

1 Testing the Usability of the Psychological Research Preregistration-Quantitative (PRP-QUANT)
2 Template

3 Lisa Spitzer¹, Michael Bosnjak², & Stefanie Mueller¹

4 ¹ Leibniz Institute for Psychology (ZPID), Universitaetsring 15, 54296 Trier, Germany

5 ² Trier University, Universitaetsring 15, 54296 Trier, Germany

6

7

8

9

10 Author note

11 Lisa Spitzer (ls@leibniz-psychology.org), <https://orcid.org/0000-0002-4925-7291>

12 Michael Bosnjak, <https://orcid.org/0000-0002-1431-8461>

13 Stefanie Mueller, <https://orcid.org/0000-0002-3611-6190>

14

Abstract

15 The Psychological Research Preregistration-Quantitative (PRP-QUANT) Template
16 provides researchers with a comprehensive list of elements to consider when planning a
17 psychological study. We assessed its usability and researchers' intention to use it.

18 We conducted a usability test (study 1) and surveyed researchers who submitted or reviewed
19 a preregistration created with the template (study 2, authors: $N = 19$, reviewers: $N = 29$) regarding
20 their impression of the template. For the usability test, we recruited participants via the mailing
21 lists of the German Psychological Society, the American Psychological Association, and the British
22 Psychological Society, and social media. Participants answered selected template and web probing
23 items and provided an overall rating ($N = 88$). Based on the Unified Theory of Acceptance and Use
24 of Technology (UTAUT), we expected that the intention to use the template is influenced by
25 performance expectancy (moderated by age), effort expectancy (moderated by age and experience),
26 and social influence (moderated by age, experience, and voluntariness, $N = 60$).

27 The results suggest that the PRP-QUANT Template is suitable for different research areas
28 within psychology, is evaluated as effective, and perceived positively. Performance expectancy
29 and all predictors combined significantly predicted researchers' intention to use the template.

30 *Keywords:* meta-research, open science, preregistration, reproducibility, replicability

31 *Word count:* 10.473

32

Introduction

33 In recent years, there has been a growing call for methods and procedures to increase the
34 transparency of research (e.g., see [1]), one of them being study preregistration [2], in which
35 researchers document and publish their study plan before data have been collected or examined.
36 Preregistrations are time-stamped and published with an independent party (e.g., a repository) so
37 that they can be accessed by others (possibly after an embargo period). This way, preregistration
38 aims to provide transparent documentation of study procedures, clear identification of deviations
39 from preregistered plans, and a clear distinction between confirmatory and exploratory research
40 [3].

41 Studies show that preregistration indeed helps reduce questionable research practices and
42 the rate of false positive findings [4,5] and, among other open science techniques, increases the
43 replication rate of research [6]. However, despite its benefits, several obstacles currently prevent
44 researchers in psychology from using preregistration. A recent study showed that uncertainty about
45 which aspects needed to be included in the preregistration was both a concern of researchers who
46 had not yet preregistered and a problem experienced by researchers with preregistration experience
47 [7]. Accordingly, better education about preregistration was one of the most common suggestions
48 to increase motivation and reduce obstacles of preregistration [7].

49 Preregistration templates can help overcome uncertainty by listing important elements that
50 researchers should address in their preregistration (e.g., hypotheses, study design, data acquisition,
51 and data analysis plan). Nowadays, a variety of templates are available, differing in scope and
52 targeted research type. Besides more universal templates, there are templates specifically focusing
53 on social psychology [8], fMRI studies [9], replication studies [10], cognitive models [11], or

54 secondary data analyses [12]. This wide range of options may lead to fragmentation and potential
55 confusion among researchers as to which template should be used.

56 To develop a common standard for psychology and reinforce the importance of
57 preregistration, members of the American Psychological Association (APA), the British
58 Psychological Society (BPS), the German Psychological Society (DGPs), the Center for Open
59 Science (COS, <https://www.cos.io/>), and the Leibniz Institute for Psychology (ZPID,
60 <https://leibniz-psychology.org/en/>) formed a Joint Psychological Societies Preregistration Task
61 Force. Together, they developed the Psychological Research Preregistration-Quantitative (PRP-
62 QUANT) Template [13], a comprehensive template that aids the preregistration of quantitative
63 studies in psychology.

64 **Testing the usability of the PRP-QUANT Template**

65 Since preregistration of research in psychology is mainly voluntary, it is essential to provide
66 a good usability to enhance acceptance. There is currently little empirical evaluation of
67 preregistration templates (for an example, see [14,15]), however, it is reasonable to not simply
68 assume usability but to test it empirically. Thus, in line with the PRP-QUANT Template's goal of
69 becoming increasingly adapted to the needs of the psychological research community (see [13]),
70 the first aim of our studies was to evaluate its usability and identify areas for improvement.

71 Various definitions of usability exist, largely sharing the same underlying concepts, but
72 highlighting different aspects. A popular definition comes from the International Organization for
73 Standardization (ISO), which measures usability along the dimensions of effectiveness, efficiency,
74 and satisfaction with regard to specific users, objectives and contexts [16]. Another definition is
75 provided by Shackel [17], who defines usability as “the capability to be used by humans easily and

76 effectively” (p. 340). According to Shackel [17], four aspects of usability should be considered:
77 learnability, flexibility, effectiveness, and attitude. These partly align with the ISO standard but
78 place more emphasis on learnability and flexibility. As the PRP-QUANT Template is intended to
79 cover a wide range of different psychological sub-disciplines, and therefore the issues of
80 learnability and flexibility are particularly relevant for assessing the template’s usability within all
81 of psychology, the aspects defined by Shackel [17] were used as the basis for this research.

82 Specifically, we were interested in the following questions:

83 A) *Learnability*: Do authors from the various sub-disciplines of psychology understand how to fill
84 in the different items of the template? Do they understand the items in the same way?

85 B) *Flexibility*: Does the template capture the main points across sub-disciplines, as indicated by
86 researchers from different sub-disciplines?

87 C) *Effectiveness*: Are the items specific enough (i.e., are researcher degrees of freedom
88 minimised)? Are items answered as expected (i.e., is the information requested in the item
89 provided by researchers in response to it)?

90 D) *Attitude*: Are users satisfied with using the template? Are costs (e.g., tiredness, personal effort)
91 acceptable? Can the goals of the template (i.e., a detailed mapping of the preregistered study)
92 be achieved with a reasonable amount of effort? Would authors recommend/use the template?

93 To assess the usability of the PRP-QUANT Template, we conducted an online study in
94 which we asked psychological researchers to think about one of their studies and create a
95 preregistration for that study using the template (see *study 1: simulation trial and intention to use*).
96 Participants did not actually submit their preregistration. Alongside the template items, several web

97 probing questions were presented. We wanted to assess the overall perceived usability, as well as
98 participants' comments and suggestions for improving individual items.

99 In addition, we conducted a survey among researchers who responded to a call for online
100 studies by submitting a preregistration created with the PRP-QUANT Template (see *study 2:*
101 *survey of preregistration authors and reviewers*). Responding to this call, researchers applied with
102 their preregistrations for funding for their data collection from ZPID's service PsychLab ONLINE.
103 PsychLab aims to encourage preregistration by offering the incentive of free-of-charge data
104 collection for high-quality preregistrations, which addresses another current obstacle to
105 preregistration, i.e., insufficient incentives [7]. The submitted preregistrations were evaluated by
106 external peer-reviewers. After the peer-reviews were completed, the authors of preregistrations
107 (i.e., the applicants) and the reviewers were surveyed about using the PRP-QUANT Template for
108 writing and reviewing, respectively.

