

# Person- and situation-specific factors in discounting science via scientific impotence excuses

Tom Rosman\*<sup>1</sup>, Martin Kerwer<sup>1</sup>, Anita Chasiotis<sup>1</sup> & Oliver Wedderhoff<sup>1</sup>

<sup>1</sup>Leibniz Institute for Psychology Information

## Introduction

The scientific impotence excuse (SIE; Munro, 2010) predicts that „people resist belief-disconfirming scientific evidence by concluding that the topic of study is not amenable to scientific investigation” (p. 597)”. While the incongruence between scientific claims and individual topic-specific beliefs is indeed a central factor in discounting science, research has, up to now, neglected other factors that might lead to scientific impotence excuses. Therefore, the present work aims at extending SIE theory by investigating factors other than belief-disconfirming evidence that may contribute to the devaluation of science. As a situational factor, it will focus on the contradictoriness of the evidence presented, and, as person-specific factor, it will investigate the incongruence between the (epistemic) nature of the scientific evidence and individual epistemic beliefs (i.e., individual beliefs about knowledge).

## General hypotheses<sup>1</sup>

We first aim at replicating the general SIE proposition (Munro, 2010), namely that individuals make use of scientific impotence excuses when the scientific evidence at hand contradicts their prior beliefs. We therefore suggest, based on the corresponding theory by Munro (2010), the following hypothesis:

*Hypothesis 1:* Scientific impotence excuses will be stronger when individuals are confronted with scientific evidence that contradicts their prior topic-specific beliefs – compared to belief-consistent evidence.

Second, we aim at investigating the effects of epistemic beliefs on scientific impotency excuses. In line with the consistency hypothesis by Muis and Franco (2010), a congruence between an individual’s epistemic beliefs and the ‘nature’ of learning materials fosters learning – hence an individual believing in knowledge as tentative and preliminary will learn better with contradictory learning materials than an individual viewing knowledge as an accumulation of absolute ‘truths’. Transferring this idea to the SIE context, we suggest that

---

<sup>1</sup> More specific operational hypotheses can be found in the ‘operational hypotheses’ section.

inconsistencies between epistemic beliefs and the nature of learning materials will lead to scientific impotence excuses – and we suggest that these effects are incremental over ‘traditional’ SIE effects.:

*Hypothesis 2:* Scientific impotence excuses will be stronger when individuals are confronted with scientific evidence that exhibits a ‘nature’ contradicting their prior topic-specific epistemic beliefs – compared to belief-consistent evidence (H2a). These effects persist when controlling for ‘traditional’ SIE effects (i.e., the effects of prior topic-specific beliefs; H2b).

Third, we aim to examine whether the contradictory nature of scientific evidence as such leads to scientific impotence excuses. Earlier research has shown that confronting individuals with conflicting scientific evidence (often labelled as ‘diverging information’) leads to views of scientific evidence as tentative and preliminary (Kienhues, Bromme, & Stahl, 2008; Kienhues, Ferguson, & Stahl, 2016) – in the present research, we aim to test whether this also leads to scientific impotence excuses:

*Hypothesis 3:* Scientific impotence excuses will be stronger when individuals are confronted with contradictory scientific evidence – compared to non-contradictory evidence (H3a). These effects persist when controlling for ‘traditional’ SIE effects (i.e., the effects of prior topic-specific beliefs; H3b).

## **Sampling plan**

### ***Data collection procedures***

A German general population sample will be recruited by means of a commercial panel provider (e.g., Respondi<sup>TM</sup>). Participants complete the data collection using their own device, and will be paid for their participation by the respective panel provider. The sample will have the following properties:

- German speaking participants aged 18-65
- Sex distribution: 50 % males, 50 % females
- Education: at least middle maturity (‘mittlere Reife’)
- No acupuncture treatments in the last 10 years
- No prior beliefs ‘against acupuncture’ (only ‘pro acupuncture’ and ‘no opinion’; see variables section)
- Quotas: Sex distribution (50/50, see above) identical across the ‘pro acupuncture’ and ‘no opinion’ groups (see variables section)

### ***Target Sample Size and Sample Size Calculation***

Target sample size:  $N = 972$

Sample size calculation was performed using GPower 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009). Since the effects regarding Hypothesis 2 might be smaller than ‘traditional’ SIE effects, especially in an online setting, sample size calculation was performed with an expected effect size of  $f = 0.10$  and yielded a total sample size of  $N = 967$  ( $f = 0.10$ ;  $\alpha = .05$ ;  $1 - \beta = .80$ ; 6 conditions). For a better configuration of the quotas (i.e., an optimal distribution of

participants across conditions), the study will be carried out with  $N = 972$  participants (the next largest number divisible by 4 and 6).

