

# Translating the evidence of psychological meta-analyses into plain language - Study 2<sup>1</sup>

Martin Kerwer<sup>a</sup>, Marlene Stoll<sup>ab</sup>, Mark Jonas<sup>a</sup>, Gesa Benz<sup>a</sup> and Anita Chasiotis<sup>a</sup>

<sup>a</sup> Leibniz Institute for Psychology (ZPID)

<sup>b</sup> Leibniz Institute for Resilience Research (LIR)

## Background

Scientific articles are often hard to understand for non-scientists. This is partly due to the fact that such articles use technical terms and statistical concepts without explaining them, since they are considered to be common knowledge for their scientific audience. Additionally, researchers may tend to pay more attention to the scientific relevance of their findings and less attention to their practical relevance to the broader public when discussing them in scientific publications. To address these problems, plain language summaries (PLS)—easily comprehensible research summaries that complement scientific abstracts—were introduced. Empirically validated guidelines on how to write such PLS focus, however, mostly on biomedical research, for example the Cochrane guidelines (Jelicic Kadic et al., 2016; Santesso et al., 2008). With the aim to develop empirically validated guidelines for writing PLS in psychology, the Leibniz Institute for Psychology (ZPID) started the project “PLan Psy”. This project aims at finding a standardized way to translate findings of psychological meta-analyses into PLS. This protocol outlines the design of the second empirical study of PLan Psy, which will, by applying an experimental design, investigate psychology-specific aspects of how to optimally communicate meta-analytical evidence to laypersons.

## Research Question

We will conduct an experimental study that systematically varies characteristics of PLS of psychological meta-analyses (i.e., approaches for communicating complex meta-analytic designs, approaches for communicating results, approaches for formal text structuring, see

---

<sup>1</sup> *Note:* This study protocol is strongly based on the study protocol of the first empirical study of project PLan Psy and, where appropriate, text passages were adopted from this protocol without modification. Citation of the first study protocol:

Kerwer, M., Stoll, M., & Chasiotis, A. (2021). Translating the evidence of psychological meta-analyses into plain language. PsychArchives. <https://doi.org/10.23668/PSYCHARCHIVES.4471>

Independent Variables) with respect to relevant outcomes (i.e., perceived accessibility, perceived understanding, content-related knowledge, perceived empowerment, see Dependent Variables). Moreover, we will examine to what extent these effects can be generalized to more or less complex PLS (see Independent Variables).

## Hypotheses

In hypotheses 1 and 2, we will investigate how characteristics of PLS that relate to communicating results, i.e. the specific approach for explaining statistical terms, affect the aforementioned outcome variables. More specifically, we make the following assumptions:

- H1a: (Perceived) Accessibility of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a qualitative statement *without* a glossary.
- H1b: (Perceived) Understanding of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a qualitative statement *without* a glossary.
- H1c: Content-related knowledge will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a qualitative statement *without* a glossary.
- H1d: (Perceived) Empowerment of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a qualitative statement *without* a glossary.
- H2a: (Perceived) Accessibility of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a glossary *without* a qualitative statement.
- H2b: (Perceived) Understanding of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a glossary *without* a qualitative statement.
- H2c: Content-related knowledge of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS compared to reporting the effect size *and* providing a glossary *without* a qualitative statement.
- H2d: (Perceived) Empowerment of PLS will be higher if an effect size is reported *and* a glossary that explains statistical terms is provided after the PLS *and* a qualitative statement on the interpretation of the effect size is provided within the text of the PLS

compared to reporting the effect size *and* providing a glossary *without* a qualitative statement.

We will also test for differences between reporting the effect size and providing a glossary *or* a qualitative statement but have no expectation on the direction of effects for this comparison. Moreover, we will compare these three conditions (reporting the effect size with a glossary and/or qualitative statement) to a condition in which no effect size is reported but only a qualitative statement on the size of the effect is provided but have no expectation on the direction of effects for these comparisons.

In hypothesis 3, we will investigate how explaining complex meta-analytic designs affects the following outcome variables: (Perceived) Accessibility (H3a), (Perceived) Understanding (H3b), Content-related Knowledge (H3c), (Perceived) Empowerment (H3d). We have no expectation on the direction of effects for these comparisons.

In hypothesis 4, we will investigate if formally structuring the text—for example, by means of bullet points—affects the aforementioned outcome variables. More specifically, we make the following assumptions:

- H4a: (Perceived) Accessibility of PLS will be higher if the text is structured compared to a condition where no structuring is conducted.
- H4b: (Perceived) Understanding of PLS will be higher if the text is structured compared to a condition where no structuring is conducted.
- H4c: Content-related knowledge will be higher if the text is structured compared to a condition where no structuring is conducted.
- H4d: (Perceived) Empowerment of PLS will be higher if the text is structured compared to a condition where no structuring is conducted.

## Method

### Sample

A general population sample ( $N = 2004$ , see Sample Size Calculation) will be recruited in this study via the panel provider Respondi.

The following specific quotas apply for age, education level and sex: The same proportion of participants will be recruited for age (18 - 44, 45 or older), sex (men, women) and education level (“Hauptschulabschluss”, “Mittlere Reife”, “Hochschulreife”). Thus, there will be 12 quota conditions and the sample size for each quota condition will be  $2004/12 = 167$ .

Additionally, the following inclusion criteria apply in this study:

- Participants possess German language skills at native speaker level.
- Participants have successfully graduated from school.
- Participants are not currently studying psychology or holding a degree in psychology.
- Participants are interested in psychological research.
  - Item: “Ich interessiere mich für psychologische Forschung.”
  - Criterion: Agreement of “4” or higher on a 1 to 8 rating scale (based on the lowest medium value observed for this item across all quota conditions of the first study of project PLan Psy)

## Design

Figure 1 provides an overview of the study design. This study employs a between-subjects design with three factors that are varied within the presented PLS: Communication of complex meta-analytic designs, approach for explaining statistical terms, formal text structuring (see Independent Variables). Additionally, there will be two study arms (i.e., an additional Independent Variable that is, however, not varied within each PLS). In study arm A, two less complex PLS (i.e., only one simple correlation/mean difference is reported) will be presented. In study arm B, two entirely different more complex PLS (i.e., with meta-moderator analyses or a large number of investigated effects) will be included. Each participant is randomly assigned to one study arm and reads the two PLS of this study arm. Both PLS are randomly assigned to one of the four approaches for explaining statistical terms, one of the two approaches for communicating complex meta-analytic designs and one of the two approaches on formal text structuring (see Independent Variables). This means that the same approach with regard to these independent variables is applied for both PLS. Thus, there are in total 16 ( $4 * 2 * 2$ ) conditions, which are the same in each study arm. All hypotheses will be tested in both study arms separately. We will also explore differences in the effects between both study arms.

All PLS that are employed in this study are given in Appendix A.

## Procedure

The study will be conducted online using the survey software Unipark. At the beginning of the study, participants complete an informed consent form. Thereafter, inclusion criteria are checked and, if they are fulfilled, participants are randomly assigned to one of the 16 experimental conditions for presenting PLS and one study arm (see above). Two PLS will be presented to each participant. The design and content of these PLS depends on the assigned experimental condition, which is the same for both PLS. Participants read each of those two PLS for at least 3 minutes and answer the outcome measures on perceived accessibility, perceived understanding, and perceived empowerment on the same webpage (see Dependent Variables). After reading each PLS, a knowledge test will be administered. Finally, participants complete a user survey (see below). All study procedures were approved by the ethics committee of Trier University.

Screening	Demographics/ Covariates	IV1: Study Arm/Level of Complexity	IV2: Explaining statistical terms	IV3: Communication of complex designs	IV4: Formal Text Structuring	DVs	User Survey/Covariates	
<ul style="list-style-type: none"> <li>German language skills at native speaker level</li> <li>successfully graduated from school</li> <li>not studying or holding a degree in psychology</li> <li>interest in psychological research</li> </ul>	<p>Equally distributed quotas:</p> <ul style="list-style-type: none"> <li>Sex (men/women)</li> <li>Education level ("Hauptschule", "Mittlere Reife", "Hochschulreife")</li> <li>Age (18 to 44/45 or older)</li> </ul> <p>Covariates:</p> <ul style="list-style-type: none"> <li>job in a science related field</li> <li>interest in PLS topics</li> </ul>	Complex PLS	Qualitative Statement	no additional statement	unstructured	<p>For each PLS</p> <ul style="list-style-type: none"> <li>Accessibility</li> <li>Understanding</li> <li>Empowerment</li> <li>Content Knowledge</li> </ul>	<p>Among others questions related to:</p> <ul style="list-style-type: none"> <li>Free text feedback on presented PLS</li> <li>Prior experience with searching for research findings</li> <li>Interest in 20 psychological topic categories</li> <li>Reasons of interest for categories rated as highly interesting</li> <li>Justification Beliefs in the context of psychology</li> <li>Trustworthiness of Scientists</li> </ul>	
				structured				
				additional statement	unstructured			
				structured				
				Effect Size + Qualitative Statement	no additional statement			unstructured
					structured			
			additional statement		unstructured			
			structured					
			Effect Size + Glossary		no additional statement			unstructured
					structured			
				additional statement	unstructured			
				structured				
		Effect Size + Glossary + Qualitative Statement		no additional statement	unstructured			
				structured				
			additional statement	unstructured				
			structured					
			Less complex PLS	Qualitative Statement	no additional statement	unstructured		
					structured			
		additional statement			unstructured			
		structured						
		Effect Size + Qualitative Statement			no additional statement	unstructured		
					structured			
				additional statement	unstructured			
				structured				
Effect Size + Glossary	no additional statement			unstructured				
	structured							
	additional statement	unstructured						
	structured							
	Effect Size + Glossary + Qualitative Statement	no additional statement	unstructured					
		structured						
additional statement		unstructured						
structured								

Figure 1. Graphical outline of the study design.