109 **Intention to use the template**

110 In addition to exploring the usability of the preregistration template, we also wanted to find
111 out whether psychological researchers plan to use the template in the future to create their
112 preregistrations. To investigate this, the theoretical framework of the Unified Theory of Acceptance
113 and Use of Technology (UTAUT) [18,19] was used. This theory postulates that performance
114 expectancy (i.e., the belief that using the system will help achieve performance gains), effort
115 expectancy (i.e., the degree of ease associated with using the system), and social influence (i.e., the
116 perception that important others believe one should use the new system) predict people's intention
117 to use a new system. According to the UTAUT, intention, combined with facilitating conditions
118 (i.e., the belief that an organisational and technical infrastructure exists which supports using the
119 system), is a predictor of actual behaviour. Therefore, it is of interest to investigate the intention to

120 use the template in more detail, as this might be an estimator of how likely psychological
121 researchers will use it in the future to create their preregistrations.

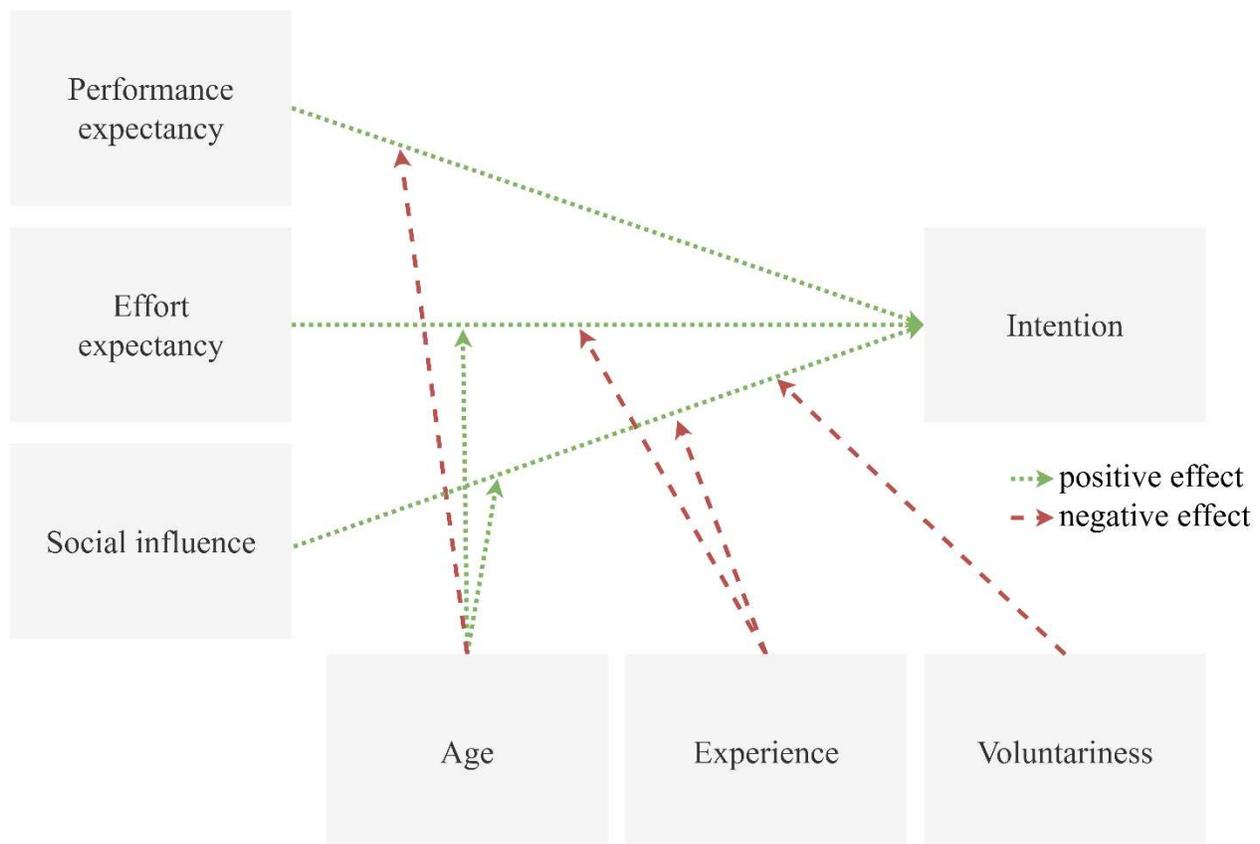
122 To examine the intention to use the PRP-QUANT Template, we asked the participants of
123 study 1 to answer various UTAUT items (see *study 1: simulation trial and intention to use*). Then,
124 we computed a moderated multiple regression model. Based on the UTAUT and the effects
125 described in [18], we had the following predictions, which are also displayed in Figure 1:

- 126 1) Performance expectancy is a positive predictor for the intention to use the template.
- 127 2) Effort expectancy is a positive predictor for the intention to use the template.
- 128 3) Social influence is a positive predictor for the intention to use the template.
- 129 4) Age negatively moderates the effect of performance expectancy on intention, as it has
130 been shown that extrinsic rewards may be more important for younger persons.
- 131 5) Age positively moderates the effect of effort expectancy on intention, since older
132 persons have more difficulties in processing complex stimuli and attention allocation.
- 133 6) Age positively moderates the effect of social influence on intention, since older persons
134 might place more importance on social influences and affiliation.
- 135 7) Experience negatively moderates the effect of effort expectancy on intention, as prior
136 experience would serve as a facilitator for using the new system.
- 137 8) Experience negatively moderates the effect of social influence on intention, since it has
138 been shown that the salience of social influences decreases with experience.

139 9) Voluntariness negatively moderates the effect of social influence on intention, as social
 140 influence is less important in settings where the decision to use the system is completely
 141 voluntary.

142 **Figure 1**

143 *Hypotheses based on the UTAUT [18,19]*



144

145

146 **Study 1: Simulation trial and intention to use**

147 In study 1, we examined the usability of the PRP-QUANT Template by asking
 148 psychological researchers to think of one of their studies and complete selected parts of the

149 template. We also presented web probing questions and examined the researchers' intention to use
150 the template in the future.

151 **Methods**

152 This study was preregistered [20]. It was conducted as preregistered, except for the
153 deviations summarised and justified in the section *deviations from the preregistration*.

154 ***Participants***

155 Participants were invited via the mailing lists of all research-oriented APA
156 (<https://www.apadivisions.org/>), BPS (<https://www.bps.org.uk/member-networks>), and DGPs
157 (<https://www.dgps.de/fachgruppen>) divisions. A reminder was sent a few weeks after the initial
158 invitation. Furthermore, the survey was advertised on social media (Facebook and Twitter).
159 Participants were not compensated.

