1 to 8, Likert scale	.511	.020	3.703	1.829	3.511	1.843	3.468	1.857
Ability to make a decision	1 to 8, Likert scale	.442	.016	3.103	1.755	2.783	1.717	2.767	1.695
Curiosity	1 to 5, Likert scale	.246	.077	2.950	1.152	2.885	1.210	2.627	1.158
Boredom	1 to 5, Likert scale	.263	.070	1.674	0.936	1.759	0.969	1.779	0.984
Confusion	1 to 5, Likert scale	.189	.115	1.602	0.855	1.738	0.951	1.997	1.098
Frustration	1 to 5, Likert scale	.283	.040	1.410	0.775	1.531	0.921	1.611	0.981
N		1947	1947	620	620	663	663	664	664

Note: M = arithmetic mean; *SD* = standard deviation; ICC = intra-class-correlation coefficient; PLS = plain language summary.
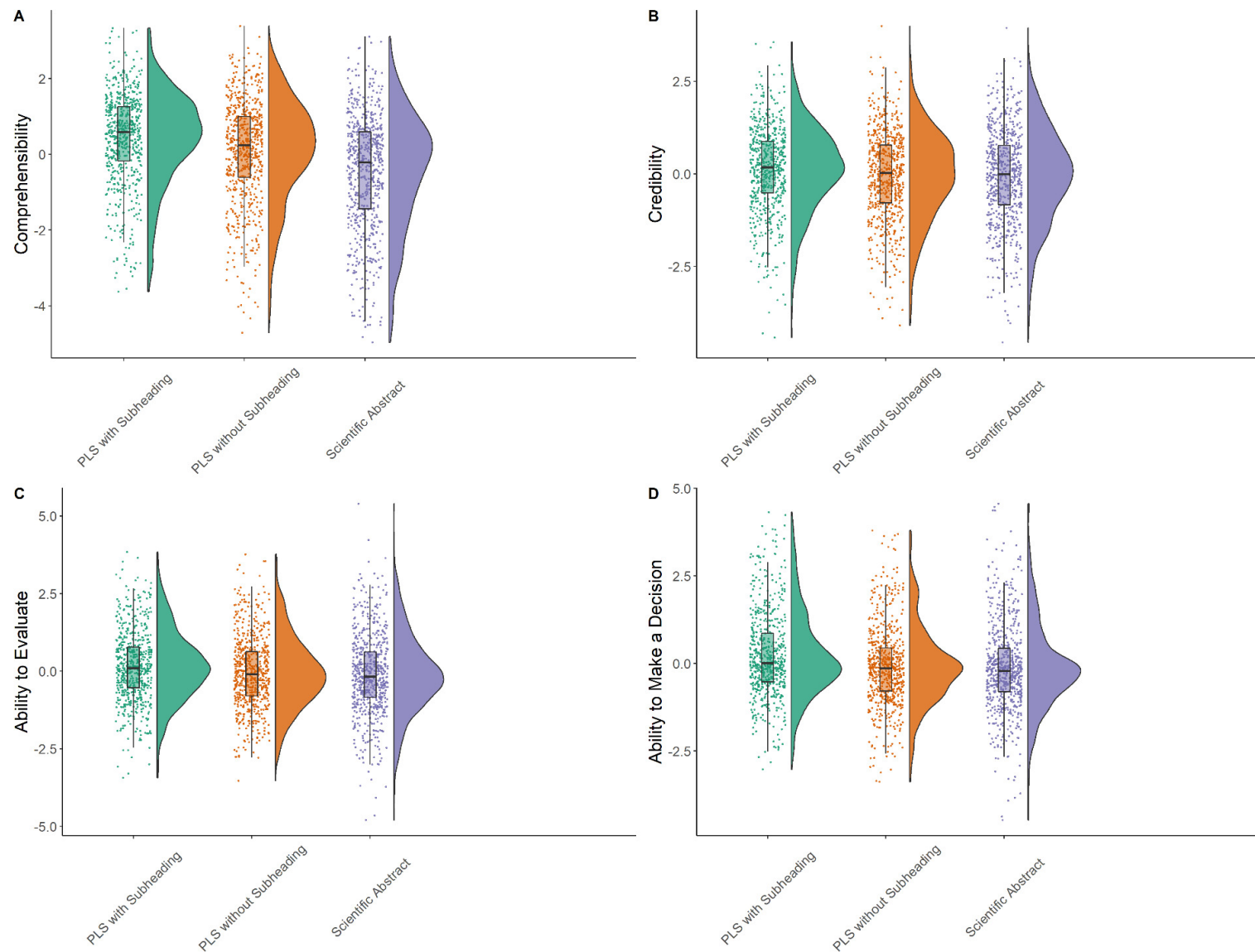
Figure 3. Raincloud plots for continuous confirmatory outcomes. Residual scores for comprehensibility (A), credibility (B), ability to evaluate (C) and ability to make a decision (D) are depicted separated by experimental condition (PLS = plain language summary). Residual scores were obtained from a mixed model that controlled for study and participant as random factors.

Table 4. Results of confirmatory analyses on comprehensibility (H1), credibility (H3), ability to evaluate (H4) and ability to make a decision (H5). Estimates are based on mixed models with (contrasts of) fixed effects for type of presented research summary (plain language summary with/without subheadings and ordinary scientific abstract) and random effects for participant and study.