# Variables

## *Independent Variables*

### Formal text structuring

First evidence suggests that structuring the text of PLS (e.g., by means of bullet points) improves the user experience for information recipients (e.g., Anzinger et al., 2020; Ellen et al., 2014; Raynor et al., 2018). To test if these findings can be transferred to PLS on psychological meta-analyses (and also can be replicated on a conceptual level in a considerably larger and more representative sample), we will compare the following approaches for structuring text blocks of PLS:

- Unstructured text: No formatting of text blocks.
- Structured text: Formatting of text blocks (e.g., by bullet points) according to the rules outlined below (see also Appendix A for structured and unstructured PLS used in our study).

The following structuring subheadings are included in both conditions: “Was war das Ziel der Übersichtsarbeit?”, “Kernaussage der Übersichtsarbeit”, “Wie wurde vorgegangen?”, “Was sind die wichtigsten Ergebnisse?”. The following additional rules apply to the formatting of structured text blocks: The subsection “Was war das Ziel der Übersichtsarbeit?” was further split into two paragraphs which were named “Hintergrund:” and “Forschungsfrage:”. The subsection “Wie wurde vorgegangen?” was split into four paragraphs (“Welche Studien wurden gesucht?”, “Welche Studien wurden gefunden?”, “Was haben die ForscherInnen gemacht?”, “Was wurde untersucht?”). Moreover, bullet points were introduced for reporting investigated variables (“Was wurde untersucht?”) and results (“Was sind die wichtigsten Ergebnisse?”).

### Approach for explaining statistical terms

Laypeople are most likely not familiar with the statistical concepts and especially effect size measures that are employed in scientific publications. In spite of the fact that first studies on communicating statistical evidence to non-scientists exist (e.g., Buljan et al., 2020; Glenton et al., 2010; Santesso et al., 2015), there is still no consensus on how to do this most efficiently in more diverse populations that are not highly educated and, thus, a need for further research in this area. We deem this to be especially true when it comes to the arguably less tangible effect size types frequently employed in psychological science (e.g., correlation coefficients which cannot be easily ‘translated’ to natural frequencies). In this study, we will therefore test four approaches for explaining statistical terms and effect sizes in psychological meta-analyses:

- Qualitative Statement: The effect size is not reported—instead a qualitative statement on the interpretation of its size is provided (e.g., “Der Zusammenhang war mittelstark.”).
- Effect Size + Qualitative Statement: The effect size is reported and a qualitative statement on the interpretation of its size is provided (e.g., “Die Korrelation betrug  $r = 0.40$ . Das ist ein mittelstarker Zusammenhang.”).
- Effect Size + Glossary: No qualitative statement is provided. The effect size is reported in the text. An explanation of the effect size, its boundaries and cut-off values for easing interpretation are provided below the PLS (e.g., What is  $r$  and what value of  $r$  implies that the effect is large?). This explanation below the PLS is not

specifically adapted to the effect size values of the PLS at hand but only to the type of effect size that is reported.

- Effect Size + Glossary + Qualitative Statement: The identical text as in condition *Effect Size + Qualitative Statement* is provided. Additionally, the same glossary as in condition *Effect Size + Glossary* is provided.

It is important to note that we will only vary approaches for explaining ‘statistical terms’ here. In contrast, technical terms that refer to (theoretical) core concepts—*psychological* constructs that are relevant with regard to the specific research questions at hand—are for all PLS explained and/or translated into everyday language.

### Complexity of the PLS

PLS might differ in their complexity for various reasons. They might be complex because the design of the meta-analyses they summarize is complex and they strive to communicate the complexity of these meta-analytic investigations in detail. For example, comparably simple meta-analytic designs might focus on one specific effect size (e.g., one treatment effect or correlation) while more complex meta-analytic designs might investigate how such relationships differ depending on third variables (e.g., via meta-regressions). The complexity of PLS might, however, also increase if meta-analyses report a high amount of effect sizes (e.g., various simple correlation coefficients) and the PLS tries to provide its reader with this high amount of information. In this study, we will investigate whether our findings on approaches for explaining statistical terms and formal text structuring can be replicated for both complex and less complex PLS. For this purpose, PLS on the following meta-analyses were written:

- Complex PLS: One PLS reports 9 correlation coefficients (based on Yule et al., 2019), the other PLS reports meta-moderator analyses (based on Bergmann & Cristia, 2016)
- Less complex PLS: One PLS reports only one correlation (based on Schwalm et al., 2021), the other PLS reports only one mean difference (based on Bucher et al., 2020; in the ZPID’s PsychOpen CAMA system, see Burgard et al., 2021).

All PLS are provided in Appendix A. It is important to note that the complexity of PLS is not varied within each PLS (i.e., the same PLS is not presented with different ‘degrees of complexity’) but that entirely different PLS are presented for different conditions of this independent variable. This approach differs from the approach pursued for the other independent variables which are varied within PLS. Consequently, investigating more and less complex PLS is in our opinion a test of the generalizability of the effects of the other independent variables introduced above.

### Communication of complex meta-analytic designs

In the first study of project P<sub>lan</sub> Psy, we provided laypersons with a statement explaining the concept of meta-analysis that was focused on simple meta-analytic designs (i.e., one synthesized effect). In this study, we will investigate how to explain more complex meta-analytic designs by extending the statement we used in the first study:

- no additional statement on complex meta-analytic designs (i.e., of moderator analyses as a methodological approach)
- additional statement on complex meta-analytic designs (see Appendix B)

## Dependent Variables

Information on the following outcomes will be assessed on the same page as the corresponding PLS.

### Perceived Accessibility

The following item will be analyzed in confirmatory analyses: "Die Sprache in dieser Zusammenfassung ist einfach zu lesen." This item will be measured on a rating scale ranging from 1 ("stimme gar nicht zu") to 8 ("stimme voll und ganz zu").

Additionally, the following three exploratory items will be included. These items might be combined to build an accessibility scale for exploratory analyses.

- "Wie viel einfacher oder schwieriger ist die Sprache der Zusammenfassung im Vergleich zu dem folgenden Text von Jäger zu lesen?"

*Als Moderatorvariablen bezeichnet man in der Regel solche Variablen oder Variablensysteme, die bestehende Abhängigkeiten zwischen wiederum Variablen oder Variablensystemen in der Größe und/oder Richtung verändern (Jäger 1974).*

Die Zusammenfassung ist ... zu lesen als der Text von Jäger."

-4 ("sehr viel schwieriger") - 0 ("genauso einfach oder schwierig") +4 ("sehr viel einfacher")

- "Wie viel einfacher oder schwieriger ist die Sprache der Zusammenfassung im Vergleich zu dem folgenden Text aus einem Kinderbuch zu lesen?"

*Ein Kind sitzt an einem Fluss. Es langweilt sich. Dann rennt ein Tier vorbei. Es ist ein Kaninchen. Das Kaninchen ist weiß. Es rennt schnell.*

Die Zusammenfassung ist ... zu lesen als der Text des Kinderbuchs."

-4 ("sehr viel schwieriger") - 0 ("genauso einfach oder schwierig") +4 ("sehr viel einfacher")

- "Angenommen, wir würden uns an einem Nachmittag in eine belebte Innenstadt stellen und zufällig 100 Leute befragen - Was denken Sie, wie viele Menschen dieser Gruppe von 100 Leuten diese Zusammenfassung einfach zu lesen finden?"

### Perceived Understanding

The following item will be analyzed in confirmatory analyses: "Ich habe die Informationen aus dieser Zusammenfassung verstanden." This item will be measured on rating scales ranging from 1 ("stimme gar nicht zu") to 8 ("stimme voll und ganz zu").

Additionally, the following three exploratory items will be included. These items might be combined to build an understandability scale for exploratory analyses.

- “Wie viel besser oder schlechter haben Sie die Informationen aus der Zusammenfassung im Vergleich zu dem folgenden Text von Jäger verstanden?”

*Als Moderatorvariablen bezeichnet man in der Regel solche Variablen oder Variablensysteme, die bestehende Abhängigkeiten zwischen wiederum Variablen oder Variablensystemen in der Größe und/oder Richtung verändern (Jäger 1974).*

Ich habe die Informationen aus der Zusammenfassung ... verstanden als die Informationen aus dem Text von Jäger.”

-4 (sehr viel schlechter) - 0 (genauso gut oder schlecht) +4 (sehr viel besser)

- “Wie viel besser oder schlechter haben Sie die Informationen aus der Zusammenfassung im Vergleich zu dem folgenden Text aus einem Kinderbuch verstanden?”

*Ein Kind sitzt an einem Fluss. Es langweilt sich. Dann rennt ein Tier vorbei. Es ist ein Kaninchen. Das Kaninchen ist weiß. Es rennt schnell.*

Ich habe die Informationen aus der Zusammenfassung ... verstanden als die Informationen aus dem Text des Kinderbuchs.”