## **Variables**

### ***Experimental factors***

Each participant will read 8 texts that describe empirical studies on the efficacy of acupuncture versus massaging in the treatment of back pain. Additionally, four filler texts irrelevant to the topic at hand will be included. All 12 texts will be presented on separate pages in the survey software, and their order will be randomized. The subset of texts that are presented will vary depending on the experimental condition (see below) – the experimental manipulation of the texts is thereby made by interchanging the words ‘acupuncture’ and ‘massaging’. This means, for example, that the studies described in EG1 (‘pro acupuncture’) and EG2 (‘contra acupuncture’) are identical except for the fact that the texts in EG2 suggest that massaging is better than acupuncture. Text length will be around 100 words for each of the 12 texts and all texts will be presented in German language (see Appendix 6). Using these texts, the following experimental conditions will be realized:

- *EG1 (‘pro acupuncture’)*. Participants will be confronted with 8 texts suggesting that acupuncture is better suited for the treatment of back pain than massaging (+ 4 distractor snippets irrelevant to the topic at hand).
- *EG2 (‘against acupuncture’)*. Participants will be confronted with 8 texts suggesting that massaging is better suited for the treatment of back pain than acupuncture (+ 4 distractor snippets irrelevant to the topic at hand).
- *EG3 (‘controversial evidence’)*. Participants will be confronted with 4 ‘pro acupuncture’ and 4 ‘against acupuncture’ texts (+ 4 distractor snippets irrelevant to the topic at hand).

### ***Quasiexperimental factors***

Participants’ topic-specific prior beliefs (i.e., beliefs on the effects of acupuncture versus massaging in the treatment of back pain) will be assessed using a screening question in a forced-choice format at the very beginning of the questionnaire (see Appendix 4). Based on this screening, two equally-sized quasiexperimental groups will be formed: (1) Participants believing in the superiority of acupuncture over massaging (‘pro acupuncture beliefs’ group, QEG1), and (2) participants having no clear opinion on the superiority or inferiority of acupuncture (‘no opinion’ group, QEG2). Sex distribution (50/50, see above) will be identical across the ‘pro acupuncture’ and ‘no opinion’ groups. To keep the design as parsimonious as possible, participants believing in the superiority of massaging will be screened out at the beginning of the data collection.

### ***Continuous independent variables***

*Epistemic beliefs*, which will be used as independent variable regarding Hypothesis 2, will be measured by a slightly adapted version of the CAEB (Stahl & Bromme, 2007), which uses adjective pairs and a five-point Likert scale to assess epistemic beliefs on two subscales: *texture* and *variability*. Scores on individual items will be aggregated to mean scores. Higher scores on the CAEB indicate views on knowledge as more tentative and evolving.

### ***Dependent variables***

*Scientific impotency excuses (SIE)*, the main dependent variable, will be measured using a six-point Likert scale and six newly developed items (see Appendix 3). Since the scale has not been piloted yet, items that produce problems with regard to the scale's reliability or factorial validity may be eliminated prior to hypothesis testing. Subsequently, scores on individual items will be aggregated to mean scores. Higher scores indicate a stronger SIE.

### ***Covariates***

Several *covariates* (e.g., demographics), which are not relevant for the preregistered hypotheses outlined above, will be additionally measured (see Appendix 5). All analyses regarding these covariates are exploratory.

### ***Manipulation checks***

Two scales will be used to test whether the experimental manipulation outlined in section 'Experimental Factors' worked as expected. First, on a content level, a semantic differential single item will be administered to assess which general 'message' the administered texts convey (7-point Likert-scale with the end points 'pro massaging' and 'pro acupuncture'; see Appendix 1). It is expected that EG1 scores higher on this scale than EG2 and EG3 (i.e., as the texts in their condition are more 'pro acupuncture'), that EG2 scores lower than EG1 and EG3 (i.e., that they view the texts as more 'pro massaging'), and that EG3 scores lower than EG1 but higher than EG2 (i.e., that they view 'pro acupuncture' and 'pro massaging' texts as more balanced out).

Furthermore, it is expected that participants in EG3 ('controversial evidence') perceive the findings they read as more inconsistent than the participants in EG1 and EG2. To test this, a slightly adapted scale from Rosman, Mayer, Merk, and Kerwer (2019) will be employed (see Appendix 2). The scale consists of three items relating to the perceived contradictoriness of the administered text materials (see Appendix 2). High scores indicate stronger perceived contradictoriness – hence, it is expected that EG3 has higher scores on this scale than EG1 and EG2.

### ***Design plan***

An experimental between-person design will be realized. In the reading task outlined above (see 'Experimental factors' section), participants will be confronted with short fictitious texts on the treatment of back pain through acupuncture versus massaging. Before and after reading, the dependent variables and covariates will be measured.

### ***Study design***

An experimental 3x2 between-person design will be realized (see Table 1). Participants will be randomly assigned to the three experimental conditions (experimental factor). Moreover, it will be taken care that the two quasiexperimental groups (prior beliefs – quasiexperimental factor) are distributed evenly across the three experimental conditions. Some covariates not relevant for the suggested hypotheses will be measured twice – once before (t1) and once after (t2) reading the texts.

Table 1. Distribution of experimental and quasiexperimental groups

		Experimental factor		
		EG1 (texts ‘pro acupuncture’)	EG2 (texts ‘against acupuncture’)	EG3 (texts ‘controversial evidence’)
Quasiexperimental factor	QEG1 (prior beliefs ‘pro acupuncture’)	<i>n</i> = 162	<i>n</i> = 162	<i>n</i> = 162
	QEG2 (prior beliefs ‘no opinion’)	<i>n</i> = 162	<i>n</i> = 162	<i>n</i> = 162

### ***Blinding***

Subjects will not know the treatment group to which they have been assigned.