160 Of the 2668 persons that clicked on the study link, 314 provided informed consent and
161 started the main body of the study. Nine participants who indicated that they were not researchers
162 or that their research did not fall within the scope of psychology were screened out at the beginning
163 of the study since we specifically aimed to collect data from psychological researchers. Of the
164 remaining participants, 88 subsequently worked on the template items and were thus included in
165 the descriptive reports, as this was the core part of the study (34.09% male, 57.95% female, 2.27%
166 other, 5.68% preferred not to answer; $Mean_{age} = 37.19$; 86.36% from Europe, 11.36% from North
167 America, 2.27% did not respond; 13.64% native English speakers). We were able to collect data
168 from all targeted academic groups, that is, 32.95% of participants were PhD students, 37.5% were
169 postdocs, 26.14% were professors, 2.27% indicated "other", and 1.14% did not respond.
170 Additionally, all the considered research areas were present (see Appendix B in the electronic
171 supplementary material). More than half of the participants indicated having preregistered before

172 (56.82%). Of the participants with preregistration experience ($n = 50$), 16% had preregistered one
173 study, 14% had preregistered two studies, 16% three studies, 8% four studies, 12% five studies,
174 and 34% more than five studies.

175 Of all participants included in the descriptive reports, 60 answered all items relevant for the
176 UTAUT model and were therefore included in the hypotheses tests (36.67% male, 56.67% female,
177 3.33% other, 3.33% preferred not to answer; $Mean_{age} = 36.12$; 33.33% PhD students, 41.67%
178 postdocs, and 25% professors; 88.33% from Europe, 10% from North America, 1.67% did not
179 respond; 13.33% native English speakers).

180 Data were collected between March 1, 2021, and April 24, 2021. As preregistered, data
181 collection was stopped one month after the initial invitation was sent to the last contacted division.
182 Originally, a sample size of $N = 89$ was targeted to be able to detect effects of $R^2 = 25\%$ with $\alpha =$
183 $\beta = .05$ for the UTAUT regression model including 12 predictors, which was determined by an a
184 priori power analysis (see preregistration [20]). This sample size was not reached within the set
185 timeframe. However, the effect found in our study was quite large ($R^2_{adjusted} = 42.79\%$) and could
186 thus also be detected with the achieved sample size.

187 ***Material and measures***

188 The online survey was created using the software SoSci Survey [21] and was supplied via
189 www.socisurvey.de. It was presented in English. In line with the two aims of this project, the study
190 items focused on assessing the usability of the PRP-QUANT Template and measuring the UTAUT
191 variables (see Appendix C in the electronic supplementary material for a complete list of items).

192 As a first measure of usability, the template's effectiveness was inspected. For this purpose,
193 participants were asked to answer the items of the PRP-QUANT Template as if they were preparing
194 a real preregistration, thinking about a study they were currently planning or conducting (or, if no

195 current project was available, a previous study). They first provided a brief description of their
196 study and answered items about its status and whether they planned to preregister it. They were
197 then asked to complete the individual template items. Effectiveness was measured by coding
198 participants' responses to the template items in terms of their fit with what was asked in the item
199 (see section *data analysis and pre-processing*) and by having participants rate the perceived
200 importance of all items.

201 To gain a deeper insight into the participants' interaction with the template items, as well as
202 collect participants' suggestions for improvements for all items, several web probing questions
203 were displayed alongside the template items (derived from [22]), probing for category-selection
204 (e.g., for items such as T11 "Code availability" which required selecting an option, participants
205 were asked to elaborate why they selected the respective category), comprehension (e.g., asking
206 participants for the meaning of terms or paraphrasing, requesting participants to rate how well they
207 understood the item, or to differentiate template items from related items), or elaboration (e.g.,
208 asking participants for examples). Some of the web probing items were displayed for all template
209 items (i.e., rating the perceived importance of the item and an open-ended question asking what
210 participants would add, change, or remove about the item), while others were specific to individual
211 items.

212 Meanwhile, participants' attitudes regarding the template, as well as learnability and
213 flexibility, were assessed using various rating items that were displayed after participants had
214 finished working on the template. These items inquired about, for example, participants'
215 satisfaction with using the template (*attitude*), how well they understood it (*learnability*), or how
216 well it covered the most important aspects of their research (*flexibility*).

217 The items used for the usability test were open text input items, single- or multiple-choice
218 items, and rating items with varying scales (see Appendix C in the electronic supplementary
219 material).

220 To measure the UTAUT variables, the following scales were assessed: a *performance*
221 *expectancy* scale (i.e., five items measuring participants' expected performance when using the
222 template), an *effort expectancy* scale (i.e., five items inquiring about the expected effort when using
223 the template, where higher scores were associated with lower expected effort), a *social influences*
224 scale (i.e., a scale of five items examining the perceived social pressure to use preregistration), and
225 a *voluntariness* scale (i.e., perceived control over the behaviour). The *intention* to use the template
226 (dependent variable) was measured with one variable, as were *age* and *experience* (operationalised
227 as the participants' academic group). In addition, *facilitating conditions* (i.e., a scale of five items)
228 were measured. The latter scale was not part of our hypotheses tests, as it is assumed to influence
229 people's actual behaviour, not their intention [18]. However, this scale was still assessed since it
230 might provide insights into factors that might help foster the practice of preregistration. The
231 UTAUT scales were measured on a seven-point rating scale with 1 = "Disagree" to 7 = "Agree"
232 [18].

233 All UTAUT items were adapted from [18]. The other items were developed based on face
234 validity and revised in consultation with the members of the Preregistration Task Force that
235 developed the PRP-QUANT Template [13]. Additionally, before data collection, a pre-test was
236 conducted with four participants (two PhD students, one postdoc, and one professor), and its results
237 were used to further improve the items (e.g., by increasing their comprehensibility).

238 ***Procedure***

239 Participants received the study link via their respective society's mailing list or social media
240 (see section *participants*). After the welcome page, participant information was presented and
241 participants were required to provide informed consent to proceed. They were informed of the
242 study objectives (i.e., evaluation of the PRP-QUANT Template).

243 At the beginning of the study, the participants provided information about their
244 sociodemographic data and general use of preregistration. The study then focused successively on
245 the usability of the template and measurement of the UTAUT items. Some items of the study were
246 only shown to participants who had preregistered before (see Appendix C in the electronic
247 supplementary material). Before any PRP-QUANT Template or web probing items were displayed,
248 all participants were shown the entire template. They opened the template in a table format in a
249 new browser tab by clicking on a link provided in the study and were asked to look at the entire
250 template to obtain a general impression. A control question regarding the content of a template
251 item had to be answered correctly to proceed. Participants were asked to keep the template open in
252 the additional tab so that they could refer to it throughout the study.

253 For answering the template and web probing items, participants were randomly assigned to
254 one of four conditions. Depending on their condition, they were requested to fill out only a
255 subsection of the template: 1) title and introduction, 2) overall methods, sampling procedure, and
256 data collection, 3) overall methods, conditions, and design, or 4) analysis plans. This aimed to
257 reduce the burden placed on each participant. Twenty-three participants were in condition 1, 29 in
258 condition 2, 19 in condition 3, and 17 in condition 4 (overall: $N = 88$).

259 No items of the study were mandatory besides the filter question at the beginning, which
260 inquired whether the participants worked in psychological research (see section *participants*).
261 However, for participants' study descriptions and the UTAUT questionnaire, participants who did
262 not respond were asked to confirm their choice to ensure that gaps were not created inadvertently.
263 Additionally, if participants did not respond to the template items, they were asked to provide a
264 reason for this (i.e., whether they thought the item was optional, made a mistake, did not know
265 what to answer, did not like the item, if the item did not fit their research, or they could provide
266 other reasons via open text input). This question itself was not mandatory.