-4 (sehr viel schlechter) - 0 (genauso gut oder schlecht) +4 (sehr viel besser)

- “Angenommen, wir würden uns an einem Nachmittag in eine belebte Innenstadt stellen und zufällig 100 Leute befragen - Was denken Sie, wie viele Menschen dieser Gruppe von 100 Leuten die Informationen aus dieser Zusammenfassung verstehen?”

#### Perceived Empowerment

The following item will be analyzed in confirmatory analyses:

- “Nach dem Lesen dieser Zusammenfassung kann ich bei diesem Thema mitreden.”

Additionally, the following two exploratory items will be included. These items might be combined to build an empowerment scale for exploratory analyses.

- “Nach dem Lesen dieser Zusammenfassung fühle ich mich kompetent.”
- “Das Lesen der Zusammenfassung hat mir dabei geholfen, solche oder ähnliche Forschungsergebnisse besser zu verstehen.”

All empowerment items will be measured on rating scales ranging from 1 (“stimme gar nicht zu”) to 8 (“stimme voll und ganz zu”).

#### Content-related knowledge

Information on content-related knowledge (whether participants acquired knowledge regarding the key message of the PLS) will be assessed on a separate page after the PLS. The exact wording of these items is given in Appendix C.

## Exploratory outcomes

As exploratory outcome variables, we will assess knowledge on the quality of evidence (see Appendix C) and whether individuals would be willing to read the corresponding full-text after having read this summary. (“Nachdem Sie diese Zusammenfassung gelesen haben, würden Sie gerne noch die Originalstudie lesen?” (ja/nein)). Participants are further asked to provide information on the reasons for their choice.

## Other Variables

**Demographics.** Information on age, sex and education will be collected (for quotas: “Hauptschulabschluss”, “Mittlere Reife”, “Hochschulreife”). Moreover, information on whether participants are holding a university degree and (if applicable) their study subject will be collected as a potential covariate.

**Covariates.** As a further potential covariate we will assess if participants are working in a field in which they have contact to science (“Ich habe in meinem Beruf viel Kontakt zu Wissenschaftsthemen oder wissenschaftlichen Texten.”). Agreement to this statement will be assessed on a rating scale that ranges from 1 to 8. After randomly assigning participants to study arm A or B, we will additionally present them with the titles of the PLS that they are going to read and they rate the personal relevance and their interest in this topic on 1 to 8 rating scales (“Das Thema X betrifft mich in meinem Alltag.”/“Das Thema X finde ich interessant.”). Their interest will also be assessed once more after they have read the corresponding PLS.

**User Survey.** Moreover, a user survey will be conducted at the end of the study. Participants’ interest in twenty psychological categories (e.g. “Developmental topics across the lifespan”) extracted from the first project study will be assessed on a 1-8 rating scale ( 1 = “Finde ich gar nicht interessant”, 8 = “Finde ich sehr interessant”). For categories with a topic interest of 7 or higher, participants will be asked to name one or multiple reasons for their interest (“Warum interessiert Sie das Thema “X”?”) via a checkbox-format and, if selected, additional free text. Furthermore, they will be asked two questions about the merit of and their future interest in PLS.

Additionally, we will probe individual epistemic beliefs on the justification of knowledge: *justification by authority* and *personal justification* via an adapted version of the domain-general-questionnaire by Klopp and Stark (2016), which is in turn based on Ferguson et al. (2013). Finally, participants’ trust in scientists based on the presented PLS will be assessed using the Muenster Epistemic Trustworthiness Inventory (METI, Hendriks et al., 2015). All user survey measures are provided in Appendix D.

## Statistical Analysis

### *Sample Size Calculation and Power Analysis*

A power analysis was conducted using the software GPower (Faul et al., 2009). As a statistical test, we selected F-Test on between-level factors in repeated measures ANOVAs (since all our hypotheses were specified with regard to main effects of between-level variables). The following parameters were specified: Small ( $f = .1$ ) effect,  $\alpha = .05$ , power = .875 with four groups (i.e., the maximum number of independent variable conditions that

will be realized for approach for explaining statistical terms), two measures and an expected correlation between measures of .50. This power analysis indicated that a sample size of 992 participants was required. Thus, we decided to recruit at least 1000 participants to test hypotheses 1 and 2. Since all hypotheses are tested twice—once for complex and once for less complex PLS—in our study (see below) and additionally quota restrictions apply (see above), a total sample size of 2004 participants will be recruited.

### *Analysis Plan*

Data will be analyzed by means of mixed models (i.e., with experimental conditions as factors) or RM ANOVAs. Hypotheses 1, 2, 3 and 4 will be tested on the subsample of participants in study arm A and B separately. Exploratory analyses will be conducted to compare effects between study arms A and B. One-sided hypothesis tests will be conducted when appropriate and the significance of effects will be tested at  $p < .05$ .

Additional exploratory analyses on trustworthiness of scientists (METI) and epistemic justification beliefs may be conducted.

## References

- Anzinger, H., Elliott, S. A., & Hartling, L. (2020). Comparative Usability Analysis and Parental Preferences of Three Web-Based Knowledge Translation Tools: Multimethod Study. *Journal of Medical Internet Research*, 22(3), e14562. <https://doi.org/10.2196/14562>
- Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, 19(6), 901–917. <https://doi.org/10.1111/desc.12341>
- Bucher, L., Tran, U. S., Prinz, G. M., Burgard, T., Bosnjak, M., & Voracek, M. (2020). *Keeping meta-analyses alive and well: Using PsychOpenCAMA to implement a community-augmented meta-analysis on the Dark Triad of personality*.
- Buljan, I., Tokalić, R., Roguljić, M., Zakarija-Grković, I., Vrdoljak, D., Milić, P., Puljak, L., & Marušić, A. (2020). Framing the numerical findings of Cochrane plain language summaries: Two randomized controlled trials. *BMC Medical Research Methodology*, 20(1), 101. <https://doi.org/10.1186/s12874-020-00990-4>
- Burgard, T., Bošnjak, M., & Studtrucker, R. (2021). Community-Augmented Meta-Analyses (CAMAs) in Psychology: Potentials and Current Systems. *Zeitschrift Für Psychologie*, 229(1), 15–23. <https://doi.org/10.1027/2151-2604/a000431>
- Ellen, M. E., Lavis, J. N., Wilson, M. G., Grimshaw, J., Haynes, R. B., Ouimet, M., Raina, P., & Gruen, R. (2014). Health system decision makers' feedback on summaries and tools supporting the use of systematic reviews: A qualitative study. *Evidence & Policy: A Journal of Research, Debate and Practice*, 10(3), 337–359. <https://doi.org/10.1332/174426514x672362>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Ferguson, L. E., Bråten, I., Strømsø, H. I., & Anmarkrud, Ø. (2013). Epistemic beliefs and comprehension in the context of reading multiple documents: Examining the role of conflict. *International Journal of Educational Research*, 62, 100–114. <https://doi.org/10.1016/j.ijer.2013.07.001>
- 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>
- Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOS ONE*, 10(10), e0139309. <https://doi.org/10.1371/journal.pone.0139309>
- Jelicic Kadic, A., 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, 61. <https://doi.org/10.1186/s12874-016-0162-y>
- Klopp, E., & Stark, R. (2016). *Entwicklung eines Fragebogens zur Erfassung domänenübergreifender epistemologischer Überzeugungen [Development of a domain-general epistemological beliefs questionnaire]* [Unpublished Manuscript].
- Krapp, A., & Prenzel, M. (2011). Research on Interest in Science: Theories, methods, and findings. *International Journal of Science Education*, 33(1), 27–50. <https://doi.org/10.1080/09500693.2010.518645>
- Raynor, D. K., Myers, L., Blackwell, K., Kress, B., Dubost, A., & Joos, A. (2018). Clinical Trial Results Summary for Laypersons: A User Testing Study. *Therapeutic Innovation & Regulatory Science*, 52(5), 606–628. <https://doi.org/10.1177/2168479017753129>
- Santesso, N., Glenton, C., & Lang, B. (2008). Evidence that patients can understand and use? *Zeitschrift für Evidenz, Fortbildung und Qualität Im Gesundheitswesen*, 102(8), 493–496. <https://doi.org/10.1016/j.zefq.2008.08.027>

- 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>
- Schiefele, U., Krapp, A., Wild, K.-P., & Winteler, A. (1993). Der "Fragebogen zum Studieninteresse" (FSI); 1993. *Diagnostika*, *39*(4), 335–351.
- Schwalm, F. D., Zandavalli, R. B., de Castro Filho, E. D., & Lucchetti, G. (2021). Is there a relationship between spirituality/religiosity and resilience? A systematic review and meta-analysis of observational studies. *Journal of Health Psychology*, *135*910532098453. <https://doi.org/10.1177/1359105320984537>
- Yule, K., Houston, J., & Grych, J. (2019). Resilience in Children Exposed to Violence: A Meta-analysis of Protective Factors Across Ecological Contexts. *Clinical Child and Family Psychology Review*, *22*(3), 406–431. <https://doi.org/10.1007/s10567-019-00293-1>

## Appendix A: PLS

### Less complex Unstructured (Study Arm A)

Schwalm, F. D., Zandavalli, R. B., de Castro Filho, E. D., & Lucchetti, G. (2021). Is there a relationship between spirituality/religiosity and resilience? A systematic review and meta-analysis of observational studies. *Journal of Health Psychology*, 135910532098453. <https://doi.org/10.1177/1359105320984537>

#### *Glaube an etwas Göttliches oder einen höheren Sinn und die Fähigkeit zur Erholung von belastenden Ereignissen*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel "Is there a relationship between spirituality/religiosity and resilience? A systematic review and meta-analysis of observational studies" von Fábio Duarte Schwalm und KollegInnen aus dem Jahr 2021.