### ***Randomization***

Assignment of participants to treatment conditions will be carried out using the randomization function of the survey software Unipark.

### ***Experimental procedure***

The study will be conducted online using the survey software Unipark. All study procedures will be carried out in one single session. After screening and some demographic questions, epistemic beliefs and covariates will be measured, followed by the reading task. After reading, SIE will be measured a second time, followed by some additional questions and covariates.

### **Operational hypotheses**

Based on the design and procedures outlined above, the following set of operational hypotheses is suggested:

*Hypothesis 1:* In the quasiexperimental group QEG1 (prior beliefs ‘pro acupuncture’), SIE will be stronger when participants are confronted with ‘against acupuncture’ texts (EG2) compared to ‘pro acupuncture’ texts (EG1).

*Hypothesis 2a:* In the quasiexperimental group QEG2 (prior beliefs ‘no opinion’), there is an interaction between the experimental factor and the CAEB: With increasing scores on the CAEB (both texture [H2a1] and variability [H2a2]), SIE will be stronger in EG1 and EG2 compared to EG3.

*Hypothesis 2b:* The effects specified in H2a1 and H2a2 (→ H2b1 and H2b2) persist when including all participants and specifying the quasiexperimental factor as a control variable in the respective analyses.

*Hypothesis 3a:* In the quasiexperimental group QEG2 (prior beliefs ‘no opinion’), SIE will be stronger in EG3 compared to EG1 (H3a1) and in EG3 compared to EG2 (H3a2).

*Hypothesis 3b:* The effects specified in H3a1 and H3a2 (→ H3b1 and H3b2) persist when including all participants and specifying the quasiexperimental factor as a control variable in the respective analyses.

## **Analysis plan**

### ***Statistical models***

*Hypothesis 1* will be tested using a *t*-test for independent samples in the QEG1 data subset (Dependent variable: SIE; Independent variable: experimental factor [EG1, EG2]).

*Hypothesis 2a* (both H2a1 and H2a2) will be tested using regression-based interaction testing (Aiken & West, 1991) by means of the PROCESS macro (Hayes, 2013) in the QEG2 data subset (Dependent variable: SIE; Independent variables: experimental factor [EG1, EG2, EG3], CAEB). Separate models will be tested for texture (H2a1) and variability (H2a2).

*Hypothesis 2b* will be tested using the same regression-based interaction testing as in H2a. This time, all participants (except those meeting the exclusion criteria) will be included in the analysis, and the quasiexperimental factor will additionally be included in the analysis, too (Dependent variable: SIE; Independent variables: experimental factor [EG1, EG2, EG3], quasiexperimental factor [QEG1, QEG2], CAEB). Again, separate models will be tested for texture (H2b1) and variability (H2b2).

*Hypothesis 3a* will be tested using a univariate ANOVA in the QEG2 data subset (Dependent variable: SIE; Independent variable: experimental factor [EG1, EG2, EG3]). In case of a significant *F*-test, Tukey HSD post-hoc testing will be employed to test H3a1 and H3a2.

*Hypothesis 3b* will again be tested using a univariate ANOVA. This time, all participants (except those meeting the exclusion criteria) will be included in the analysis, and the quasiexperimental factor will additionally be included in the analysis, too (Dependent variable: SIE; Independent variables: experimental factor [EG1, EG2, EG3], quasiexperimental factor [QEG1, QEG2]). Again, in case of a significant *F*-test, Tukey HSD post-hoc testing will be employed to test H3b1 and H3b2.

Additional exploratory analyses may be conducted.

### ***Inference criteria***

We will use the standard  $p < .05$  criterion for determining if the estimated effects of (dummy coded) intervention components are significantly different from those expected if the null hypothesis was correct. Where appropriate, one-sided hypothesis testing will be performed.

### ***Exclusion criteria***

*Missing, erroneous, or overly consistent responses:* Participants may be excluded from analyses if major protocol deviations occur (e.g., premature study termination, implausibly fast survey completion, selecting the same response category implausibly often). The decision to eliminate such cases will be made prior to analyzing the data.

*Data-based outlier criteria:* Analyses will be performed without univariate outliers (based on *z*-scores; criterion:  $p(z) < .001$ ).