Outcome	Parameter	EST	SE	p	STD.EST
Comprehensibility	Random Effect Participant	0.655			
	Random Effect Study	0.669			
	Residual	2.034			
	Intercept (PLS with Subheadings)	6.305	0.251	<.001	
	No Subheadings	-0.352	0.081	<.001	-0.247
	OSA	-0.894	0.080	<.001	-0.627
	OSA – No Subheadings	-0.542	0.078	<.001	-0.380
	Marginal R^2/Conditional R^2	.038/.418			
Credibility	Random Effect Participant	0.897			
	Random Effect Study	0.115			
	Residual	1.605			
	Intercept (PLS with Subheadings)	5.365	0.133	<.001	
	No Subheadings	-0.171	0.072	.017	-0.135
	OSA	-0.182	0.071	.011	-0.143
	OSA – No Subheadings	-0.011	0.070	.877	-0.009
	Marginal R^2/Conditional R^2	.003/.388			
Ability to evaluate	Random Effect Participant	1.748			
	Random Effect Study	0.083			
	Residual	1.573			
	Intercept (PLS with Subheadings)	3.726	0.142	<.001	
	No Subheadings	-0.220	0.071	.002	-0.175
	OSA	-0.267	0.071	<.001	-0.213
	OSA – No Subheadings	-0.047	0.069	.492	-0.038
	Marginal R^2/Conditional R^2	.004/.540			
Ability to make a decision	Random Effect Participant	1.328			
	Random Effect Study	0.060			
	Residual	1.582			
	Intercept (PLS with Subheadings)	3.117	0.125	<.001	
	No Subheadings	-0.336	0.071	<.001	-0.267
	OSA	-0.361	0.071	<.001	-0.287
	OSA – No Subheadings	-0.025	0.069	.717	-0.020
	Marginal R^2/Conditional R^2	.009/.472			

Note. EST = estimates for variances of residuals and random effects, unstandardized regression weights and marginal/conditional R^2 (see Nakagawa et al., 2013); *SE* = standard error, *p* = *p*-value of two-tailed significance test; *STD. EST* = standardized regression weights; PLS = plain language summary, OSA = ordinary scientific abstract; comprehensibility/credibility were measured on 1 to 8 semantic differentials and ability to evaluate/to make a decision on 1 to 8 Likert scales

Exploratory Analyses

Epistemic Emotions

Descriptively, we found stronger negative epistemic emotions for ordinary scientific abstracts compared to plain language summaries with or without subheadings (see Table 3). Moreover, subjects that read plain language summaries with subheadings were significantly less bored, less frustrated, and less confused than subjects reading plain language summaries without subheadings or ordinary scientific abstracts, but were (descriptively, compared to plain language summaries without subheadings, see Table 3 and Table 5) more curious. Moreover, apart from boredom, all corresponding differences between plain language summaries without subheadings and ordinary scientific abstracts also reached significance (see Table 5). In terms of residual standard deviations, the largest differences between plain language summaries and ordinary scientific abstracts were obtained for confusion (0.334 without sub-

Table 5. Results of exploratory analyses on epistemic emotions. Estimates are based on mixed models with (contrasts of) fixed effects for type of presented research summary (plain language summary with/without subheadings and ordinary scientific abstract) and random effects for participant and study.

Epistemic Emotion	Parameter	EST	SE	*p*	STD.EST
Curiosity	Random Effect Participant	0.353			
	Random Effect Study	0.114			
	Residual	0.924			
	Intercept (PLS with Subheadings)	2.978	0.115	<.001	
	No Subheadings	-0.100	0.054	.066	-0.104
	OSA	-0.359	0.054	<.001	-0.374
	OSA – No Subheadings	-0.259	0.053	<.001	-0.270
	Marginal R^2/Conditional R^2	.016/.346			
Boredom[1]	Random Effect Participant	0.248			
	Random Effect Study	0.063			
	Residual	0.620			
	Intercept (PLS with Subheadings)	1.645	0.083	<.001	
	No Subheadings	0.120	0.045	.007	0.152
	OSA	0.136	0.044	.002	0.173
	OSA – No Subheadings	0.017	0.043	.697	0.021
	Marginal R^2/Conditional R^2	.004/.336			
Confusion	Random Effect Participant	0.198			
	Random Effect Study	0.125			
	Residual	0.643			
	Intercept (PLS with Subheadings)	1.553	0.113	<.001	
	No Subheadings	0.190	0.045	<.001	0.237
	OSA	0.458	0.045	<.001	0.571
	OSA – No Subheadings	0.268	0.044	<.001	0.334
	Marginal R^2/Conditional R^2	.035/.358			
Frustration	Random Effect Participant	0.234			
	Random Effect Study	0.036			
	Residual	0.546			
	Intercept (PLS with Subheadings)	1.391	0.073	<.001	
	No Subheadings	0.142	0.042	<.001	0.193
	OSA	0.228	0.042	<.001	0.309
	OSA – No Subheadings	0.086	0.041	.035	0.116
	Marginal R^2/Conditional R^2	.011/.339			

Note: EST = estimates for variances of residuals and random effects, unstandardized regression weights and marginal/conditional R^2 (see Nakagawa et al., 2013); *SE* = standard error, *p* = *p*-value of two-tailed significance test, *STD. EST* = standardized regression weights, PLS = plain language summary, OSA = ordinary scientific abstract; all epistemic emotions were measured on 1 to 5 Likert scales

[1]We employed ML optimization (instead of REML), as we encountered convergence issues using the REML criterion.

headings and 0.571 with subheadings), whereas effects on boredom were considerably smaller (0.021 without subheadings and 0.173 with subheadings).

Full Text Access

Overall, participants requested links for 335 out of 1,947 presented studies (17.21 %). This rate was lower when ordinary scientific abstracts were presented (15.21 %) compared to plain language summaries with subheadings (18.23 %) or plain language summaries without subheadings (18.25 %). Inferential analyses revealed that differences between both plain language summary types (*z* = - 0.52, *p* = .605), as well as differences between plain language summaries without subheadings and ordinary scientific abstracts were non-significant (*z* = -1.73, *p* = .083), while differences between plain language summaries with subheadings and ordinary scientific abstracts were significant (*z* = -2.20, *p* = .028). In oth-

er words, subjects were more likely to request article full texts when they read plain language summaries with subheadings compared to ordinary scientific abstracts.