Was war das Ziel der Übersichtsarbeit?

Es gibt zahlreiche einzelne Studien, die darauf hindeuten, dass es einen Zusammenhang zwischen bestimmten Formen des Glaubens, nämlich Religiosität, also dem Glauben an etwas Göttliches, oder Spiritualität, also dem Glauben an einen höheren Sinn oder höhere Mächte, und der Fähigkeit zur Erholung von belastenden Ereignissen gibt. Die ForscherInnen untersuchten in der Übersichtsarbeit folgende Forschungsfrage: Erholen sich gläubige Menschen besser oder schlechter von belastenden Ereignissen?

Kernaussage der Übersichtsarbeit

Es gibt einen positiven Zusammenhang zwischen Religiosität oder Spiritualität und der Fähigkeit zur Erholung von belastenden Ereignissen: Je gläubiger Menschen sind, desto besser erholen sie sich von belastenden Ereignissen.

Wie wurde vorgegangen?

Die ForscherInnen suchten Studien über die Beziehung zwischen Religiosität oder Spiritualität und der Fähigkeit zur Erholung von belastenden Ereignissen. Sie fanden 34 Studien, deren Ergebnisse sie zu einer Meta-Analyse zusammenfassen konnten. Dann betrachteten sie den Zusammenhang zwischen Religiosität oder Spiritualität und der Fähigkeit, sich von belastenden Ereignissen zu erholen, über alle 34 Studien hinweg.

Was sind die wichtigsten Ergebnisse?

Über die 34 Studien hinweg gab es einen Zusammenhang zwischen Religiosität oder Spiritualität und der Fähigkeit zur Erholung von belastenden Ereignissen. [Condition Qualitative Statement: Der Zusammenhang war mittelstark.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Korrelation betrug  $r = 0.40$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Das ist ein mittelstarker Zusammenhang.] Gläubige erholten sich also besser von belastenden Ereignissen.

Statistical Glossary (included in conditions: *Effect Size + Glossary* and *Effect Size + Glossary + Qualitative Statement*)

**Korrelation r:** Gibt die Richtung und Stärke des Zusammenhangs zwischen zwei Merkmalen (z.B. Dauer einer Therapie und Wohlbefinden) an. Eine Korrelation kann Werte von -1 bis +1 annehmen. Wenn r größer als 0 ist, so besteht ein positiver Zusammenhang. Dann steigt z.B. mit einer höheren Anzahl an Therapiesitzungen das Wohlbefinden immer weiter an. Wenn r hingegen kleiner als 0 ist, dann besteht ein negativer Zusammenhang. Das ist z.B. der Fall, wenn mit einer höheren Anzahl an Therapiesitzungen das Wohlbefinden immer weiter abnimmt.

Je näher r dabei +1 oder -1 kommt, desto stärker ist der positive oder negative Zusammenhang. Dabei kann man folgende Faustregel nutzen:

- Bis zu Werten von  $r = +/- .10$  ist der Zusammenhang sehr schwach und praktisch nicht bedeutsam
- Ab Werten von  $r = +/- .10$  ist der Zusammenhang schwach
- Ab Werten von  $r = +/- .30$  ist der Zusammenhang mittelstark
- Ab Werten von  $r = +/- .50$  ist der Zusammenhang stark

Bucher, L., Tran, U. S., Prinz, G. M., Burgard, T., Bosnjak, M., & Voracek, M. (2020). Keeping meta-analyses alive and well: Using PsychOpenCAMA to implement a community-augmented meta-analysis on the Dark Triad of personality.

*Die Neigung, andere zum eigenen Vorteil zu beeinflussen und auszunutzen: Gibt es Geschlechtsunterschiede?*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel "Keeping meta-analyses alive and well: Using PsychOpenCAMA to implement a community-augmented meta-analysis on the Dark Triad of personality" von Bucher und KollegInnen aus dem Jahr 2020.

Was war das Ziel der Übersichtsarbeit?

Jeder Mensch weist verschiedene Eigenschaften auf, die seine Persönlichkeit ausmachen. Eine solche Eigenschaft ist der sogenannte Machiavellismus. Diese Eigenschaft besitzt jeder Mensch in unterschiedlich starker Ausprägung. Personen mit hohen Ausprägungen von Machiavellismus sind eher bereit andere zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen. Sie werden daher oft als berechnend oder kalt bezeichnet. Auf Basis bisheriger Forschung wollten die ForscherInnen in der Übersichtsarbeit herausfinden: Ist diese Eigenschaft bei Männern stärker ausgeprägt als bei Frauen?

Kernaussage der Übersichtsarbeit

Männer sind eher bereit, andere Personen zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen. Frauen zeigen diese Eigenschaft nicht so stark.

Wie wurde vorgegangen?

Die ForscherInnen suchten Studien, die verglichen, ob Männer oder Frauen eher bereit waren, andere zu beeinflussen und auszunutzen, um ihre eigenen Ziele zu erreichen. Die ForscherInnen fanden 207 Studien, deren Ergebnisse sie zu einer Metaanalyse zusammenfassen konnten. In allen Studien schauten die ForscherInnen, welches Geschlecht diese Eigenschaft eher zeigte.

Was sind die wichtigsten Ergebnisse?

Über 207 Studien hinweg zeigten Männer eine stärkere Bereitschaft, andere Personen zu manipulieren, um ihre Ziele zu erreichen. Bei Frauen war diese Eigenschaft nicht so stark ausgeprägt. **[Condition Qualitative Statement: Der Unterschied zwischen Frauen und Männern war klein, aber vorhanden.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke Hedges g betrug 0.34.] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Unterschied zwischen Frauen und Männern ist demnach klein, aber vorhanden.]** Das bestätigt die ursprüngliche Annahme der ForscherInnen, dass Machiavellismus bei Männern stärker ausgeprägt ist als bei Frauen.

Statistical Glossary (included in conditions: *Effect Size + Glossary* and *Effect Size + Glossary + Qualitative Statement*)

**Hedges g:** Ist ein Maß für den Unterschied zwischen zwei Gruppen im Hinblick auf ein Merkmal. Zum Beispiel kann man damit prüfen, ob eine Therapie bei Männern zu mehr Wohlbefinden führt als bei Frauen. Dabei wird zunächst der Unterschied zwischen den Geschlechtern im Wohlbefinden berechnet. Anschließend wird dieser Unterschied noch

einmal umgerechnet, sodass Hedges  $g$  über mehrere Betrachtungen hinweg (z.B. zwischen verschiedenen Studien) vergleichbar ist.

Hedges  $g$  kann sowohl Werte größer als auch kleiner 0 annehmen. Je stärker dieser Wert von 0 abweicht, desto größer ist der Unterschied zwischen den Gruppen. Dabei kann man folgende Faustregel nutzen:

- Bis zu Werten von  $g = \pm 0.20$  ist der Unterschied sehr klein und praktisch nicht bedeutsam
- Ab Werten von  $g = \pm 0.20$  ist der Unterschied klein
- Ab Werten von  $g = \pm 0.50$  ist der Unterschied mittelgroß
- Ab Werten von  $g = \pm 0.80$  ist der Unterschied groß

## Less complex Structured (Study Arm A)

Schwalm, F. D., Zandavalli, R. B., de Castro Filho, E. D., & Lucchetti, G. (2021). Is there a relationship between spirituality/religiosity and resilience? A systematic review and meta-analysis of observational studies. *Journal of Health Psychology*, 135910532098453. <https://doi.org/10.1177/1359105320984537>

### *Glaube an etwas Göttliches oder einen höheren Sinn und die Fähigkeit zur Erholung von belastenden Ereignissen*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel "Is there a relationship between spirituality/religiosity and resilience? A systematic review and meta-analysis of observational studies" von Fábio Duarte Schwalm und KollegInnen aus dem Jahr 2021.

Was war das Ziel der Übersichtsarbeit?

**Hintergrund:** Es gibt zahlreiche einzelne Studien, die darauf hindeuten, dass es einen Zusammenhang zwischen bestimmten Formen des Glaubens, nämlich Religiosität, also dem Glauben an etwas Göttliches, oder Spiritualität, also dem Glauben an einen höheren Sinn oder höhere Mächte, und der Fähigkeit zur Erholung von belastenden Ereignissen gibt.

**Forschungsfrage:** Die ForscherInnen untersuchten in der Übersichtsarbeit folgende Forschungsfrage: Erholen sich gläubige Menschen besser oder schlechter von belastenden Ereignissen?

Kernaussage der Übersichtsarbeit

Es gibt einen positiven Zusammenhang zwischen Religiosität oder Spiritualität und der Fähigkeit zur Erholung von belastenden Ereignissen: Je gläubiger Menschen sind, desto besser erholen sie sich von belastenden Ereignissen.

Wie wurde vorgegangen?

**Welche Studien wurden gesucht?** Die ForscherInnen suchten Studien über die Beziehung zwischen Religiosität oder Spiritualität und der Fähigkeit zur Erholung von belastenden Ereignissen.