## References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. London: SAGE Publications.
- Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63(1), 1–49. <https://doi.org/10.3102/00346543063001001>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach. Methodology in the social sciences*. New York, NY: Guilford Press.
- 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>
- Kienhues, D., Bromme, R., & Stahl, E. (2008). Changing epistemological beliefs: The unexpected impact of a short-term intervention. *British Journal of Educational Psychology*, 78(4), 545–565. <https://doi.org/10.1348/000709907X268589>
- Kienhues, D., Ferguson, L. E., & Stahl, E. (2016). Diverging information and epistemic change. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of epistemic cognition* (pp. 318–330). London: Routledge.
- Muis, K. R., & Franco, G. M. (2010). Epistemic profiles and metacognition: Support for the consistency hypothesis. *Metacognition and Learning*, 5(1), 27–45. <https://doi.org/10.1007/s11409-009-9041-9>
- Munro, G. D. (2010). The Scientific Impotence Excuse:  $\frac{1}{2}$ Discounting Belief-Threatening Scientific Abstracts. *Journal of Applied Social Psychology*, 40(3), 579–600. <https://doi.org/10.1111/j.1559-1816.2010.00588.x>
- Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The epistemically-related emotion scales. *Cognition & Emotion*, 31(6), 1268–1276. <https://doi.org/10.1080/02699931.2016.1204989>
- Rosman, T., Mayer, A.-K., Merk, S., & Kerwer, M. (2019). On the benefits of ‘doing science’: Does integrative writing about scientific controversies foster epistemic beliefs? *Contemporary Educational Psychology*, 58, 85–101. <https://doi.org/10.1016/j.cedpsych.2019.02.007>
- Stahl, E., & Bromme, R. (2007). The CAEB: An instrument for measuring connotative aspects of epistemological beliefs. *Learning and Instruction*, 17(6), 773–785. <https://doi.org/10.1016/j.learninstruc.2007.09.016>

## **Appendix**

### ***Appendix 1: Manipulation check – content***

Insgesamt legen die präsentierten Befunde nahe, dass ...  
Massage ... .. Akupunktur  
... besser zur Behandlung von Rückenschmerzen geeignet ist.

Response format: bipolar 7-point Likert scale

Time of measurement: t2 – after reading



## ***Appendix 2: Manipulation check – perceived contradictoriness***

- Die Befunde zum Thema “Akupunktur vs. Massage” kamen mir sehr widersprüchlich vor.
- Die Studienergebnisse zum Thema “Akupunktur vs. Massage” erschienen mir logisch unvereinbar.
- Die dargestellte Befundlage zum Thema “Akupunktur vs. Massage” wirkte auf mich sehr unstimmig.

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Time of measurement: t2 – after reading

### *Appendix 3: SIE scale*

Ob Akupunktur besser zur Behandlung von Rückenschmerzen geeignet ist als Massagen...

- ... lässt sich mit wissenschaftlichen Methoden gar nicht untersuchen.
- ... wird der Wissenschaft auch in Zukunft verborgen bleiben.
- ... muss jeder für sich selbst herausfinden – die Wissenschaft kann darauf keine Antworten liefern.
- ... ist einer wissenschaftlichen Analyse nicht zugänglich.
- ... können nur praktizierende Ärzte und Ärztinnen durch ihre Erfahrung beurteilen – nicht die Wissenschaft.
- ... hängt von so vielen verschiedenen Einflussgrößen ab, dass die Wissenschaft darauf keine Antwort finden kann.

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Time of measurement: t1 and t2 – before and after reading

#### *Appendix 4: Screening question on prior beliefs*

In letzter Zeit wird häufiger diskutiert, ob Akupunktur besser zur Behandlung von Rückenschmerzen geeignet ist als vergleichbare konventionelle Therapien (z. B. Massagen). Was glauben Sie?

- Akupunktur ist zur Therapie von Rückenschmerzen besser geeignet als Massagen.
- Akupunktur ist zur Therapie von Rückenschmerzen schlechter geeignet als Massagen.
- Ich habe zu diesem Thema keine klare Meinung.

Response format: Forced choice format – only one category may be selected

Time of measurement: t1 – before reading

## ***Appendix 5: Covariates***

**Comprehension of text content** (administered once for each text):

Die Studie lässt folgenden Schluss zu ...

- “Akupunktur ist besser als Massage”
- “Massage ist besser als Akupunktur”
- “Weder/Noch: Die Studie geht nicht auf den Vergleich zwischen Akupunktur und Massage ein”

Response format: Forced choice format – only one category may be selected

Time of measurement: after each text

**Ignoring anomalous data** (alternative dependent variable for exploratory analyses; Chinn & Brewer, 1993)

- Eigentlich möchte ich gar nicht wissen, was die Wissenschaft zu diesem Thema sagt.
- Die wissenschaftliche Sichtweise zu diesem Thema interessiert mich ehrlich gesagt nicht.
- Ich vertraue meiner Intuition und Erfahrung - nicht der Wissenschaft.
- Der Standpunkt der Wissenschaft zu dem Thema hat für mich keine Bedeutung.
- Es ist mir egal, was in wissenschaftlichen Studien zu diesem Thema gefunden wurde.

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Time of measurement: t2 – after reading

### **Reasons for distrust:**

(filter question – only administered if mean SIE at t2 is larger than  $M = 3.50$ )

Sie haben angegeben, dass die Wissenschaft eher keine Antwort darauf liefern kann, ob Akupunktur besser zur Behandlung von Rückenschmerzen geeignet ist als Massagen. Wir interessieren uns dafür, warum Sie zu dieser Einschätzung kommen. Bitte geben Sie uns eine kurze Einschätzung:

Subscale ‘epistemic / topic-related reasons’

- Die komplexen Wirkmechanismen von Akupunktur sind der kalten, nüchternen Wissenschaftslogik nicht zugänglich.
- Die Methoden der Wissenschaft sind zu schlecht, um ein komplexes Thema wie die Akupunktur zu untersuchen.
- Da Akupunktur ganzheitlich wirkt, ist eine isolierte wissenschaftliche Betrachtung einzelner Wirkungen nicht möglich.