267 On average, it took participants approximately 31 minutes to complete the study ($SD = 12$
268 min, $range = 54$ min, times adjusted for interruptions). The procedure was approved by the ethics
269 committee of Trier University, Germany (approval number: 27/2020). An example screen
270 recording of the procedure is available in Appendix D in the electronic supplementary material.
271 Additionally, a PDF of the questionnaire can be found in Appendix E.

272 ***Data analysis and pre-processing***

273 We used R (Version 4.2.2) [23] and the R-packages *corrplot* (version 0.92) [24], *lm.beta*
274 (version 1.7-2) [25], *olsrr* (version 0.5.3) [26], *psych* (version 2.2.9) [27], *RColorBrewer* (version
275 1.1-3) [28], *readxl* (version 1.4.2) [29], *Rmisc* (version 1.5.1) [30], *tidyverse* (version 2.0.0) [31],
276 and *writexl* (version 1.4.2) [32] for all analyses. All analysis scripts and anonymised data (including
277 meta-data about variables and values) are publicly accessible in the digital research repository
278 PsychArchives [33,34].

279 Data were pre-processed by recoding responses from multiple-choice questions (originally:
280 1 = "not selected" and 2 = "selected"; new: 0 = "not selected" and 1 = "selected") and turning

281 single-choice items into factors. The polarity of negatively poled scale items was reversed. All
282 UTAUT items as well as some other items of the web probing and overall evaluation were recoded
283 from “1 to 7” to “-3 to 3”, yielding a middle category which has absolute meaning (i.e., 0 = neutral
284 opinion, neither agreement nor disagreement). As specified in the preregistration, empty data and
285 nonsense responses (e.g., random key pressing) were excluded.

286 **Coding of open text input.** Participants’ responses to the template items and other open
287 text input items were coded for the analysis. Three coders were involved in this process, but coding
288 was split between coders item-wise so that only one individual coded all responses for one item.

289 For the participants’ responses to the template items, it was coded if the given response
290 matched what was requested in the item (0 = “not applicable”, 1 = “fits poorly”, 2 = “fits
291 moderately”, 3 = “fits well”, -9 = “nonsense answer”). For this, a coding scheme was used, which
292 was developed and published prior to data collection alongside the preregistration [20]. To improve
293 the pre-specified coding scheme and represent as many potential responses as possible, the template
294 responses of 25% of participants per condition were randomly selected and coded, while the coding
295 scheme was revised in the process (e.g., the coding categories 0 = “not applicable” and -9 =
296 “nonsense answer” were added). Subsequently, the improved coding scheme was applied to the
297 remaining datasets. The final coding scheme is available in Appendix F in the electronic
298 supplementary material.

299 Next, open web probing questions and other open text input items were evaluated by coding
300 common themes. Responses were shuffled and the coder read the first 10% of the shuffled
301 responses. They identified common themes mentioned by the participants, which were then
302 transferred to new columns in a coding sheet. Then, it was coded for all other responses if the theme
303 was mentioned (= 1) or not mentioned (= 0). If new relevant topics appeared to the coder that they

304 had not coded before, these were added as categories as the coding continued and were coded later.
305 For the item “What would you add, change, or remove about the item?”, common themes were
306 categorised into 1) things to add, 2) things to change, and 3) things to remove.

307 When asked for definitions, explanations, or examples, a different coding was implemented.
308 For definitions, it was coded if the term was correctly described in the response (= 1) or not (= 0),
309 and for examples it was coded if the examples fit the requested term (= 1) or not (= 0). All coded,
310 anonymised comments are published alongside the data [33].

311 **Quality check of UTAUT scale items.** For the items of the UTAUT scales and the overall
312 evaluation, floor and ceiling effects were inspected, that is, items for which $\geq 90\%$ of participants
313 selected the lowest or highest category. No floor or ceiling effects were found for the overall
314 sample, nor the sample used for the UTAUT analyses. Furthermore, considering only the data of
315 participants included in the hypotheses tests, the reliability of the UTAUT scales was inspected.
316 The reliability analyses showed high to excellent reliability for the performance expectancy ($\alpha =$
317 $.87$) and effort expectancy scales ($\alpha = .9$), adequate reliability for the social influence scale ($\alpha =$
318 $.76$), and moderate reliability for the voluntariness scale ($\alpha = .62$).

319 *Deviations from the preregistration*

320 All deviations from the preregistered plan are displayed in Table 1 below. For each
321 deviation, a justification is provided.

322

323 **Table 1**324 *Deviations from the preregistration*

Section	Description and justification
Recruitment	Reminder emails were sent later than anticipated (not after one week, but after five weeks for the DGPs, and two and a half weeks for the BPS). For the APA, no reminder was sent, instead the study was also advertised in their newsletter.
Participants	Originally, it was planned to include all participants that started the main body of the study in the descriptive reports. However, since many participants dropped out before starting to work on the template items, and these are the core part of the study, we decided to report all descriptive reports for this sub-sample ($N = 88$).
Pre-processing	In addition to the preregistered pre-processing steps, further quality checks were conducted, but did not result in any modifications in item inclusion. Specifically, reliability as well as floor and ceiling effects were inspected (i.e., it was checked for items of the UTAUT scales and overall evaluation, if $\geq 90\%$ of participants answered the lowest/highest category). Reliability analyses showed moderate to excellent reliability, no items needed to be excluded. No floor or ceiling effects were found.
UTAUT	<p>Since the assumption tests showed a high multicollinearity due to the interaction terms, for the hypotheses test the UTAUT scales were centred instead of recoding them from “1 to 7” to “-3 to 3”.</p> <p>In the UTAUT sample, for the “academic group” variable, the option “other” was excluded ($n = 1$) because it holds no information for the regression model (heterogeneous group).</p> <p>For the scales, instead of displaying means and standard deviations, these were displayed in a plot showing the mean and confidence interval, for easier inspection.</p> <p>It was not clearly defined a priori that one-sided tests would be used for the regression weights, however, since directional hypotheses were tested, this was implemented. This had no impact on the results.</p>
Coding of open comments	For the web probing, it was originally planned to code common themes for questions of the type “how is this item different from another item”. However, it makes more sense to code whether the reported differences were perceived correctly (= 1) or incorrectly (= 0).

325

326 **Results**327 *Usability of the PRP-QUANT Template*328 **Satisfaction, perceived effectiveness, fit to research area, and comprehensiveness.**