**Welche Studien wurden gefunden?** Die ForscherInnen fanden 34 Studien, deren Ergebnisse sie zu einer Meta-Analyse zusammenfassen konnten.

**Was haben die ForscherInnen gemacht?** Die ForscherInnen betrachteten den Zusammenhang zwischen Religiosität oder Spiritualität und der Fähigkeit, sich von belastenden Ereignissen zu erholen, über alle 34 Studien hinweg.

**Was wurde untersucht?**

- Spiritualität
- Religiosität
- Fähigkeit zur Erholung von belastenden Ereignissen

Was sind die wichtigsten Ergebnisse?

- Über die 34 Studien hinweg gab es einen Zusammenhang zwischen Religiosität oder Spiritualität und der Fähigkeit zur Erholung von belastenden Ereignissen. **[Condition Qualitative Statement: Der Zusammenhang war mittelstark.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative**

Statement: Die Korrelation betrug  $r = 0.40$ . [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Das ist ein mittelstarker Zusammenhang.]  
Gläubige erholten sich also besser von belastenden Ereignissen.

<statistical glossary as provided for the unstructured condition>

Bucher, L., Tran, U. S., Prinz, G. M., Burgard, T., Bosnjak, M., & Voracek, M. (2020). Keeping meta-analyses alive and well: Using PsychOpenCAMA to implement a community-augmented meta-analysis on the Dark Triad of personality.

*Die Neigung, andere zum eigenen Vorteil zu beeinflussen und auszunutzen: Gibt es Geschlechtsunterschiede?*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel "Keeping meta-analyses alive and well: Using PsychOpenCAMA to implement a community-augmented meta-analysis on the Dark Triad of personality" von Bucher und KollegInnen aus dem Jahr 2020.

Was war das Ziel der Übersichtsarbeit?

**Hintergrund:** Jeder Mensch weist verschiedene Eigenschaften auf, die seine Persönlichkeit ausmachen. Eine solche Eigenschaft ist der sogenannte Machiavellismus. Diese Eigenschaft besitzt jeder Mensch in unterschiedlich starker Ausprägung. Personen mit hohen Ausprägungen von Machiavellismus sind eher bereit andere zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen. Sie werden daher oft als berechnend oder kalt bezeichnet.

**Forschungsfrage:** Auf Basis bisheriger Forschung wollten die ForscherInnen in der Übersichtsarbeit herausfinden: Ist diese Eigenschaft bei Männern stärker ausgeprägt als bei Frauen?

Kernaussage der Übersichtsarbeit

Männer sind eher bereit, andere Personen zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen. Frauen zeigen diese Eigenschaft nicht so stark.

Wie wurde vorgegangen?

**Welche Studien wurden gesucht?** Die ForscherInnen suchten Studien, die verglichen, ob Männer oder Frauen eher bereit waren, andere zu beeinflussen und auszunutzen, um ihre eigenen Ziele zu erreichen.

**Welche Studien wurden gefunden?** Die ForscherInnen fanden 207 Studien, deren Ergebnisse sie zu einer Metaanalyse zusammenfassen konnten.

**Was haben die ForscherInnen gemacht?** In allen Studien schauten die ForscherInnen, welches Geschlecht diese Eigenschaft eher zeigte.

**Was wurde untersucht?**

- Machiavellismus
- Geschlecht: männlich oder weiblich

Was sind die wichtigsten Ergebnisse?

- Über 207 Studien hinweg zeigten Männer eine stärkere Bereitschaft, andere Personen zu manipulieren, um ihre Ziele zu erreichen. Bei Frauen war diese Eigenschaft nicht so stark ausgeprägt. [Condition Qualitative Statement: Der Unterschied zwischen Frauen und Männern war klein, aber vorhanden.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke Hedges g betrug 0.34.] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Unterschied zwischen Frauen und Männern ist demnach klein, aber vorhanden.] Das bestätigt die ursprüngliche Annahme der ForscherInnen, dass Machiavellismus bei Männern stärker ausgeprägt ist als bei Frauen.

<statistical glossary as provided for the unstructured condition>

## Complex Unstructured (Study Arm B)

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, 19(6), 901–917. <https://doi.org/10.1111/desc.12341>

### *Wie Babys Sprache lernen - Mögen Babys lieber bekannte oder unbekannte Wörter?*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel "Development of infants' segmentation of words from native speech: a meta-analytic approach" von Christina Bergmann und Alejandrina Cristia aus dem Jahr 2016.

Was war das Ziel der Übersichtsarbeit?

Babys lernen bereits im Alter von 6 Monaten erste Wörter. Dafür müssen sie Wörter in Sätzen erkennen können. Es wird angenommen, dass Babys mit zunehmendem Alter besser im Erkennen von Wörtern werden. Jüngere Babys scheinen bereits bekannte Wörter zu bevorzugen. Ältere Babys scheinen unbekannte Wörter lieber zu mögen. In dieser Übersichtsarbeit wollten die ForscherInnen herausfinden: Stimmt es tatsächlich, dass jüngere Babys bekannte, ältere Babys aber unbekannte Wörter lieber mögen?

Kernaussage der Übersichtsarbeit

Babys mögen bekannte Wörter lieber als unbekannte. Auch bei älteren Babys ist das so. Ältere Babys mögen also nicht, wie zuvor angenommen, unbekannte Wörter lieber.

Wie wurde vorgegangen?

Die ForscherInnen suchten Studien, die verglichen, wie Babys auf bekannte und neue Wörter in Sätzen reagieren. Die Studien mussten echte, gesprochene Wörter benutzen und keine erfundenen Wörter. Die ForscherInnen fanden insgesamt 168 Studien, deren Ergebnisse sie zu einer Meta-Analyse zusammenfassen konnten. Die Babys waren zwischen 6 und 24 Monaten alt. In allen Studien schauten die ForscherInnen, ob die Babys Sätze lieber mochten, die bekannte Wörter statt unbekanntem Wörtern enthalten. Auch das Alter der Kinder interessierte die ForscherInnen. Sie gingen davon aus, dass ältere Babys unbekannte Wörter irgendwann lieber mögen, da sie schon Profis im Erkennen von Wörtern sind.

Was sind die wichtigsten Ergebnisse?

Über alle Studien hinweg mochten die Babys bekannte Wörter lieber als unbekannte. [Condition Qualitative Statement: Der Unterschied zwischen bekannten und unbekanntem Wörtern war klein, aber vorhanden.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke Hedges g betrug 0.22.] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Unterschied zwischen bekannten und unbekanntem Wörtern war also klein, aber vorhanden.] Anders als erwartet, mochten auch ältere Babys bekannte Wörter lieber als unbekannte. [Condition Qualitative Statement: Der Einfluss des Alters war unbedeutend.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der zugehörige Regressionskoeffizient Beta betrug 0.0003.] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Einfluss des Alters war also unbedeutend.] Die ForscherInnen gehen deshalb davon aus, dass ältere Babys andere Wörter lieber mögen als bisher angenommen.

**Hedges g:** Ist ein Maß für den Unterschied zwischen zwei Gruppen im Hinblick auf ein Merkmal. Zum Beispiel kann man damit prüfen, ob eine Therapie bei Männern zu mehr Wohlbefinden führt als bei Frauen. Dabei wird zunächst der Unterschied zwischen den Geschlechtern im Wohlbefinden berechnet. Anschließend wird dieser Unterschied noch einmal umgerechnet, sodass Hedges g über mehrere Betrachtungen hinweg (z.B. zwischen verschiedenen Studien) vergleichbar ist.

Hedges g kann sowohl Werte größer als auch kleiner 0 annehmen. Je stärker dieser Wert von 0 abweicht, desto größer ist der Unterschied zwischen den Gruppen. Dabei kann man folgende Faustregel nutzen:

- Bis zu Werten von  $g = \pm 0.20$  ist der Unterschied sehr klein und praktisch nicht bedeutsam
- Ab Werten von  $g = \pm 0.20$  ist der Unterschied klein
- Ab Werten von  $g = \pm 0.50$  ist der Unterschied mittelgroß
- Ab Werten von  $g = \pm 0.80$  ist der Unterschied groß

**Moderator/Moderation:** Bezeichnet den Einfluss, den ein Merkmal auf den Zusammenhang zwischen zwei anderen Merkmalen hat. Eine Psychotherapie kann z.B. allgemein zu verbesserter Stimmung führen. Im Detail zeigt sich aber, dass das Alter dabei eine Rolle spielt: Die Therapie verbessert die Stimmung stärker für jüngere Menschen als für Ältere. Alter ist in diesem Beispiel der Moderator, da es auf den Zusammenhang zwischen Therapie und Verbesserung der Stimmung einwirkt.

**Regressionsgewicht Beta:** Beschreibt, wie sehr ein Merkmal (z.B. Anzahl Sitzungen einer Psychotherapie) ein zweites Merkmal (z.B. Wohlbefinden) vorhersagen kann. B kann Werte größer oder kleiner 0 annehmen. Wenn B größer als 0 ist, so besteht ein positiver Zusammenhang. Mehr absolvierte Therapiesitzungen sagen in diesem Fall ein höheres Wohlbefinden vorher. Wenn Beta hingegen kleiner als 0 ist, so besteht ein negativer Zusammenhang. Das ist z.B. der Fall, wenn mehr Sitzungen einer Psychotherapie ein geringeres Wohlbefinden vorhersagen.