Subscale ‘research-related reasons’ (derived from Hendriks, Kienhues, & Bromme, 2015)

#### Subsubscale “Expertise”

- Wissenschaftler/-innen machen häufig Fehler. Deswegen kann die Wissenschaft keine klaren Antworten liefern.
- Wissenschaftler/-innen sind zu spezialisiert und blicken nicht über den Tellerrand. Deswegen kann die Wissenschaft keine klaren Antworten liefern.
- Wissenschaftliche Untersuchungen sind zu alltagsfern. Deswegen lassen sich wissenschaftliche Ergebnisse nicht auf konkrete Fragestellungen anwenden.

#### Subsubscale “Integrity”

- Wissenschaftler/-innen sind stark abhängig von ihren Geldgebern. Deswegen kann die Wissenschaft keine klaren Antworten liefern.
- Wissenschaftliche Erkenntnisse, die der gängigen Auffassung widersprechen, werden häufig unter den Teppich gekehrt. Deswegen kann die Wissenschaft keine klaren Antworten liefern.
- Wissenschaftler/-innen passen ihre Ergebnisse oft ihren eigenen Erwartungen an. Deswegen kann die Wissenschaft keine klaren Antworten liefern.

#### Subsubscale “Benevolence”

- Wissenschaftler/-innen sind hauptsächlich an ihrer eigenen Karriere interessiert - sie übernehmen keine gesellschaftliche Verantwortung. Deswegen kann die Wissenschaft keine klaren Antworten liefern.
- Wissenschaftler/-innen verfolgen ihre eigenen Ziele - die Suche nach der Wahrheit gerät zur Nebensache. Deswegen kann die Wissenschaft keine klaren Antworten liefern.
- Wissenschaftler/-innen haben keine ethischen Prinzipien. Deswegen kann die Wissenschaft keine klaren Antworten liefern.

Weitere Gründe (falls zutreffend): \_\_\_\_\_

Time of measurement: t2 – after reading

### **Epistemic emotions**

(see Pekrun, Vogl, Muis, & Sinatra, 2017, Introduction pertains to reading of the texts)

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Time of measurement: t2 – after reading

### **Strength of prior beliefs**

- Mit Akupunktur lassen sich Rückenschmerzen deutlich reduzieren.
- Akupunktur ist gut geeignet, um Rückenprobleme in den Griff zu kriegen.
- Bevor man wegen Rückenschmerzen ärztliche Hilfe aufsucht, sollte man ausprobieren, ob sich die Beschwerden möglicherweise durch Akupunktur lindern lassen.
- Akupunktur führt zu einer langfristigen und dauerhaften Verringerung von Rückenschmerzen.
- Akupunktur ist gefährlich, da sie Personen mit Rückenproblemen davon abhalten kann, eine adäquate Therapie zu beginnen. (-)
- Akupunktur sollte nur bei leichten Rückenschmerzen angewandt werden - bei stärkeren Beschwerden sollte man ärztliche Hilfe aufsuchen. (-)
- Akupunktur dient dazu, Personen mit Rückenschmerzen das Geld aus der Tasche zu ziehen. (-)
- Akupunktur sollte nur dann zum Einsatz kommen, wenn alle konventionellen Rückenschmerz-Therapien erfolglos sind. (-)

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Time of measurement: t1 – before reading

### **Trust in science**

- Ich habe wenig Vertrauen in die Wissenschaft.
- Erkenntnisse von Wissenschaftler/innen sind vertrauenswürdig.
- Ich vertraue darauf, dass Wissenschaftler/-innen das Richtige tun.
- Wissenschaftler/-innen sagen der Öffentlichkeit oft nicht die Wahrheit.
- Ich finde das gegenwärtige Wissenschaftssystem fragwürdig.

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Time of measurement: t1 and t2 – before and after reading

### **Beliefs in conspiracy theories:**

Nun folgen einige Aussagen, die von den etablierten Medien mehr oder weniger stark als "unwahr" oder "Verschwörungstheorien" angesehen werden. Bitte geben Sie uns eine Einschätzung, wie Sie zu diesen Themen stehen.

Bitte bedenken Sie: In dem Fragebogen gibt es keine objektiv "richtigen" oder "falschen" Antworten. Wir sind lediglich an Ihrer persönlichen Meinung interessiert.

- Der Klimawandel ist nicht menschengemacht.
- Impfungen sind schädlich.
- Die amerikanische Regierung verschleiert die Landung von Außerirdischen.
- Die Homöopathie kann viele Krankheiten heilen.
- HIV wurde in biochemischen Labors erschaffen.
- Die Erde ist in Wirklichkeit flach (statt kugelförmig).
- Es waren noch nie Menschen auf dem Mond.