329 Participants’ responses concerning the overall evaluation of the template are displayed in Figure 2.

330 On average, they were rather satisfied with using the template ($Mean = 0.72$, $Median = 1$, $SD =$

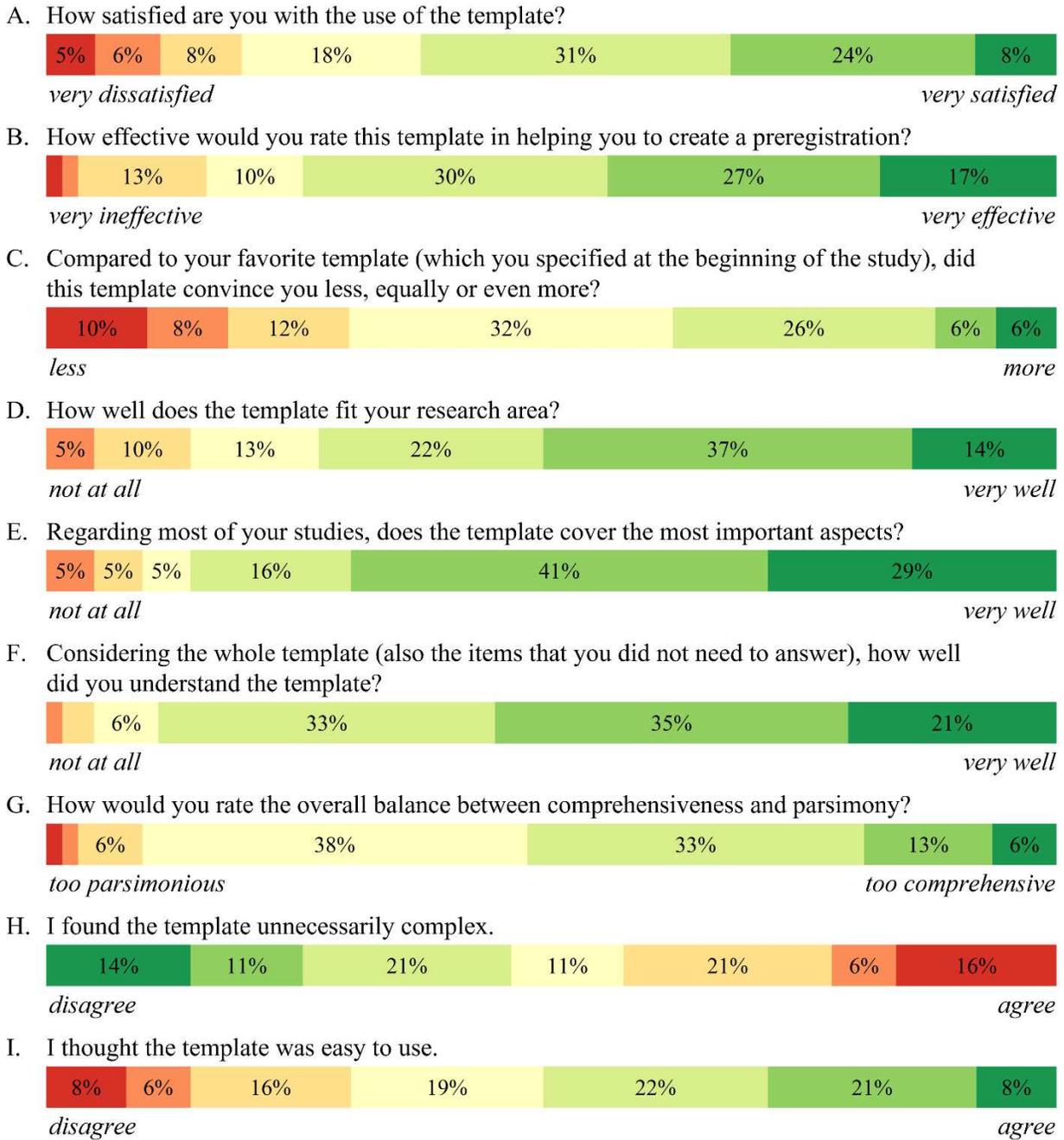
331 1.54, $IQR = 2$, $range = 6$, on a scale from -3 = “very dissatisfied” to 3 = “very satisfied”, see Figure
332 2A). They rated it as being effective for helping them create a preregistration ($Mean = 1.18$, $Median$
333 $= 1$, $SD = 1.41$, $IQR = 1.25$, $range = 6$, on a scale from -3 = “very ineffective” to 3 = “very
334 effective”, see Figure 2B). Compared to their favourite preregistration template, the PRP-QUANT
335 Template convinced the participants to about the same extent ($Mean = 0.04$, $Median = 0$, $SD =$
336 1.51 , $IQR = 2$, $range = 6$, on a scale from -3 = “less” to 3 = “more”, see Figure 2C). When asked
337 how likely they would use the template in the future to create their preregistrations, participants
338 indicated an average probability of 61.47% ($Median = 68$, $SD = 28.51$, $IQR = 33.75$, $range = 100$).
339 Additionally, they indicated an average probability of 64.67% ($Median = 72$, $SD = 30.25$, $IQR =$
340 40 , $range = 100$) for recommending the template to a colleague.

341 The PRP-QUANT Template fit quite well to the participants’ research areas ($Mean = 1.22$,
342 $Median = 2$, $SD = 1.38$, $IQR = 1.25$, $range = 5$, on a scale from -3 = “not at all” to 3 = “very well”,
343 see Figure 2D) and covered the most important aspects of their studies ($Mean = 1.65$, $Median = 2$,
344 $SD = 1.34$, $IQR = 2$, $range = 5$, on a scale from -3 = “not at all” to 3 = “very well”, see Figure 2E).
345 Additionally, the participants understood the template well ($Mean = 1.58$, $Median = 2$, $SD = 1.11$,
346 $IQR = 1$, $range = 5$, on a scale from -3 = “not at all” to 3 = “very well”, see Figure 2F).

347 When asked to rate the template’s overall balance between comprehensiveness and
348 parsimony, they rated it as rather comprehensive ($Mean = 0.68$, $Median = 1$, $SD = 1.08$, $IQR = 1$,
349 $range = 6$, on a scale from -3 = “too parsimonious” to 3 = “too comprehensive”, see Figure 2G).
350 However, they did not find it unnecessarily complex ($Mean = -0.08$, $Median = 0$, $SD = 1.99$, IQR
351 $= 3$, $range = 6$, on a scale from -3 = “disagree” to 3 = “agree”, see Figure 2H). Instead, they tended
352 towards finding it easy to use ($Mean = 0.35$, $Median = 0.5$, $SD = 1.71$, $IQR = 3$, $range = 6$, on a
353 scale from -3 = “disagree” to 3 = “agree”, see Figure 2I).

354 **Figure 2**

355 *Rating of overall satisfaction, perceived effectiveness, fit to research area, and comprehensiveness*
 356 *in percent*



357
 358 *Note.* All items were rated on a seven-point rating scale. Percentages are based on all responses to
 359 each item (A: $N = 62$; B: $N = 63$; C: $N = 50$; D – I: $N = 63$). Only percentages above 5% are labelled.

360 **Participants' suggestions for improving the template.** Participants were invited to
361 provide open text input suggestions to improve the template. Only the themes mentioned more than
362 once are included here, but all coded comments can be inspected online [33]. Of the 37 participants
363 who responded to this item, 13.51% complimented the comprehensiveness of the template and
364 found it to be a good guide for preregistration beginners and early career researchers. However,
365 32.43% pointed out that the template was very long and specific and that it might be beneficial to
366 reduce its complexity. Correspondingly, 10.81% suggested providing a shorter basic template,
367 where you specify the study type at the beginning and then get more specific items matching your
368 study type. Additionally, 10.81% of the participants indicated that some items seemed redundant
369 and that it would be helpful if the instructions provided additional information to clarify their
370 differences. The participants also provided suggestions for the practical implementation of the
371 template: for example, 8.11% suggested offering it in different formats (which is currently already
372 the case, see <https://doi.org/10.23668/psycharchives.4584>) and to provide predefined options
373 which can be adjusted if you deviate from them. Another 5.41% wished for example answers that
374 could be used as support when filling in the questionnaire. Furthermore, 16.22% of the participants
375 indicated that, while the template fits best with confirmatory and experimental studies, items for
376 other research types might be added. Paying attention to interdisciplinarity was also suggested by
377 participants in the general comment section at the end of the study (8.7% of 23 responses), while
378 most responses were praise to the template (34.78%) or more general comments about
379 preregistration or the study.