The Role of Justification Beliefs and English Proficiency

LR tests (see Table 6) showed that overall significant individual differences existed in perceived comprehensibility. Moreover, individuals differed in the effects ordinary scientific abstracts had compared to plain language summaries (with or without subheadings). English proficiency had no significant effect on these interindividual differences. Subjects with higher English proficiency levels perceived summaries to be more comprehensible for all types of presented research summaries ($b = .17$, $p = .015$). Justification beliefs predicted these interindividual differences in the effects ordinary scientific abstracts had compared to plain language summaries (with or without subheadings). More specifically, for plain language summaries with and without subheadings effects of justification by authority ($b = .14$, $p = .072$) and justification by multiple sources ($b = .15$, $p = .053$) (closely) failed to reach significance, whereas the effect of beliefs in personal justification on perceived comprehensibility was significant and negative ($b = -.18$, $p = .020$). Interestingly, these effects reversed (at least on a descriptive level) for ordinary scientific abstracts, where a significant interaction was found for justification by multiple sources ($b = -.16$, $p = .035$), but not for personal justification ($b = .14$, $p = .080$) or justification by authority ($b = -.11$, $p = .173$). Since all these effects were of similar size and power issues likely exist for exploratory analyses on reader characteristics (see Discussion), we strongly advocate against overinterpreting the statistical significance of findings here.

LR tests on the knowledge acquisition test revealed significant overall differences in random intercepts (participants differed across research summary types in their ability to answer items correctly), but no individual differences in effects of plain language summaries on knowledge acquisition. While justification beliefs were not able to predict differences in knowledge acquisition (as indicated by a non-significant LR test, see Table 6), higher levels of English proficiency resulted in a higher likelihood of answering correctly to knowledge test items regardless of the type of research summary that was presented ($b = .25$, $p < .001$).

Regarding perceived credibility, perceived ability to evaluate knowledge claims, and perceived ability to make informed decisions, we found individual differences in the effects of plain language summaries (a random slope, see Table 6) for perceived credibility and ability to evaluate. However, neither effects of justification beliefs nor effects of English proficiency could predict this variation in the effects of plain language summaries (all $p > .190$, see Table 6). Yet, subsequent analyses revealed that justification beliefs predicted individual differences across research sum-

mary types in these outcomes (i.e., the variance of the random intercept), while English language proficiency did not (see Table 6). More specifically, beliefs in justification by authority had significant effects on all three outcomes ($b = .30 - .34$, all $p < .01$). Effects of personal justification ($b = -.13 - .09$, $p = .09 - .39$) and justification by multiple sources ($b = -.03 - .0002$, $p = .77 - .98$) were non-significant.

Discussion

Confirmatory Findings

Based on our data, we were able to fully confirm Hypotheses 1 and 2. Individuals perceived plain language summaries without subheadings and plain language summaries with subheadings as more comprehensible than the corresponding ordinary scientific abstracts and answered a higher amount of knowledge test items for both types of plain language summaries correctly. As expected, participants also rated plain language summaries as even more comprehensible when subheadings were included and acquired more knowledge from plain language summaries with subheadings compared to plain language summaries without subheadings. This means that writing plain language summaries and facilitating knowledge acquisition by including subheadings fulfilled the intended purpose of a better understanding and amended accessibility of the presented research.

The practical relevance of these findings is quite high since we did not use artificial materials to illustrate the benefits of plain language summaries but employed published (and thus, "real") plain language summaries and scientific abstracts. These 'gains' in the ecological validity and practical significance of our findings are, however, at the expense of a less strict control of differences between plain language summaries and scientific abstracts. We do not know exactly what differs between those conditions—or what drives this effect—as the authors wrote both types of research summaries themselves based on a short guidance provided by JSPP. Adherence to such guidance has been shown to be rather low (see Kadic et al., 2016) and different authors might differ fundamentally in their understanding of how to write plain language summaries. Differences between plain language summaries and scientific abstracts may have existed, for example, in linguistic characteristics (e.g., usage of technical terms), formal characteristics (i.e., sentence length) or content (e.g., a more comprehensive introduction to the background of the research question at hand or qualitative descriptions of statistical results). Analyzing in what way exactly published plain language summaries and scientific abstracts differ regarding these aspects is an intriguing research question of its own that was, however, beyond the scope of this study.

Table 6. Likelihood ratio tests on individual differences and the role of reader characteristics (i.e., justification beliefs, English proficiency).

	Likelihood Ratio Test	Comprehensibility			Knowledge Acquisition			Credibility			Ability to Evaluate			Ability to Make a Decision		
	Added Parameters	χ^2	df	p	χ^2	df	p	χ^2	df	p	χ^2	df	p	χ^2	df	p
1	Random Intercept Participant	282.67	1	<.001	39.10	1	<.001	527.41	1	<.001	1014.50	1	<.001	789.07	1	<.001
2	Random Slope OSA by Participant	8.87	2	.012	_[1]	_[1]	_[1]	8.07	2	.018	17.90	2	<.001	_[1]	_[1]	_[1]
3	Random Slope No Subheadings by Participant	0.52	2	.771	0.84	2	.656	3.01	2	.222	_[1]	_[1]	_[1]	3.55	2	.170
4	Justification Beliefs on Random Intercept	6.24	3	.101	4.37	3	.224	25.99	3	<.001	9.56	3	.023	11.20	3	.011
5	Justification Beliefs on Random Slope OSA	15.45	6	.017	_[2]	_[2]	_[2]	2.04	3	.564	3.80	3	.284	_[2]	_[2]	_[2]
6	English Proficiency on Random Intercept	6.01	1	.014	24.75	1	<.001	1.548	1	.214	0.01	1	.933	0.00	1	.998
7	English Proficiency on Random Slope OSA	0.56	1	.455	_[2]	_[2]	_[2]	3.31	2	.191	0.03	2	.990	_[2]	_[2]	_[2]