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu

Time of measurement: t2 – after reading

### **Text comprehensibility and study evaluation:**

Die 12 Texte zum Thema "Behandlung von Rückenschmerzen" ...

- ... waren zu kompliziert.
- ... enthielten zu viel Fachjargon.
- ... waren schwer verständlich.
- ... waren langweilig.
- ... erschienen mir wenig glaubwürdig.
- Die einzelnen Texte waren zu lang.
- Es waren zu viele Texte.
- Die Studie war insgesamt zu lang.

Falls Sie noch weitere Anmerkungen zu dieser Studie haben, können sie sie hier eingeben:

---

Response format: 6-point Likert scale ranging from „stimme überhaupt nicht zu“ to „stimme voll und ganz zu“

Die Inhalte dieser Studie weisen für mich eine hohe persönliche Relevanz auf.

Response format: 6-point Likert scale ranging from „lehne vollständig ab“ to „stimme vollständig zu“

Time of measurement: t2 – at the very end of the survey



## Appendix 6: Text materials

**Instruction:** Auf den nächsten Seiten erhalten Sie insgesamt 12 kurze Studienbeschreibungen zum Thema "Behandlung von Rückenschmerzen". Die genannten Studien basieren auf tatsächlichen Befunden aus der medizinischen Forschung und spiegeln den Forschungsstand korrekt wider. Allerdings haben wir diverse Aspekte der Studien vereinfacht dargestellt. Beispielsweise werden in den Beschreibungen immer nur zwei Behandlungsmethoden miteinander verglichen, während die Originalstudien teilweise deutlich komplexer sind.

Bitte lesen Sie jede Studienbeschreibung genau durch und entscheiden Sie dann, ob die Studie suggeriert, dass

- Akupunktur besser zur Behandlung von Rückenschmerzen geeignet ist als Massage
- Massage besser zur Behandlung von Rückenschmerzen geeignet ist als Akupunktur
- die Studie nicht auf den Vergleich zwischen Akupunktur und Massage eingeht (also eine ganz andere Forschungsfrage hat)

<b>Krankheitstage</b>	<p>Eine Studie von Collins und Kollegen (2017) untersuchte die Wirkung von Akupunktur bei N = 149 berufstätigen Patienten und Patientinnen mit chronischen Rückenschmerzen. Dabei erhielten 75 Versuchspersonen im Zeitraum von 10 Wochen 6 Akupunktur-Behandlungen. Die restlichen 74 Personen erhielten in demselben Zeitraum insgesamt 6 Rückenmassagen. Alle Behandlungen fanden in derselben Praxis statt. Eine Patientenbefragung nach sechs Monaten zeigte, dass die Personen der Akupunkturgruppe nach der Therapie etwa 30 Prozent weniger Fehltage bei der Arbeit verzeichneten. In der Massagegruppe reduzierten sich die Fehltage nur um 8 Prozent. Auf Grundlage dieser Daten sehen die Forschenden es als erwiesen an, dass Akupunktur bei chronischen Rückenschmerzen effektiver ist als Massage.</p> <p>Eine amerikanische Forschergruppe (Huber et al., 2017) verglich die Wirkung von Akupunktur und Massagetherapie bei Personen mit chronischen Rückenbeschwerden. N = 142 berufstätige Versuchspersonen beteiligten sich an der Studie. Sie erhielten entweder eine über 10 Wochen andauernde Akupunkturbehandlung (72 Personen), oder aber eine gleichlange Massagetherapie (70 Personen). Nach einem halben Jahr wurden die Versuchspersonen dazu befragt, wie häufig sie nach Therapieende aufgrund von Rückenschmerzen krankgeschrieben waren. Dabei berichteten die Personen der Akupunkturgruppe deutlich weniger Krankheitstage als die Massagetherapie-Personen. Haynes und Kollegen schließen aus ihren Ergebnissen, dass bei chronischen Rückenschmerzen Akupunktur effektiver ist als Massage.</p>
<b>subj. Schmerz</b>	<p>Andersen et al. (2019) führten eine klinische Studie durch, um die Wirksamkeit von verschiedenen Behandlungsmethoden bei Rückenschmerzen zu erfassen. Über verschiedene orthopädische Praxen rekrutierten sie insgesamt N = 159 Rückenschmerzpatienten und -patientinnen. Diese wurden im Anschluss entweder mit Akupunktur oder mit Massagen behandelt. Vor und nach der sechswöchigen Behandlungsphase wurde die Schmerzintensität der Versuchspersonen auf einer numerischen Ratingskala (NRS) abgefragt. Der Prä-Post-Vergleich zeigt, dass die Schmerzintensität bei</p>