380 **Individual template items and web probing.** Next, the participants' responses and
381 comments regarding the individual template items were inspected. Overall, 88 participants worked
382 on the template items (see section *participants*), whose responses were consequently used for the
383 analysis of the individual template items and web probing questions. Of these, 21.59% had just

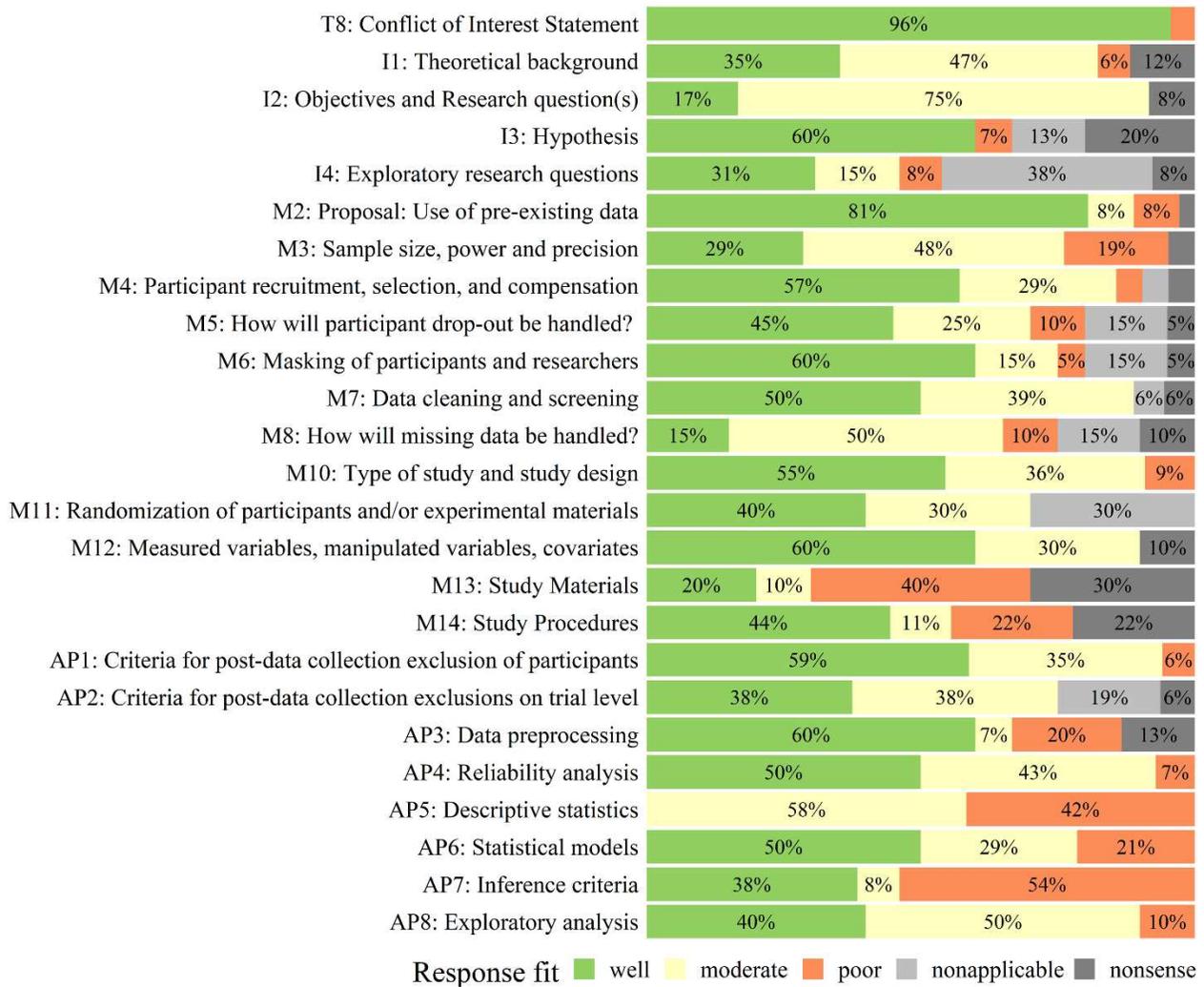
384 started planning the study, 35.23% had planned the study in detail, 13.64% were currently
385 conducting their study, 28.41% had already completed their study, and 1.14% did not indicate their
386 study status. Moreover, 29.55% planned to preregister their study, 19.32% were currently working
387 on the preregistration, 21.59% had already preregistered, 28.41% did not (plan to) preregister the
388 study, and 1.14% did not indicate the preregistration status.

389 ***Overall good fit of responses.*** Of all responses provided for the template items, 48.61% fit
390 well with the item, 27.96% fit moderately, 11.34% fit poorly, 6.05% indicated that the item was
391 not applicable to the participants' studies, and 6.05% were nonsense answers. Figure 3 provides an
392 overview of the response fit for all template items presented that required an open-ended response
393 from participants. Inspecting the plot reveals that for some items, answers were primarily well-
394 fitting (e.g., item T8 "Conflict of Interest Statement", M2 "Use of pre-existing data"), and that most
395 items showed a moderate to good response fit. However, a higher proportion of poor answers were
396 given for the items M13 "Study Materials", M14 "Study Procedures", AP3 "Data preprocessing",
397 AP5 "Descriptive statistics", AP6 "Statistical models" and AP7 "Inference criteria" (i.e., $\geq 20\%$ of
398 responses were coded as "poor"). The means, standard deviations, medians, and ranges for each
399 item's fit are displayed in Appendix G in the electronic supplementary material.