Je größer der Wert von Beta in positive oder negative Richtung wird, desto stärker der Zusammenhang. Dabei kann man folgende Faustregel nutzen:

- Bis zu Werten von  $Beta = \pm 0.10$  ist der Zusammenhang sehr schwach und praktisch nicht bedeutsam
- Ab Werten von  $Beta = \pm 0.10$  ist der Zusammenhang schwach
- Ab Werten von  $Beta = \pm 0.30$  ist der Zusammenhang mittelstark
- Ab Werten von  $Beta = \pm 0.50$  ist der Zusammenhang stark

Yule, K., Houston, J., & Grych, J. (2019). Resilience in Children Exposed to Violence: A Meta-analysis of Protective Factors Across Ecological Contexts. *Clinical Child and Family Psychology Review*, 22(3), 406–431.

### *Was befähigt Kinder dazu, sich besser von Gewalterfahrungen zu erholen?*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel “Resilience in Children Exposed to Violence: A Meta-analysis of Protective Factors Across Ecological Contexts” von Kristen Yule und KollegInnen aus dem Jahr 2019.

#### Was war das Ziel der Übersichtsarbeit?

Manche Kinder erleben Gewalt im Alltag. Kinder, die Gewalt erleben, entwickeln häufig auffälliges Verhalten und psychische Probleme. Das gilt aber nicht für alle Kinder, die Gewalt erleben. Manche Kinder entwickeln sich trotz Gewalterfahrung unauffällig. Diese Kinder können sich besonders gut von belastenden Ereignissen erholen. Die ForscherInnen untersuchten in der Übersichtsarbeit folgende Forschungsfragen: Warum zeigen manche Kinder diese Fähigkeit zur Erholung von belastenden Ereignissen und andere nicht? Welche Faktoren spielen bei der Verarbeitung von Gewalterfahrungen eine bedeutsame Rolle?

#### Kernaussage der Übersichtsarbeit

Es wurden fünf Schutzfaktoren gefunden, die Kindern halfen, Gewalterfahrungen zu verarbeiten: Unterstützung durch die Familie; Unterstützung durch die Schule; Unterstützung durch Gleichaltrige; Einbindung in eine religiöse Gemeinschaft; die Fähigkeit, eigene Gefühle und eigenes Verhalten zu steuern.

#### Wie wurde vorgegangen?

Die ForscherInnen suchten nach Studien, die den Zusammenhang zwischen Gewalterfahrung und der Fähigkeit, sich von belastenden Ereignissen zu erholen, bei Minderjährigen untersuchten. Insgesamt fanden die ForscherInnen 118 Studien, deren Ergebnisse sie zu einer Meta-Analyse zusammenfassen konnten. Im Rahmen der Meta-Analyse wurde dann geprüft, welche sogenannten Schutzfaktoren mit einer höheren Fähigkeit, sich von belastenden Ereignissen zu erholen, zusammenhingen. Die untersuchten Schutzfaktoren waren familiäre Faktoren, Unterstützung durch Gleichaltrige oder die Schule sowie persönliche oder gesellschaftliche Faktoren. In der Meta-Analyse wurde außerdem betrachtet, ob die Zusammenhänge zwischen Gewalterfahrung, Schutzfaktoren und der Fähigkeit, sich von belastenden Ereignissen zu erholen, nicht nur zum Zeitpunkt der Befragung (also gegenwärtig) besteht, sondern sich auch über die Zeit hinweg (also langfristig) bei den betroffenen Kindern zeigt.

#### Was sind die wichtigsten Ergebnisse?

Über 118 Studien hinweg wurden 11 verschiedene Schutzfaktoren getestet. Davon waren 5 besonders bedeutsam, weil sich ihre schützende Wirkung sowohl gegenwärtig als auch langfristig zeigte. Die Unterstützung durch die Familie war einer dieser Schutzfaktoren.

[Condition Qualitative Statement: Der Zusammenhang war hier schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.16$  und  $R = 0.18$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also schwach.] Ein weiterer Schutzfaktor ist die Unterstützung durch die Schule. [Condition Qualitative Statement: Der Zusammenhang war hier schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.20$  und  $R = 0.21$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also schwach.] Auch die Unterstützung durch Gleichaltrige war ein Schutzfaktor. [Condition

Qualitative Statement: Der Zusammenhang war hier erneut schwach. [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke betrug hier  $R = 0.12$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Das ist erneut ein schwacher Zusammenhang.] Bei den persönlichen Faktoren zeigte sich, dass die Fähigkeit, eigene Gefühle und eigenes Verhalten selbst zu steuern, ein Schutzfaktor ist. [Condition Qualitative Statement: Der Zusammenhang war hier mittelstark.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.30$  und  $R = 0.45$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also mittelstark.] Ein gesellschaftlicher Schutzfaktor war die Eingebundenheit in eine religiöse Gemeinschaft. [Condition Qualitative Statement: Der Zusammenhang war hier sehr schwach bis schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.05$  und  $R = 0.16$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also sehr schwach bis schwach.] Diese fünf Schutzfaktoren scheinen also bei Kindern mit Gewalterfahrungen die Fähigkeit, sich von belastenden Ereignissen zu erholen, zu stärken.

Statistical Glossary (included in conditions: *Effect Size + Glossary* and *Effect Size + Glossary + Qualitative Statement*)

**Gewichtete Effektstärke R:** Gibt die Richtung und Stärke des Zusammenhangs zwischen zwei Merkmalen (z.B. Dauer einer Therapie und Wohlbefinden) an. R kann Werte von -1 bis +1 annehmen. Wenn R größer als 0 ist, so besteht ein positiver Zusammenhang. Dann steigt z.B. mit einer höheren Anzahl an Therapiesitzungen das Wohlbefinden immer weiter an. Wenn R hingegen kleiner als 0 ist, dann besteht ein negativer Zusammenhang. Das ist z.B. der Fall, wenn mit einer höheren Anzahl an Therapiesitzungen das Wohlbefinden immer weiter abnimmt.

Je näher R dabei +1 oder -1 kommt, desto stärker ist der positive oder negative Zusammenhang. Dabei kann man folgende Faustregel nutzen:

- Bis zu Werten von  $R = +/- .10$  ist der Zusammenhang sehr schwach und praktisch nicht bedeutsam
- Ab Werten von  $R = +/- .10$  ist der Zusammenhang schwach
- Ab Werten von  $R = +/- .30$  ist der Zusammenhang mittelstark
- Ab Werten von  $R = +/- .50$  ist der Zusammenhang stark

“Gewichtet” bedeutet, dass die Autor\*Innen R über alle Studien hinweg vergleichbar machten. So spielt es dann etwa keine Rolle, wenn Studien unterschiedlich viele Teilnehmer\*Innen hatten.

## Complex Structured (Study Arm B)

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, 19(6), 901–917. <https://doi.org/10.1111/desc.12341>

### *Wie Babys Sprache lernen - Mögen Babys lieber bekannte oder unbekannte Wörter?*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel "Development of infants' segmentation of words from native speech: a meta-analytic approach" von Christina Bergmann und Alejandrina Cristia aus dem Jahr 2016.

Was war das Ziel der Übersichtsarbeit?

**Hintergrund:** Babys lernen bereits im Alter von 6 Monaten erste Wörter. Dafür müssen sie Wörter in Sätzen erkennen können. Es wird angenommen, dass Babys mit zunehmendem Alter besser im Erkennen von Wörtern werden. Jüngere Babys scheinen bereits bekannte Wörter zu bevorzugen. Ältere Babys scheinen unbekannte Wörter lieber zu mögen.

**Forschungsfrage:** In dieser Übersichtsarbeit wollten die ForscherInnen herausfinden: Stimmt es tatsächlich, dass jüngere Babys bekannte, ältere Babys aber unbekannte Wörter lieber mögen?

Kernaussage der Übersichtsarbeit

Babys mögen bekannte Wörter lieber als unbekannte. Auch bei älteren Babys ist das so. Ältere Babys mögen also nicht, wie zuvor angenommen, unbekannte Wörter lieber.

Wie wurde vorgegangen?

**Welche Studien wurden gesucht?** Die ForscherInnen suchten Studien, die verglichen, wie Babys auf bekannte und neue Wörter in Sätzen reagieren. Die Studien mussten echte, gesprochene Wörter benutzen und keine erfundenen Wörter.

**Welche Studien wurden gefunden?** Die ForscherInnen fanden insgesamt 168 Studien, deren Ergebnisse sie zu einer Meta-Analyse zusammenfassen konnten. Die Babys waren zwischen 6 und 24 Monaten alt.

**Was haben die ForscherInnen gemacht?** In allen Studien schauten die ForscherInnen, ob die Babys Sätze lieber mochten, die bekannte Wörter statt unbekanntem Wörtern enthalten. Auch das Alter der Kinder interessierte die ForscherInnen. Sie gingen davon aus, dass ältere Babys unbekannte Wörter irgendwann lieber mögen, da sie schon Profis im Erkennen von Wörtern sind.

**Was wurde untersucht?**

- Vorliebe für unbekannte und bekannte Wörter
- Alter der Babys

Was sind die wichtigsten Ergebnisse?