Note: OSA = ordinary scientific abstract; χ^2 = chi-square statistic; *df* = degrees of freedom, *p* = *p*-value of the chi-squared test. In column 1, information on the parameters that were tested in the corresponding LR test are provided. First, tests on the significance of random intercepts (row 1) and random slopes for both experimental conditions (i.e., with plain language summary with subheadings as a reference category) were conducted (row 2, 3). If these tests on random effects were significant, we conducted subsequent LR tests to examine if epistemic justification beliefs (row 4, 5) or English proficiency (row 6, 7) were able to explain variation in random intercepts (row 4, 6) or random slopes for ordinary scientific abstracts (row 5, 7). Since all LR tests on random slopes for the effect of plain language summaries without subheadings were non-significant (row 3), the corresponding tests with epistemic justification beliefs and English proficiency were not conducted and these rows are therefore omitted from this table.

[1]No test because model estimation problems occurred.

[2]No test because random slope was non-significant in preceding analyses.

Furthermore, findings on H3a - H5a indicate that the easiness effect of science popularization showed for differences between plain language summaries with subheadings and ordinary scientific abstracts: Our participants found information in plain language summaries with subheadings to be more credible and relied more confidently on their findings. Even though achieving a higher trust in psychological findings might be in line with the aims of plain language summaries, this also implies that presenting individuals with plain language summaries instead of ordinary scientific abstracts might result in an overinterpretation of research findings. Taking into account results of the Open Science Collaboration (2015), which cast doubt on the replicability of various individual psychological findings, such an overinterpretation could be construed as a dangerous side effect of providing laypersons with plain language summaries. It is even conceivable that plain language summaries could be misused to serve not only the purpose of making high quality science more accessible, but also to provide doubtful information with a "scientific" label in a way that is easily understandable to most individuals. In line with these notions, the APA guidance for Translational Abstracts and Public Significance Statements stresses that "it is imperative that you [the authors] do not overstate or oversimplify your findings or conclusions." (American Psychological Association, n.d.).

The authors' ability to transparently communicate their findings without exaggerations is, therefore, of utmost importance when it comes to justifying this increased trust in plain language summaries. We argue that psychological science needs to establish (more detailed) guidance and best practices on how to communicate and report the quality of evidence to lay audiences to support researchers in this task. Journals might further improve the adequateness of the claims provided in plain language summaries by making them part of the review process. This is already done in some journals (e.g., Diabetes Therapy, n.d.). A different approach at the journal level is to have independent writers compose plain language summaries (e.g., King et al., 2017).

Taking a more positive stance, one might also argue that participants were simply able to more strongly appreciate the quality of the studies we presented when they received plain language summaries with subheadings. This would imply that trust can be already earned by means of providing plain language summaries. As a theoretical underpinning for this argument, we draw on the theoretical distinction between first-hand and second-hand evaluations of knowledge claims put forward by Bromme et al. (2010). First-hand evaluations focus on directly evaluating the veracity of knowledge claims. For example, one might investigate the logical coherence of a study's argument, evaluate its methodological approach (e.g., design, sample size, etc.), or compare it with other studies on the same subject (Bromme et al., 2010). In contrast, second-hand evaluations do not focus on the knowledge claim itself, but on the credibility of the corresponding source (Bromme et al., 2010). One might, for example, check whether a certain claim was brought forward by a known conspiracy theorist or by a renowned scientist—and choose to only believe the claim in the latter case. We argue that plain language summaries may be regarded as a way to facilitate first-hand evaluations

by helping individuals to overcome barriers associated with technical terminology (cf. Nunn & Pinfield, 2014). In fact, understanding what was (not) done in a study, enables laypeople to make, at least to a certain extent, an informed decision about its respective quality. Current findings by Hoogeveen et al. (2020) support this view as they suggest that laypersons are indeed able to evaluate the credibility of research findings (to a certain extent) if they are presented in plain language. In their study, based on non-technical summaries, laypersons were able to predict how likely study findings were to be replicated and also took adequately into account (if provided) information on the strength of evidence. Since our participants reported an increased ability to evaluate the corresponding studies for plain language summaries with subheadings (H4a), more first-hand evaluations will likely have taken place if this type of research summary was presented. If we assume that most peer-reviewed journals, such as JSPP, publish, at least to a large extent, high-quality research, first-hand evaluations of our participants were likely to be mostly positive. Consequently, such evaluations would result in a higher perceived credibility of the corresponding findings. Naturally, to enable laypersons to make such (basic) first-hand evaluations on the quality of the research at hand, researchers need more guidance on how to present and discuss their research as transparently and accurately as possible in plain language summaries. We may also provide laypeople with decision aids that further support them in their first-hand evaluations. We argue that open science badges might be a viable option to provide laypeople with this information as adherence to open science practices is considered to be an indicator of scientific rigor (Prager et al., 2019). In fact, open science practices reduce researchers' degrees of freedom, which effectively reduces questionable research practices such as HARKing (hypothesizing after the results are known; Chambers, 2019), for example.