400

401 **Figure 3**

402 *Fit of the participants' responses to the PRP-QUANT Template items*

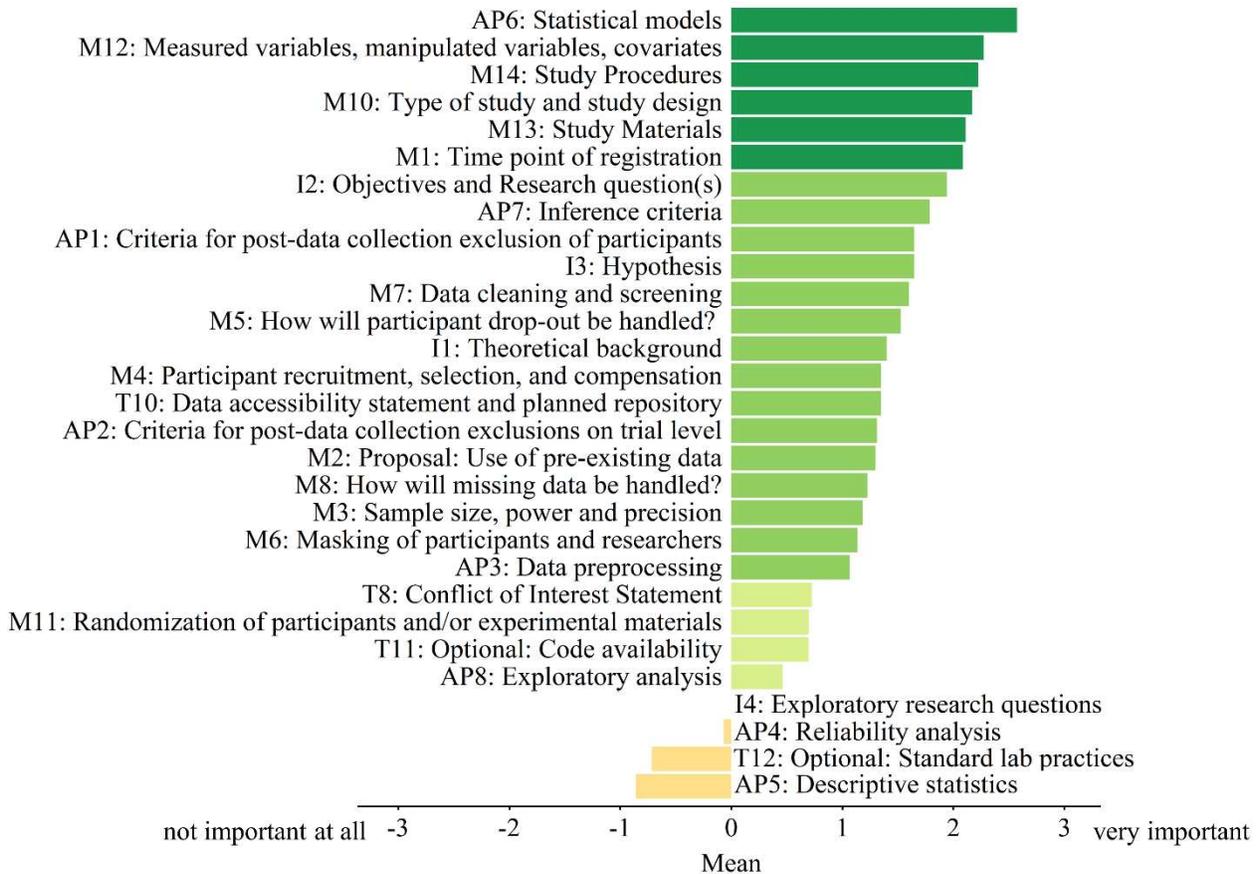


404 Whenever participants did not answer the template items, they were prompted to provide
 405 reasons for doing so (see section *procedure*). Of the 44 responses to these prompts, 50% indicated
 406 that the participants did not know what to answer, 11.36% of the participants said that it was a
 407 mistake, 6.82% thought the item was optional, 4.55% did not like the item, 2.27% said that it did
 408 not apply to their research, and 25% gave other reasons, most of which aligned with the given
 409 options.

410 ***Template items perceived as important for preregistration.*** Participants felt that most of
411 the items in the PRP-QUANT Template were important for preregistering their studies (see Figure
412 4). The items rated most important were AP6 “Statistical models”, M12 “Measured variables,
413 manipulated variables, covariates”, M14 “Study Procedures”, M10 “Type of study and study
414 design”, M13 “Study Materials”, and M1 “Time point of registration” (i.e., their mean was above
415 2 on a scale from -3 = “not important at all” to 3 = “very important”). Most other items were also
416 rated as important (i.e., their mean was above 0, for most items above 1). The item I4 “Exploratory
417 research questions” was rated as neither important nor unimportant. Still, most participants felt that
418 including exploratory research questions and analyses in the preregistration was appropriate (i.e.,
419 for research questions, 56.25% indicated “definitely yes” or “maybe yes”, and 57.14% did so for
420 analyses). The items AP4 “Reliability analysis”, T12 “Optional: Standard lab practices”, and AP5
421 “Descriptive statistics” were rated as relatively unimportant (i.e., their mean was below 0). This
422 makes sense in that reliability analyses are not applicable to all studies, providing standard lab
423 practices is optional (and only three of the 20 participants who answered even had a standard lab
424 practices document), and descriptive statistics have no direct impact on hypotheses testing.

425 **Figure 4**

426 *Importance rating of PRP-QUANT Template items*



427

428 ***Participants’ suggestions for individual items and web probing.*** For each template item

429 that the participants worked on, they were asked what they would add, change, or remove. The

430 participants offered a variety of comments and suggestions, which are summarised in Appendix H

431 in the electronic supplementary material. Additionally, participants responded to several other web

432 probing items (see Appendix C in the electronic supplementary material for a complete list) which

433 queried, for example, why they had selected an answer, whether they correctly understood the

434 concepts underlying the items and which were unclear, how they perceived the link between items,

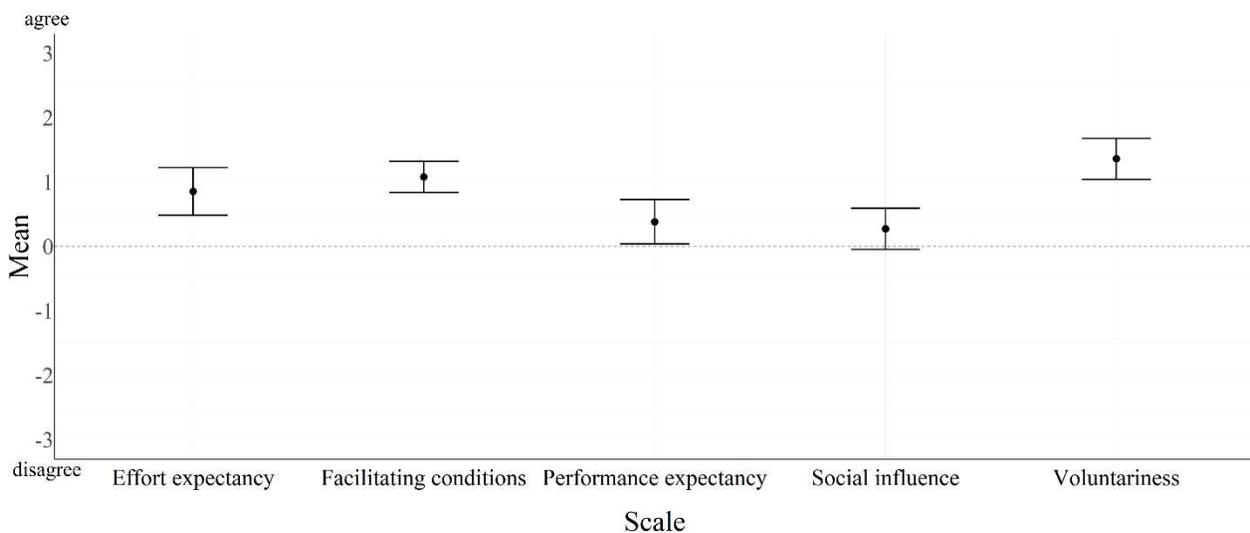
435 and whether they could distinguish items from others. These are presented in detail in Appendix I
 436 in the electronic supplementary material.

437 *Significant prediction of intention by performance expectancy and all predictors combined*

438 To investigate participants' intention to use the PRP-QUANT Template in the future, the
 439 UTAUT items were analysed. For each participant, the mean scores were computed for all UTAUT
 440 scales (i.e., performance expectancy, effort expectancy, social influence, facilitating conditions,
 441 and voluntariness). The means and confidence intervals for all scales are displayed in Figure 5.