- Über alle Studien hinweg mochten die Babys bekannte Wörter lieber als unbekannte. [Condition Qualitative Statement: Der Unterschied zwischen bekannten und unbekanntem Wörtern war klein, aber vorhanden.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke Hedges g betrug 0.22.] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Unterschied zwischen bekannten und unbekanntem Wörtern war also klein, aber vorhanden.]
- Anders als erwartet, mochten auch ältere Babys bekannte Wörter lieber als unbekannte. [Condition Qualitative Statement: Der Einfluss des Alters war unbedeutend.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der zugehörige

Regressionskoeffizient Beta betrug 0.0003.] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Einfluss des Alters war also unbedeutend.] Die ForscherInnen gehen deshalb davon aus, dass ältere Babys andere Wörter lieber mögen als bisher angenommen.

<statistical glossary as provided for the unstructured condition>

Yule, K., Houston, J., & Grych, J. (2019). Resilience in Children Exposed to Violence: A Meta-analysis of Protective Factors Across Ecological Contexts. *Clinical Child and Family Psychology Review*, 22(3), 406–431.

### *Was befähigt Kinder dazu, sich besser von Gewalterfahrungen zu erholen?*

Diese Zusammenfassung bezieht sich auf die Übersichtsarbeit mit dem Titel “Resilience in Children Exposed to Violence: A Meta-analysis of Protective Factors Across Ecological Contexts” von Kristen Yule und KollegInnen aus dem Jahr 2019.

Was war das Ziel der Übersichtsarbeit?

**Hintergrund:** Manche Kinder erleben Gewalt im Alltag. Kinder, die Gewalt erleben, entwickeln häufig auffälliges Verhalten und psychische Probleme. Das gilt aber nicht für alle Kinder, die Gewalt erleben. Manche Kinder entwickeln sich trotz Gewalterfahrung unauffällig. Diese Kinder können sich besonders gut von belastenden Ereignissen erholen.  
**Forschungsfrage:** Die ForscherInnen untersuchten in der Übersichtsarbeit folgende Forschungsfragen: Warum zeigen manche Kinder diese Fähigkeit zur Erholung von belastenden Ereignissen und andere nicht? Welche Faktoren spielen bei der Verarbeitung von Gewalterfahrungen eine bedeutsame Rolle?

Kernaussage der Übersichtsarbeit

Es wurden fünf Schutzfaktoren gefunden, die Kindern halfen, Gewalterfahrungen zu verarbeiten: Unterstützung durch die Familie; Unterstützung durch die Schule; Unterstützung durch Gleichaltrige; Einbindung in eine religiöse Gemeinschaft; die Fähigkeit, eigene Gefühle und eigenes Verhalten zu steuern.

Wie wurde vorgegangen?

**Welche Studien wurden gesucht?** Die ForscherInnen suchten nach Studien, die den Zusammenhang zwischen Gewalterfahrung und der Fähigkeit, sich von belastenden Ereignissen zu erholen, bei Minderjährigen untersuchten.

**Welche Studien wurden gefunden?** Insgesamt fanden die ForscherInnen 118 Studien, deren Ergebnisse sie zu einer Meta-Analyse zusammenfassen konnten.

**Was haben die ForscherInnen gemacht?** Im Rahmen der Meta-Analyse wurde dann geprüft, welche sogenannten Schutzfaktoren mit einer höheren Fähigkeit, sich von belastenden Ereignissen zu erholen, zusammenhängen. Die untersuchten Schutzfaktoren waren familiäre Faktoren, Unterstützung durch Gleichaltrige oder die Schule sowie persönliche oder gesellschaftliche Faktoren. In der Meta-Analyse wurde außerdem betrachtet, ob die Zusammenhänge zwischen Gewalterfahrung, Schutzfaktoren und der Fähigkeit, sich von belastenden Ereignissen zu erholen, nicht nur zum Zeitpunkt der Befragung (also gegenwärtig) besteht, sondern sich auch über die Zeit hinweg (also langfristig) bei den betroffenen Kindern zeigt.

**Was wurde untersucht?**

- Fähigkeit zur Erholung von belastenden Ereignissen (Gewalterfahrung)
- familiäre Faktoren
- Unterstützung durch Gleichaltrige
- Unterstützung durch die Schule
- persönliche Faktoren
- gesellschaftliche Faktoren
- Zeitraum: gegenwärtig oder langfristig

Was sind die wichtigsten Ergebnisse?

Über 118 Studien hinweg wurden 11 verschiedene Schutzfaktoren getestet. Davon waren 5 besonders bedeutsam, weil sich ihre schützende Wirkung sowohl gegenwärtig als auch

langfristig zeigte.

- Die Unterstützung durch die Familie war einer dieser Schutzfaktoren. [Condition Qualitative Statement: Der Zusammenhang war hier schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.16$  und  $R = 0.18$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also schwach.]
- Ein weiterer Schutzfaktor ist die Unterstützung durch die Schule. [Condition Qualitative Statement: Der Zusammenhang war hier schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.20$  und  $R = 0.21$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also schwach.]
- Auch die Unterstützung durch Gleichaltrige war ein Schutzfaktor. [Condition Qualitative Statement: Der Zusammenhang war hier erneut schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke betrug hier  $R = 0.12$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Das ist erneut ein schwacher Zusammenhang.]
- Bei den persönlichen Faktoren zeigte sich, dass die Fähigkeit, eigene Gefühle und eigenes Verhalten selbst zu steuern, ein Schutzfaktor ist. [Condition Qualitative Statement: Der Zusammenhang war hier mittelstark.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.30$  und  $R = 0.45$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also mittelstark.]
- Ein gesellschaftlicher Schutzfaktor war die Eingebundenheit in eine religiöse Gemeinschaft. [Condition Qualitative Statement: Der Zusammenhang war hier sehr schwach bis schwach.] [Conditions Effect Size + Qualitative Statement/Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Die Effektstärke lag hier zwischen  $R = 0.05$  und  $R = 0.16$ .] [Conditions Effect Size + Glossary/Effect Size + Glossary + Qualitative Statement: Der Zusammenhang war also sehr schwach bis schwach.]

Diese fünf Schutzfaktoren sollten also bei Kindern mit Gewalterfahrungen die Fähigkeit, sich von belastenden Ereignissen zu erholen, stärken.

<statistical glossary as provided for the unstructured condition>

## Appendix B: Extension of the statement on the quality of evidence

The following statement on the quality of evidence was presented:

### <Statement presented in Study 1> **Gut zu wissen:**

Die Übersichtsarbeiten, zu denen Sie heute Zusammenfassungen lesen, sind Meta-Analysen. ForscherInnen, die eine Meta-Analyse machen, suchen zuerst die Ergebnisse aller Studien zu einer bestimmten Frage (z. B. "Wie gut hilft eine bestimmte Psychotherapie?"). Dann fassen sie die Ergebnisse dieser Studien zusammen.

Warum macht man das? Eine Meta-Analyse hat zwei Ziele:

- 1) Die Meta-Analyse gibt einen Überblick über alle Studien, die diese Frage schon untersucht haben. Sie beschreibt, wer schon zu einer Frage geforscht hat (z. B., wer schon alles diese Psychotherapie untersucht hat). Außerdem steht dort, was in den einzelnen Studien herauskam (z. B., ob und wie gut die Psychotherapie geholfen hat).
- 2) Die Meta-Analyse liefert außerdem einen Wert, der die Ergebnisse aller Studien zusammenfasst. Dafür nehmen die ForscherInnen die Einzelergebnisse aller gefundenen Studien und berechnen ein Gesamtergebnis. Es ist dabei wichtig, dass die gefundenen Studien zu einem Thema ziemlich ähnlich sind. Wenn sich die Studien sehr unterscheiden, wird das berechnete Gesamtergebnis der Meta-Analyse ungenau. Das berechnete Gesamtergebnis liefert den ForscherInnen die Antwort auf die anfangs gestellte Frage (z. B. "Über alle Studien hinweg hilft die Psychotherapie sehr gut."). **Dieses Gesamtergebnis hat eine viel höhere Aussagekraft als die Ergebnisse der einzelnen Studien.**

<Statement presented in Study 1>

<Extension in the present study> In manchen Fällen können die ForscherInnen aus den Daten sogar neue Werte berechnen (z. B. wie gut die Psychotherapie speziell jungen oder speziell älteren Erwachsenen hilft). Mit den einzelnen neuen Werten können sie dann neue Rechnungen machen, die vorher nicht möglich waren (z. B., ob die Psychotherapie bei jungen Erwachsenen besser hilft als bei älteren). Oft sind die einzelnen Studien für solche Berechnungen zu klein, weil sie zu wenige Daten haben. Das ist also ein weiterer Vorteil der Meta-Analyse: dass sie durch die große Menge an Daten Zusammenhänge untersuchen kann, die in den einzelnen Studien nicht untersucht werden können. <Extension in the present study>

This statement was presented on the same page as the PLS and on a separate page before the first PLS was presented.

## Appendix C Knowledge Test Items

- Knowledge on Quality of Evidence
  - Was ist eine Meta-Analyse?
    - Eine Meta-Analyse beschäftigt sich mit der Meta-Ebene von Psychologie (z. B. mit Intelligenz).
    - Eine Meta-Analyse heißt so, wenn mehr als 20 ForscherInnen in einer Forschergruppe sind.