Contrary to our expectations, the easiness effect did not show for plain language summaries without subheadings compared to ordinary scientific abstracts. The corresponding differences did not emerge in our confirmatory analyses (H3b - H5b). Moreover, further exploratory analyses also revealed significant differences between plain language summaries without and with subheadings. This might indicate that perceived 'text easiness' results from a complex interplay of text length, text comprehensibility, and text structure. Future studies on the role of headlines with regard to the easiness effect of science popularization might draw on theoretical frameworks on information processing, such as the heuristic-systematic model (Chen & Chaiken, 1999) to decompose the underlying psychological mechanisms (e.g., subheadings might serve as heuristic cues). In this context, one might also examine how different types of subheadings affect information processing for plain language summaries (as has been done for different types of news headlines, Scacco & Muddiman, 2020).

Exploratory Findings

Exploratory analyses indicated that individuals experienced less negative emotions when reading plain language summaries and were more likely to request corresponding

full texts than when reading ordinary scientific abstracts. However, our participants knew beforehand that research on social and political psychology was the topic of our study, which is why self-selection might have occurred. This might also lead to more positive emotions in general and a higher amount of requested links (even though the proportion was still quite low) compared to a general population sample. Therefore, the extent to which these findings can be transferred to other populations remains an open question for future research.

Nonetheless, these differences in full text access might point towards the practical relevance of providing plain language summaries. Providing plain language summaries did not only influence response patterns in self-report measures, but also how individuals dealt with and accessed research findings. On the other hand, one might ask what happens when individuals actually access the corresponding scientific full texts after being 'lured' to them by easily comprehensible plain language summaries. Would they become even more frustrated? Is accessing full texts the kind of behavior that we want to promote by providing plain language summaries in the first place? As can be seen here, the future role of plain language summaries remains to be determined and largely depends on the efforts publishers, editors and authors of research articles are willing to make.

Furthermore, we found a main effect of English proficiency—but no interactions with the type of research summary—on comprehensibility and knowledge acquisition. However, good English language skills were required to participate in our study—the general population of countries in which English is not the first language will almost certainly *not* possess these language skills (see also Cochrane Collaboration, 2019). Although these findings are exploratory and should therefore be interpreted with adequate caution, they point towards English language proficiency as a major obstacle for comprehending any kind of research summary in non-native speakers. Consequently, making science accessible to the non-English speaking public requires additional measures, such as translating ordinary scientific abstracts and plain language summaries. This underlines the importance of early attempts to solve this issue, which are fortunately currently undertaken in various areas (e.g., by Cochrane in the fields of medicine and public health).

Limitations and Future Directions

This study has some limitations. Most importantly, we analyzed only twelve ordinary scientific abstracts and plain language summaries of *one* journal. Although it seems reasonable to assume that our results might apply to other journals in psychology and the social sciences as well (especially if they use similar or identical author guidelines), the generalizability of our findings remains a question for future research.

How should the non-significant findings in our study be interpreted? Since our confirmatory hypotheses were preregistered and the power calculation for these hypotheses was based on a Cohen's *d* of *.5*, we are reasonably sure that we did not fail to find any medium-sized effects (at text-level). Especially evidence on the absence of effects for differences between the plain language summaries without subheadings and the ordinary scientific abstracts condition (H3b-H5b) seems quite strong. On three related outcomes, all findings were clearly non-significant and the data (distribution) on these outcomes appears very similar (see [Figure 3](#)). Notwithstanding, given that we only report evidence of one single study, we cannot fully rule out that (possibly smaller) effects may exist at the population-level, which we were unable to detect. We are less confident regarding the robustness of our non-significant (and significant) exploratory findings on participant-level covariates (i.e., justification beliefs and English language proficiency) as evidence of absence. This is due to the fact that the calculations of our power analyses did not explicitly target such higher-level effects in mixed models. Therefore, effects might be estimated with less precision on this level and should be interpreted with great caution.

Furthermore, we forced participants to read specific research summaries in our study (instead of letting them choose articles based on personal preferences[6]). On the one hand, one might argue that observed differences in outcomes, for example, in full text access rates, might be even larger when individuals do an information search (e.g., on Google Scholar) out of their free will and on a self-selected topic with an information need of their own. For instance, they might be more likely to access full texts of plain language summaries as they are more confident that this comprehensibly presented information will address their information need. On the other hand, this information need might already be fully addressed by well-comprehensive plain language summaries, whereas ordinary scientific abstracts might evoke a need for more information due to reduced understanding. In this case, reading a plain language summary would lead to lower full text access rates in 'real life'. As can be seen here, future research needs to pay closer attention to individuals' motivation for reading plain language summaries by examining different scenarios of information access in order to shed more light on this issue. In the same vein, one might argue that providing easily comprehensible information alone does not necessarily ensure that individuals will want to understand this information (see Brossard & Lewenstein, 2009). What our study shows, however, is that plain language summaries support individuals in correctly understanding scientific information—at least when they are explicitly required to engage with the information. The will to engage with the information, in contrast, should be investigated in future studies.

Additionally, our study population consisted of university students, of which a large proportion (38.55 %) studied psychology. Even though these students are clearly non-

---

6 This effect might, however, be counteracted by the fact that our participants knew beforehand that research on social and political psychology was the topic of our study. As a consequence, self-selection might have occurred regarding personal interest in the topic.

scientists (and therefore a target group of plain language summaries), they are to some extent familiar with summaries of empirical research. However, practitioners, even those holding a degree, are also commonly considered to be "laypersons" with regard to receiving and processing scientific information (e.g., Barnes & Patrick, 2019). Nonetheless, for other groups of laypersons that are less familiar with reading research findings (e.g., the public), two scenarios are conceivable: (a) they might struggle to understand plain language summaries which were comprehensible for our population, leading to practically no differences between ordinary scientific abstracts and plain language summaries, or (b) the difference between plain language summaries and ordinary scientific abstracts might even be larger in other audiences since ordinary scientific abstracts are even harder to understand for this group of individuals. It is, therefore, in our opinion, of utmost importance to examine how familiarity with research as well as the ability to understand and evaluate research ("research literacy", Beaudry & Miller, 2016) may influence effects of plain language summaries by comparing effects in different groups of laypersons.