	<p>beiden Gruppen abnahm. Jedoch war die Schmerzreduktion lediglich bei der Akupunkturgruppe klinisch bedeutsam; bei mit Massagen therapierten Personen war sie sehr gering. Anderson et al. schlussfolgern, dass Akupunktur besser zur Therapie von Rückenschmerzen geeignet ist als Massagen.</p> <p>In einer klinischen Studie von Michelsen et al. (2019) wurden Akupunktur- und Massagebehandlungen hinsichtlich ihrer Wirksamkeit bei der Therapie von Rückenschmerzen verglichen. Zu Beginn der Studie schätzten die N = 158 teilnehmenden Versuchspersonen ihre Schmerzintensität ein. Dazu wurde eine Skala zur Selbsteinschätzung der Schmerzbeschwerden (VAS) genutzt. Anschließend wurden die Versuchspersonen entweder einer Akupunktur- oder einer Massagebehandlung (Gesamtdauer: 6 Wochen) unterzogen. Nach sechs Behandlungswochen wurde die Schmerzintensität (VAS-Skala) erneut abgefragt. Es zeigte sich, dass es in der Akupunkturgruppe zu einer klinisch bedeutsamen Schmerzverringerung kam – die Massagegruppe wies hingegen nur eine sehr geringe Schmerzreduktion auf. Daraus folgern Michelsen et al., dass Akupunktur bei Rückenschmerzpatienten deutliche Vorteile gegenüber Massagen hat.</p>
<b>subj. Wohlbefinden</b>	<p>Kanadische Forschende (Brian et al., 2018) testeten unterschiedliche Therapieansätze gegen Schmerzen im unteren Rückenbereich. Insgesamt nahmen N = 143 Patienten und Patientinnen mit chronischen Rückenschmerzen an der Studie teil. Davon erhielt die Hälfte einen Monat lang jede Woche eine 45-minütige Massage; die andere Hälfte erhielt jede Woche 45 Minuten Akupunktur. Die Behandlungsdauer war bei beiden Therapien identisch (insgesamt 4 Sitzungen). Nach Abschluss der Behandlungen wurden Einschätzungen über das subjektive Wohlbefinden der Versuchspersonen (sog. WHO-5 Well-Being Index) vor und nach der Behandlungsphase verglichen. Die Ergebnisse zeigen, dass im Rahmen der Massagetherapie das subjektive Wohlbefinden weder zu- noch abnahm. Mit Akupunktur therapierte Personen erfuhren hingegen eine Steigerung ihres subjektiven Wohlbefindens. Auf dieser Grundlage schlussfolgern die Forschenden, dass Akupunktur bei Schmerzen im unteren Rücken sinnvoller ist als Massagen.</p> <p>Witt et al. (2018) untersuchten im Rahmen einer klinischen Studie, wie man Menschen mit Schmerzen im unteren Rücken helfen kann. Dabei widmeten sie sich insbesondere der Frage, inwiefern Akupunktur- und Massagebehandlungen unterschiedliche Effekte auf das Wohlbefinden haben. N = 159 Versuchspersonen wurden für einen Monat entweder einer wöchentlichen Akupunkturbehandlung, oder aber einer wöchentlichen Massagebehandlung unterzogen. Mit 45 Minuten pro Sitzung war die Behandlungsdauer bei beiden Therapien exakt gleich. Das Wohlbefinden wurde vor und nach der Therapie mithilfe eines standardisierten Fragebogens, dem WHO-5 Well-Being Index, gemessen. Es zeigte sich, dass es bei Personen mit Massagebehandlung zu keinerlei Veränderungen im Wohlbefinden kam, während eine Akupunkturbehandlung zu einer Steigerung des</p>

	<p>subjektiven Wohlbefindens führte. Aus diesem Grund folgern Witt et al., dass Akupunktur besser zur Therapie von Schmerzen im unteren Rücken geeignet sei als Massagen.</p>
<b>Mobilität</b>	<p>In einer klinischen Studie zur Behandlung von Rückenschmerzen (Hofer, 2016) wurden die Therapiemethoden „Akupunktur“ und „Massage“ verglichen. Insgesamt nahmen N = 172 Patienten und Patientinnen teil. Davon wurden 90 Personen einer 30-minütigen Akupunkturbehandlung unterzogen und 82 Personen 30 Minuten lang massiert. Eine ausgewählte Orthopädin führte vor und nach der Behandlung eine Bewertung über die Mobilität (also die Beweglichkeit) der Patienten durch. Über die Studie hinweg verbesserte sich die Mobilität in der Akupunktur-Gruppe im Schnitt deutlich stärker als die Mobilität in der Massage-Gruppe. Deswegen sehen die Forschenden es als erwiesen an, dass Akupunktur bei der Behandlung von Rückenschmerzen sinnvoller ist als Massage – zumindest was die Beweglichkeit angeht.</p> <p>Ulrich et al. (2016) untersuchten die Wirksamkeit unterschiedlicher Rückenschmerz-Therapien. Dazu untersuchten sie N = 132 Patienten und Patientinnen. Diese durften entweder an einer halbstündigen Massage oder an einer halbstündigen Akupunkturbehandlung teilnehmen. Vor und nach der Behandlung wurde die Beweglichkeit der Versuchspersonen von einer ausgewählten Orthopädin gemessen. So konnten die Forscher zeitliche Veränderungen in der Beweglichkeit der Teilnehmenden untersuchen. Im Rahmen der Analysen zeigte sich, dass sich die Beweglichkeit der mit Akupunktur behandelten Personen viel stärker verbesserte als die der Massagetherapie-Versuchspersonen. Laut Ulrich et al. (2016) sprechen die Ergebnisse für die Wirksamkeit von Akupunktur und belegen darüber hinaus, dass Akupunktur die Beweglichkeit besser fördert als Massagen.</p>
<b>Kontrollsnippet Novadolor</b>	<p>In einer Studie der Universität Ulm (Schilz et al., 2019) wurde die Wirksamkeit unterschiedlicher Schmerztherapien an N = 137 Rücken- und Kopfschmerzpatient/-innen untersucht. Zu Beginn der Studie wurde die Schmerzintensität mit einer Skala zur Selbsteinschätzung der Schmerzbeschwerden (VAS) erfasst. Dann wurden die Patienten und Patientinnen zwei Gruppen zugeteilt: Personen der ersten Gruppe erhielten über zwei Monate ein neuartiges schmerzstillendes Medikament (NovaDolor 30mg retard), während Personen der zweiten Gruppe eine zweimonatige Psychotherapie (Klassische Psychoanalyse nach Freud) absolvierten. Es zeigte sich, dass die Schmerzsymptome in der Medikamenten-Gruppe im Vergleich zur Psychotherapie-Gruppe stärker abnahmen. Daraus schließen die Schilz et al., dass Novadolor 30mg retard besser zur Schmerztherapie geeignet ist als eine Psychoanalyse.</p>