- *Eine Meta-Analyse ist, wenn ForscherInnen verschiedener Institutionen zusammenarbeiten.*
- *Eine Meta-Analyse beschäftigt sich mit dem Grenzbereich zwischen Psychologie und Metaphysik.*
- *Eine Meta-Analyse ist ein Experiment, in dem zwei Gruppen miteinander verglichen werden.*
- *Eine Meta-Analyse sucht alle Studien zu einer Forschungsfrage und fasst deren Ergebnisse zusammen.*
- *Welche Aussage zu der Übersichtsarbeit, die Sie eben gelesen haben, stimmt? (je PLS)*
  - *Bergmann*
    - *Die ForscherInnen führten in ihrer Übersichtsarbeit 168 Studien dazu durch, ob Babys bekannte Wörter lieber mögen als unbekannte Wörter.*
    - *168 ForscherInnen untersuchten in einem großen Experiment, ob Babys bekannte Wörter lieber mögen als unbekannte Wörter und veröffentlichten die Ergebnisse dieser Studie.*
    - *Die ForscherInnen untersuchten insgesamt 168 Babys, um herauszufinden, ob Babys bekannte Wörter lieber mögen als unbekannte Wörter.*
    - *Es gab schon 168 ähnliche Studien dazu, ob Babys bekannte Wörter lieber mögen als unbekannte Wörter. Die ForscherInnen fassten die Ergebnisse dieser 168 Studien zusammen.*
  - *Bucher*
    - *Die ForscherInnen führten in ihrer Übersichtsarbeit über 170 Studien zu Geschlechterunterschieden in der Neigung, andere zum eigenen Vorteil zu beeinflussen und auszunutzen, durch.*
    - *170 ForscherInnen untersuchten in einem großen Experiment Geschlechterunterschiede in der Neigung, andere zum eigenen Vorteil zu beeinflussen und auszunutzen, und veröffentlichten die Ergebnisse dieser Studie.*
    - *Die ForscherInnen untersuchten insgesamt über 170 Personen, um mehr über Geschlechterunterschiede in der Neigung, andere zum eigenen Vorteil zu beeinflussen und auszunutzen, herauszufinden.*
    - *Es gab schon über 170 ähnliche Studien zu Geschlechterunterschieden in der Neigung, andere zum eigenen Vorteil zu beeinflussen und auszunutzen. Die ForscherInnen fassten die Ergebnisse dieser 170 Studien zu einem Wert zusammen.*
  - *Schwalm*
    - *Die ForscherInnen führten in der Übersichtsarbeit 34 verschiedene Studien dazu durch, ob höheres Sinnerleben oder auch der Glaube an etwas Göttliches mit der Fähigkeit zur Erholung von belastenden Lebensereignissen zusammenhängen. Die ForscherInnen werteten jede einzelne Studie aus.*

- *34 ForscherInnen untersuchten in einem großen Experiment, ob höheres Sinnerleben oder auch der Glaube an etwas Göttliches mit der Fähigkeit zur Erholung von belastenden Lebensereignissen zusammenhängen und veröffentlichten die Ergebnisse dieser Studie.*
- *Die ForscherInnen untersuchten insgesamt über 34 Personen, um mehr über den Zusammenhang zwischen höherem Sinnerleben oder auch dem Glauben an etwas Göttliches und der Fähigkeit zur Erholung von belastenden Lebensereignissen herauszufinden.*
- *Es gab schon 34 ähnliche Studien dazu, wie höheres Sinnerleben oder auch der Glaube an etwas Göttliches mit der Fähigkeit zur Erholung von belastenden Lebensereignissen zusammenhängen. Die ForscherInnen fassten die Ergebnisse dieser 34 Studien zu einem Wert zusammen.*
- *Yule*
  - *Die ForscherInnen führten in ihrer Übersichtsarbeit 118 verschiedene Studien zu Schutzfaktoren von Kindern, die Gewalt erlebt haben, durch. Die ForscherInnen werteten jede einzelne Studie aus.*
  - *118 ForscherInnen untersuchten in einem großen Experiment Schutzfaktoren von Kindern, die Gewalt erlebt haben, und veröffentlichten die Ergebnisse der Studie.*
  - *Die ForscherInnen untersuchten bei 118 Personen Schutzfaktoren von Kindern, die Gewalt erlebt haben. Sie fassten die Ergebnisse zu einem Wert zusammen.*
  - *Es gab schon 118 ähnliche Studien zu Schutzfaktoren von Kindern, die Gewalt erlebt haben. Die ForscherInnen fassten die Ergebnisse dieser 118 Studien zusammen.*
- *Content-related Knowledge*
  - *Welche Aussage zu den Ergebnissen der Übersichtsarbeit, die Sie eben gelesen haben, stimmt? (je PLS)*
    - *Bergmann*
      - *Babys mögen bekannte Wörter lieber als unbekannte Wörter. Auch bei älteren Babys ist das so. Ältere Babys mögen also nicht, wie zuvor angenommen, unbekannte Wörter lieber.*
      - *Babys mögen unbekannte Wörter lieber als bekannte Wörter. Auch bei älteren Babys ist das so. Ältere Babys mögen also nicht, wie zuvor angenommen, bekannte Wörter lieber.*
      - *Babys mögen unbekannte Wörter lieber als bekannte Wörter. Bei älteren Babys ist das nicht so. Ältere Babys mögen, wie zuvor angenommen, bekannte Wörter lieber.*
      - *Babys mögen bekannte Wörter lieber als unbekannte Wörter. Bei älteren Babys ist das nicht so. Ältere Babys mögen, wie zuvor angenommen, unbekannte Wörter lieber.*
    - *Bucher*

- *Männer sind eher bereit andere zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen. Frauen zeigen diese Eigenschaft nicht so stark.*
- *Männer sind eher nicht bereit andere zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen. Frauen zeigen diese Eigenschaft stärker.*
- *Männer sind genauso bereit wie Frauen, andere zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen.*
- *Männer und Frauen sind eher nicht bereit, andere zu beeinflussen und auszunutzen, um eigene Ziele zu erreichen.*
- **Schwalm**
  - *Je gläubiger Menschen waren, desto besser erholten sie sich von belastenden Ereignissen.*
  - *Je gläubiger Menschen waren, desto schlechter erholten sie sich von belastenden Ereignissen.*
  - *Es gab keinen Zusammenhang dazwischen, ob Menschen gläubig sind, und wie gut sie sich von belastenden Ereignissen erholen.*
  - *Ob Menschen sich gut von belastenden Ereignissen erholten, hing neben der Gläubigkeit auch davon ab, wie oft sie in die Kirche gehen.*
- **Yule**
  - *Es wurden fünf Faktoren gefunden, die wirksame Therapien für Kinder, die Gewalt erfahren haben, unterscheiden.*
  - *Es wurden fünf Faktoren gefunden, die Kindern helfen, Gewalterfahrungen zu verarbeiten.*
  - *Es wurden fünf Faktoren gefunden, die Gewalterfahrungen unterscheiden.*
  - *Es wurden fünf Faktoren gefunden, die Kinder davor schützen, Gewalt zu erfahren.*

## Appendix D User Survey Measures

- Interest in psychological topics
  - For twenty interest categories synthesized from study 1: “Wie sehr interessiert Sie das Thema “X” ?” (1 = “Finde ich gar nicht interessant”, 8 = “Finde ich sehr interessant”).
  - For categories rated 7 or above: “Warum interessiert Sie das Thema “X” ?”. Participants are presented with 5 possible reasons in a checkbox-format based on previous work by Schiefele et al. (1993) and Krapp and Prenzel (2011).
    - Weil mir die Auseinandersetzung mit dem Thema Spaß macht.
    - Weil mir das Thema persönlich sehr wichtig ist.
    - Weil das Thema für die Gesellschaft sehr wichtig ist.
    - Weil ich mein Wissen zu dem Thema vertiefen möchte.

- Andere, und zwar:
      - Interest in PLS:
        - “Wobei können Ihrer Meinung nach solche Zusammenfassungen von Forschung, wie Sie sie heute gelesen haben, helfen?” ( 1 = “stimme gar nicht zu”, 8 = “stimme voll zu”)
          - Um Forschung besser zu verstehen
          - Um sich Wissen besser aneignen zu können
          - Um zu erfahren, was Forscher machen oder wie Wissenschaft geht
          - Als Hilfe für Entscheidungen im Alltag
          - Um mit Expertinnen mitreden zu können
          - Um sich eine eigene Meinung bilden zu können
        - Hätten Sie Interesse, in Zukunft weitere solcher Zusammenfassungen zu psychologischen Themen zu lesen? (1 = nein, 2 = ja, 3 = weiß nicht)
- Domain specific justification beliefs (Klopp & Stark, 2016; Ferguson et al., 2013). Participants rate their beliefs in justification by authority and personal justification via three items each on a 1-6 rating scale (1 = “stimmt überhaupt nicht”, 2 = “stimmt weitgehend nicht”, 3 = “stimmt eher nicht”, 4 = “stimmt ein wenig”, 5 = “stimmt weitgehend”, 6 = “stimmt genau”). The additional dimension of justification by multiple sources was excluded for the purpose of this study.
- Muenster Epistemic Trustworthiness Inventory (METI, Hendriks et al., 2015). Participants rate the trustworthiness of the scientists whose work is presented in PLS-format according to 14 adjective pairs on a semantic differential from 1-7 (e.g. 1 = competent, 7 = incompetent) corresponding to the three dimensions Expertise (six items), Integrity (four items) and Benevolence (four Items).