Consequently, we concede that the set of covariates that was included in our study is by no means comprehensive. Besides epistemic justification beliefs, English language proficiency, and research literacy, other reader characteristics for example, need for cognitive closure, trust in science, or prior topic knowledge might also influence how individuals deal with different types of research summaries. Likewise, we do not suggest that epistemic emotions are the only process-related variables which should be considered when it comes to understanding how exactly plain language summaries work (e.g., one might examine different types of epistemic aims that readers pursue when reading research summaries, Chinn et al., 2011).

Conclusion

This study demonstrated that providing individuals with plain language summaries is a promising approach for communicating research findings to a broader audience. First, we showed that laypeople perceived plain language summaries as more comprehensible compared to ordinary scientific abstracts, but also that laypeople actually understood the corresponding information more correctly when presented in plain language summaries compared to ordinary scientific abstracts. Second, in line with the easiness effect of science popularization, we found that individuals perceived plain language summaries with subheadings as more credible and had higher confidence in their ability to make decisions based on those plain language summaries. If this increased perceived credibility is always necessarily a good thing is, however, debatable—at least if there are no effective measures in place to ensure that the claims put forward in plain language summaries are actually warrant-

ed by the empirical evidence of the corresponding study. Third, individuals experienced less negative and more positive emotions when reading plain language summaries instead of ordinary scientific abstracts and were also more likely to access the corresponding full texts. In sum, there are many good theoretical and practical reasons for providing plain language summaries and our empirical research further strengthens these arguments by demonstrating that plain language summaries actually work in practice for psychological research. Arguments for not including plain language summaries in journals should become increasingly hard to find. Thus, we strongly encourage the scientific community—and especially journal editors and publishers in fields with high societal interest—to implement them.

REFERENCES

Alderdice, F., McNeill, J., Lasserson, T., Beller, E., Carroll, M., Hundley, V., Sunderland, J., Devane, D., Noyes, J., Key, S., Norris, S., Wyn-Davies, J., & Clarke, M. (2016). Do Cochrane summaries help student midwives understand the findings of Cochrane systematic reviews: The BRIEF randomised trial. *Systematic Reviews*, *5*(40). https://doi.org/10.1186/s13643-016-0214-8

American Psychological Association. (n.d.). *Guidance for translational abstracts and public significance statements: Demonstrating the public significance of research*. https://www.apa.org/pubs/journals/resources/translational-messages

Barnes, A., & Patrick, S. (2019). Lay summaries of clinical study results: An overview. *Pharmaceutical Medicine*, *33*(4), 261–268. https://doi.org/10.1007/s40290-019-00285-0

Barzilai, S., & Strømsø, H. I. (2018). Individual differences in multiple document comprehension. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.), *Handbook of multiple source use* (pp. 99–116). Routledge.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Beaudry, J. S., & Miller, L. (2016). *Research literacy: A primer for understanding and using research*. Guilford Publications.

Behmen, D., Marušić, A., & Puljak, L. (2019). Capacity building for knowledge translation: A survey about the characteristics and motivation of volunteer translators of Cochrane plain language summaries. *Journal of Evidence-Based Medicine*, *12*(2), 147–154. https://doi.org/10.1111/jebm.12345

Bråten, I., Ferguson, L. E., Strømsø, H. I., & Anmarkrud, Ø. (2013). Justification beliefs and multiple-documents comprehension. *European Journal of Psychology of Education*, *28*(3), 879–902. https://doi.org/10.1007/s10212-012-0145-2

Bromme, R., Kienhues, D., & Porsch, T. (2010). Who knows what and who can we believe? Epistemological beliefs are beliefs about knowledge (mostly) to be attained from others. In L. D. Bendixen & F. C. Feucht (Eds.), *Personal epistemology in the classroom: Theory, research, and implications for practice* (pp. 163–193). Cambridge University Press.

Brossard, D., & Lewenstein, B. V. (2009). A critical appraisal of models of public understanding of science: Using practice to inform theory. In *Communicating science* (pp. 25–53). Routledge.

Buljan, I., Malički, M., Wager, E., Puljak, L., Hren, D., Kellie, F., West, H., Alfirević, Ž., & Marušić, A. (2018). No difference in knowledge obtained from infographic or plain language summary of a Cochrane systematic review: Three randomized controlled trials. *Journal of Clinical Epidemiology*, *97*, 86–94. https://doi.org/10.1016/j.jclinepi.2017.12.003

Callaway, E. (2011). Report finds massive fraud at Dutch universities. *Nature*, *479*(7371), 15. https://doi.org/10.1038/479015a

Carvalho, F. A., Elkins, M. R., Franco, M. R., & Pinto, R. Z. (2019). Are plain-language summaries included in published reports of evidence about physiotherapy interventions? Analysis of 4421 randomised trials, systematic reviews and guidelines on the Physiotherapy Evidence Database (PEDro). *Physiotherapy*, *105*(3), 354–361. https://doi.org/10.1016/j.physio.2018.11.003

Chambers, C. (2019). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton University Press.

Chen, S., & Chaiken, S. (1999). The heuristic-systematic. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73–96). Guilford Press.