<b>Kontrollsnippet Kraftsport</b>	In einer Studie des Instituts für Orthopädie der Universität Mannheim (Wesner et al., 2018) wurde untersucht, inwiefern sich Rückenschmerzen bei Kraftsportlerinnen reduzieren lassen. Dazu wurden in verschiedenen Fitnessstudios insgesamt N = 188 Versuchspersonen rekrutiert. Die Hälfte dieser Personen wurde einem speziellen Hüftbeuger-Dehnprogramm unterzogen. Die andere Hälfte der Versuchspersonen erhielt eine Bewegungsschulung, unter anderem zur Optimierung der Ausführung von Kraftsport-Übungen. Mithilfe von MRT-Untersuchungen wurde nach 6 Monaten der Wirbelsäulenzustand der Versuchspersonen gemessen. Allerdings zeigten sich dabei keine Unterschiede zwischen den beiden Gruppen: Der durchschnittliche Wirbelsäulenzustand war in beiden Gruppen (Dehnprogramm vs. Bewegungsschulung) etwa gleich. Aus diesem Grund sehen die Forschenden es als nicht abschließend geklärt an, inwiefern die jeweiligen Maßnahmen einen Effekt auf die Rückengesundheit haben können.
<b>Kontrollsnippet TENS</b>	Weber et al. (2017) untersuchten, inwiefern Rücken- und Kopfschmerzpatient/-innen unterschiedlich auf eine sogenannte TENS-Behandlung reagieren. Bei der TENS-Behandlung werden Stromimpulse über Hautelektroden an die Nerven geleitet, was die Signalweiterleitung ins Gehirn und damit die Schmerzwahrnehmung reduzieren soll. Im Rahmen ihrer Studie wurden insgesamt N = 170 Versuchspersonen in verschiedenen orthopädischen Praxen einer TENS-Behandlung unterzogen. Etwa die Hälfte davon (83 Personen) hatte Rückenbeschwerden, die andere Hälfte (87 Personen) war wegen Knieschmerzen in Behandlung. In den 3 Wochen vor und nach der TENS-Behandlung wurde die Häufigkeit der Medikamenteneinnahme mittels eines standardisierten Fragebogens (KFM - Kurzfragebogen für Medikamentengebrauch) erfasst. Allerdings zeigte sich weder bei Rücken- noch bei Knieschmerzpatient/-innen eine klinisch relevante Abnahme des Medikamentengebrauchs. Daraus schließen die Forschenden, dass TENS möglicherweise nur kurzzeitig wirksam ist.
<b>Kontrollsnippet Thermotherapie</b>	In einer klinischen Studie untersuchten Willach et al. (2016) Haltungsveränderungen von N = 129 Rückenschmerzpatienten und -patientinnen in Abhängigkeit der Behandlungsmethode. Verglichen wurden dabei die Behandlungsmethoden „Thermotherapie“ und „Gymnastik“. Die 80 Personen der Thermotherapie-Gruppe erhielten über einen Zeitraum von 8 Wochen regelmäßige Wärme- und Kältebehandlungen, während die 49 Personen der Gymnastik-Gruppe einem 8-wöchigen Aquajogging-Aufbautraining unterzogen wurden. Ein speziell geschulter Orthopäde führte vor und nach der Therapie bei allen Versuchspersonen eine Haltungsdiagnostik durch. Bei Personen, die am Aquajogging teilgenommen hatten, konnte dabei eine klinisch relevante Haltungsverbesserung nachgewiesen werden. In der Thermotherapie-Gruppe kam es hingegen zu keiner Haltungsveränderung. Die Forschenden schlussfolgern, dass eine Haltungskorrektur eher durch Gymnastik erzielt werden kann.

*Note:* Depending on the experimental condition, the texts will be altered as specified in ‘Experimental factors’ section.