Chinn, C. A., Buckland, L. A., & Samarapungavan, A. L. A. (2011). Expanding the dimensions of epistemic cognition: Arguments from philosophy and psychology. *Educational Psychologist*, *46*(3), 141–167. https://doi.org/10.1080/00461520.2011.587722

Clayman, M. L., Manganello, J. A., Viswanath, K., Hesse, B. W., & Arora, N. K. (2010). Providing health messages to Hispanics/Latinos: Understanding the importance of language, trust in health information sources, and media use. *Journal of Health Communication*, *15*(3), 252–263. https://doi.org/10.1080/10810730.2010.522697

Cochrane Collaboration. (2019). *Translated Cochrane evidence: Bringing you Cochrane evidence in 15 different languages*. https://www.cochrane.org/news/translated-cochrane-evidence

Cochrane Methods. (2013). *Methodological Expectations of Cochrane Intervention Reviews (MECIR): Standards for the reporting of plain language summaries in new Cochrane Intervention Reviews 2013*. https://methods.cochrane.org/sites/default/files/public/uploads/pleacs_2019.pdf

Cochrane Norway. (2019). *How to write a plain language summary of a Cochrane intervention review*. https://www.cochrane.no/sites/cochrane.no/files/public/uploads/how_to_write_a_cochrane_pls_12th_february_2019.pdf

Diabetes Therapy. (n.d.). *Submission guidelines: Guidelines for digital features and plain language summaries*. https://www.springer.com/journal/13300/submission-guidelines

European Medicines Association. (2019). *Clinical trial regulation*. https://www.ema.europa.eu/en/human-regulatory/research-development/clinical-trials/clinical-trial-regulation

Expert group on clinical trials for the implementation of Regulation (EU) No 536/2014. (n.d.). *Summaries of clinical trial results for laypersons: Recommendations of the expert group on clinical trials for the implementation of regulation (EU) No 536/2014 on clinical trials on medicinal products for human use*. https://ec.europa.eu/health/sites/health/files/files/clinicaltrials/2016_06_pc_guidelines/gl_3_consult.pdf

FitzGibbon, H., King, K., Piano, C., Wilk, C., & Gaskarth, M. (2020). Where are biomedical research plain-language summaries? *Health Science Reports*, *3*(3), e175. https://doi.org/10.1002/hsr2.175

Glenton, C., Santesso, N., Rosenbaum, S., Nilsen, E. S., Rader, T., Ciapponi, A., & Dilkes, H. (2010). Presenting the results of Cochrane systematic reviews to a consumer audience: A qualitative study. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *30*(5), 566–577. https://doi.org/10.1177/0272989x10375853

Grand, A., Wilkinson, C., Bultitude, K., & Winfield, A. F. T. (2012). Open science: A new "Trust Technology"? *Science Communication*, *34*(5), 679–689. https://doi.org/10.1177/1075547012443021

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, *44*(3), 687–691. https://doi.org/10.1016/j.jesp.2007.04.005

Hauck, S. A., II. (2019). Sharing planetary science in plain language. *Journal of Geophysical Research: Planets*, *124*(10), 2462–2464. https://doi.org/10.1029/2019je006152

Hesse, B. W. (2018). Can psychology walk the walk of open science? *The American Psychologist*, *73*(2), 126–137. https://doi.org/10.1037/amp0000197

Hoogeveen, S., Sarafoglou, A., & Wagenmakers, E.-J. (2020). Laypeople can predict which social-science studies will be replicated successfully. *Advances in Methods and Practices in Psychological Science*, *1*, 251524592091966. https://doi.org/10.1177/2515245920919667

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*(1), 601–625. https://doi.org/10.1146/annurev-psych-122414-033702

Kadic, A. J., Fidahic, M., Vujcic, M., Saric, F., Propadalo, I., Marelja, I., Dosenovic, S., & Puljak, L. (2016). Cochrane plain language summaries are highly heterogeneous with low adherence to the standards. *BMC Medical Research Methodology*, *16*(1), 444. https://doi.org/10.1186/s12874-016-0162-y

Kaslow, N. J. (2015). Translating psychological science to the public. *American Psychologist*, *70*(5), 361–371. https://doi.org/10.1037/a0039448

King, S. R., Pewsey, E., & Shailes, S. (2017). An inside guide to eLife digests. *ELife*, *6*. https://doi.org/10.7554/elife.25410

Klein, O., Hardwicke, T. E., Aust, F., Breuer, J., Danielsson, H., Hofelich Mohr, A., IJzerman, H., Nilsonne, G., Vanpaemel, W., & Frank, M. C. (2018). A practical guide for transparency in psychological science. *Collabra: Psychology*, *4*(1), 20. https://doi.org/10.1525/collabra.158

Klopp, E., & Stark, R. (2016). *Entwicklung eines Fragebogens zur Erfassung domänenübergreifender epistemologischer Überzeugungen [Development of a domain-general epistemological beliefs questionnaire]* [Unpublished manuscript]. Department of Educational Science, Saarland University, Saarbrücken, Germany.

Kuehne, L. M., & Olden, J. D. (2015). Opinion: Lay summaries needed to enhance science communication. *Proceedings of the National Academy of Sciences*, *112*(12), 3585–3586. https://doi.org/10.1073/pnas.1500882112

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Liepmann, D., Heinitz, K., Nettelnstroth, W., & Smolka, S. (2013). *E-PA: Englischtest für die Personalauswahl* (1. Auflage). Hogrefe.

Maguire, L. K., & Clarke, M. (2014). How much do you need: A randomised experiment of whether readers can understand the key messages from summaries of Cochrane Reviews without reading the full review. *Journal of the Royal Society of Medicine*, *107*(11), 444–449. https://doi.org/10.1177/0141076814546710

Mc Laughlin, G. H. (1969). SMOG grading - a new readability formula. *Journal of Reading*, *12*(8), 639–646. http://www.jstor.org/stable/40011226

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, *15*, 1–20. https://doi.org/10.1080/00461520.2017.1421465

Nakagawa, S., Schielzeth, H., & O'Hara, R. B. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nunn, E., & Pinfield, S. (2014). Lay summaries of open access journal articles: Engaging with the general public on medical research. *Learned Publishing*, *27*(3), 173–184. https://doi.org/10.1087/20140303

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The epistemically-related emotion scales. *Cognition and Emotion*, *31*(6), 1268–1276. https://doi.org/10.1080/02699931.2016.1204989

Pittinsky, T. L. (2015). America's crisis of faith in science. *Science*, *348*(6234), 511. https://doi.org/10.1126/science.348.6234.511-a

Prager, E. M., Chambers, K. E., Plotkin, J. L., McArthur, D. L., Bandrowski, A. E., Bansal, N., Martone, M. E., Bergstrom, H. C., Bespalov, A., & Graf, C. (2019). Improving transparency and scientific rigor in academic publishing. *Journal of Neuroscience Research*, *97*(4), 377–390. https://doi.org/10.1002/jnr.24340

R Core Team. (2019). *R: A language and environment for statistical computing [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Rakedzon, T., Segev, E., Chapnik, N., Yosef, R., & Baram-Tsabari, A. (2017). Automatic jargon identifier for scientists engaging with the public and science communication educators. *PloS ONE*, *12*(8), e0181742. https://doi.org/10.1371/journal.pone.0181742

Santesso, N., Rader, T., Nilsen, E. S., Glenton, C., Rosenbaum, S., Ciapponi, A., Moja, L., Pardo, J. P., Zhou, Q., & Schünemann, H. J. (2015). A summary to communicate evidence from systematic reviews to the public improved understanding and accessibility of information: A randomized controlled trial. *Journal of Clinical Epidemiology*, *68*(2), 182–190. https://doi.org/10.1016/j.jclinepi.2014.04.009

Scacco, J. M., & Muddiman, A. (2020). The curiosity effect: Information seeking in the contemporary news environment. *New Media & Society*, *22*(3), 429–448. https://doi.org/10.1177/1461444819863408

Scharrer, L., Britt, M. A., Stadtler, M., & Bromme, R. (2013). Easy to understand but difficult to decide: Information comprehensibility and controversiality affect laypeople's science-based decisions. *Discourse Processes*, *50*(6), 361–387. https://doi.org/10.1080/0163853x.2013.813835

Scharrer, L., Bromme, R., Britt, M. A., & Stadtler, M. (2012). The seduction of easiness: How science depictions influence laypeople's reliance on their own evaluation of scientific information. *Learning and Instruction*, *22*(3), 231–243. https://doi.org/10.1016/j.learninstruc.2011.11.004

Scharrer, L., Rupieper, Y., Stadtler, M., & Bromme, R. (2017). When science becomes too easy: Science popularization inclines laypeople to underrate their dependence on experts. *Public Understanding of Science*, *26*(8), 1003–1018. https://doi.org/10.1177/0963662516680311

Scharrer, L., Stadtler, M., & Bromme, R. (2014). You'd better ask an expert: Mitigating the comprehensibility effect on laypeople's decisions about science-based knowledge claims. *Applied Cognitive Psychology*, *28*(4), 465–471. https://doi.org/10.1002/acp.3018

Scharrer, L., Stadtler, M., & Bromme, R. (2019). Judging scientific information: Does source evaluation prevent the seductive effect of text easiness? *Learning and Instruction*, *63*, 101215. https://doi.org/10.1016/j.learninstruc.2019.101215

Seidel, T., Mok, S. Y., Hetmanek, A., & Knogler, M. (2017). Meta-Analysen zur Unterrichtsforschung und ihr Beitrag für die Realisierung eines Clearing House Unterricht für die Lehrerbildung [Meta-analyses on teaching effectiveness and their contribution to the realization of a Clearing House Unterricht for teacher education]. *Zeitschrift für Bildungsforschung, 7*(3), 311–325. https://doi.org/10.1007/s35834-017-0191-6

Shailes, S. (2017). Plain-Language Summaries of research: Something for everyone. *ELife, 6*(e25411), 1–5. https://doi.org/10.7554/elife.25411

Stricker, J., Chasiotis, A., Kerwer, M., & Günther, A. (2020). Scientific abstracts and plain language summaries in psychology: A comparison based on readability indices. *PloS ONE, 15*(4), e0231160. https://doi.org/10.1371/journal.pone.0231160

Stricker, J., & Günther, A. (2019). Scientific misconduct in psychology. *Zeitschrift für Psychologie, 227*(1), 53–63. https://doi.org/10.1027/2151-2604/a000356

Strømsø, H., & Kammerer, Y. (2016). Epistemic cognition and reading for understanding in the internet age. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 230–246). Routledge.

Świątkowski, W., & Dompnier, B. (2017). Replicability crisis in social psychology: Looking at the past to find new pathways for the future. *International Review of Social Psychology, 30*(1), 111. https://doi.org/10.5334/irsp.66

Thomm, E., & Bromme, R. (2012). "It should at least seem scientific!" Textual features of "scientificness" and their impact on lay assessments of online information. *Science Education, 96*(2), 187–211. https://doi.org/10.1002/sce.20480

Westfall, J. (2016, March 25). *Five different "Cohen's d" statistics for within-subject designs.* http://jakewestfall.org/blog/index.php/2016/03/25/five-different-cohens-d-statistics-for-within-subject-designs/

SUPPLEMENTARY MATERIALS

Peer Review History

Download: https://collabra.scholasticahq.com/article/18898-straight-from-the-scientist-s-mouth-plain-language-summaries-promote-laypeople-s-comprehension-and-knowledge-acquisition-when-reading-about-individual-research-findings-in-psychology/attachment/50281.docx?auth_token=Mb9l84Wkq-UG60oJzneG