442 **Figure 5**

443 *UTAUT Scales*



444

445 *Note.* Scales ranged from 1 = “Disagree” to 7 = “Agree” and were recoded to “-3 to 3” (however,
 446 these were centred for the hypotheses tests, see below). The parameters were calculated based on
 447 the sample used for the UTAUT hypotheses tests (i.e., participants who responded to all items used
 448 in these analyses, $N = 60$). Error bars indicate 95% confidence intervals. Higher effort expectancy
 449 scores are associated with lower expected effort. Facilitating conditions were considered
 450 descriptively but were not included in the hypotheses tests.

451 We expected that the intention to use the template in the future is predicted by performance
452 expectancy (moderated by age), effort expectancy (moderated by age and experience, i.e.,
453 academic group), and social influence (moderated by age, experience, i.e., academic group, and
454 voluntariness of use, see Figure 1). To test these hypotheses, a moderated multiple regression
455 model was computed, which is a method that has been frequently used to test the UTAUT (see
456 [35]).

457 Behavioural intention (i.e., the answer to the item “How likely would you use the template
458 in the future to create your preregistrations?”) was included as the dependent variable. The
459 following predictors were included in the model: 1) the score on the performance expectancy scale,
460 2) performance expectancy \times age, 3) the score on the effort expectancy scale, 4) effort expectancy
461 \times age, 5) effort expectancy \times experience (i.e., academic group), 6) the score on the social influence
462 scale, 7) social influence \times age, 8) social influence \times experience, 9) social influence \times
463 voluntariness, 10) age, 11) experience, and 12) voluntariness. The significance of the overall
464 model, as well as of individual predictors and moderating effects, was evaluated at $\alpha = .05$. Because
465 of our directional hypotheses, the regression weights were tested in a one-tailed fashion.

466 Before computing the moderated regression model, its assumptions were tested: linearity,
467 uncorrelated predictors, independence and normality of residuals, and homogeneity of variance.
468 The assumption tests showed high multicollinearity among the predictors, as judged based on the
469 variance inflation factors (VIF > 10 for seven predictor/interaction terms). Because of this, all
470 predictors except for experience (nominal variable) were centred, which drastically reduced
471 multicollinearity (VIF < 10 for all except one predictor).

472 As expected, the predictors of the UTAUT combined were able to significantly predict
473 researchers' intention to use the PRP-QUANT Template in the future, $F(15, 44) = 3.94, p < .001$,

474 $R^2_{\text{adjusted}} = 42.79\%$. However, of the individual predictors, only performance expectancy was a
475 significant predictor for intention, $t(44) = 2.28$, $p_{\text{one-sided}} = .014$, $\beta = .36$. When this model is
476 compared to the preregistered model with un-centred predictors, performance expectancy becomes
477 non-significant, $t(44) = 1.58$, $p_{\text{one-sided}} = .06$, $\beta = .87$.

478 **Study 2: Survey of preregistration authors and reviewers**

479 In study 2, we surveyed psychological researchers who had used the PRP-QUANT
480 Template to create a preregistration and submitted it when applying for free-of-charge data
481 collection in ZPID's call for online studies. In addition, call reviewers were asked how they felt
482 about reviewing preregistrations based on the PRP-QUANT Template. Instead of being constructed
483 based on theoretical considerations, these surveys were designed to provide a quick exploration of
484 the participants' experiences with the template and processes within the call. Thus, the items of
485 study 2 did not cover all aspects of usability measured in study 1, but instead focused largely on
486 participants' attitudes regarding the template.

487 **Methods**

488 This study was not preregistered as it was conducted on a short notice. It does not include
489 hypotheses tests. Instead, we report the survey results descriptively. As for study 1, the data and
490 analysis scripts are available online [33,34].

491 ***Participants***

492 Twenty-eight preregistrations reached the review stage of the call. After this stage, all
493 submitting authors (i.e., mostly the first author) were invited to participate in the author survey.

494 Nineteen authors participated in this study. Data were collected between May 16, 2022, and May
495 29, 2022. After the initial invitation, two reminders were sent to all potential participants.

496 Meanwhile, 44 researchers were reviewers of an accepted proposal and were invited to
497 participate in the reviewer survey, of whom 29 participated. Their data were collected between
498 March 11, 2022, and March 26, 2022, with one reminder sent to them. Participants were not
499 compensated.

500 *Material and measures*

501 Both surveys were created using Google Forms (<https://docs.google.com/forms>). The
502 author survey consisted of 21 items and the reviewer survey consisted of 17 items. There were five
503 different sections of items in the author survey: 1) items about the participants' previous and future
504 use of preregistration and PsychLab, 2) items regarding their experiences with using the structured
505 template compared to a continuous format (i.e., a normal report) and 3) their experiences during
506 the review process, 4) questions about their general opinion of preregistration, and 5) additional
507 comments. The shorter reviewer survey had three different sections of items: 1) items asking
508 participants to compare their experiences with reviewing this structured format versus a continuous
509 text, 2) questions concerning their general impression of reviewing preregistrations instead of full
510 manuscripts, and 3) additional comments. Some of the items related more generally to the call's
511 processes and participants' general impression of preregistration, however, we only present those
512 responses related to the usability of the template in this article.

513 Of the various aspects of usability measured in study 1 [17], attitudes were the focus of
514 study 2. Authors were asked whether they felt that the structure of the template facilitated creating
515 their preregistration, whether it made them think of details that were important for planning their

516 study (which can also be seen as an indicator of flexibility), whether the template helped them
517 include all relevant information, or whether they felt that the items should be reduced. Similarly,
518 reviewers were asked whether they felt that the structure of the template facilitated their evaluation,
519 helped them find the information more easily, focus their attention on the relevant sections, and
520 assess the completeness of the information, or whether, in contrast, the template hindered their
521 reading flow or contained too many elements irrelevant to the review. Participants were also asked
522 whether they would have preferred to prepare/review the proposal in this structured preregistration
523 format or in a normal report format. As a further aspect, learnability was measured by asking the
524 authors if they had difficulty understanding what they were supposed to fill in for some items.

525 Most items were 7-point rating items with 1 = “strongly disagree” and 7 = “strongly agree”.
526 In addition, participants were given the opportunity to express their opinions in several open text
527 input items (e.g., “Anything else you would like to add about the template? Please comment here”).
528 See Appendix C in the electronic supplementary material for an overview of the items in the author
529 and reviewer surveys.

530 *Procedure*

531 The surveys took approximately five minutes to complete. Participants were invited via
532 personal email, and data were collected anonymously. In both surveys, participants successively
533 completed the different item sections (for authors: use of preregistration and PsychLab, comparison
534 of structured template versus continuous text, review process, general opinion of preregistration,
535 and additional comments; for reviewers: comparison of structured template versus continuous text,
536 comparison of reviewing preregistrations versus complete manuscripts, and additional comments),
537 which were each displayed on a new page. None of the items were mandatory.

538 *Data analysis and pre-processing*

539 Again, R (Version 4.2.2) [23] was used for the analysis. Data of the rating items were pre-
540 processed similarly to study 1, that is, they were recoded from “1 to 7” to “-3 to 3” to facilitate
541 interpretation. Scales that deviated from this format were not recoded (these are labelled
542 accordingly below). For all rating items, the mean, median, standard deviation, interquartile range,
543 and range were calculated. Percentages were computed to examine preregistration experience and
544 intention. Open-ended comments were manually reviewed and summarised.

545 **Results**

546 *Author responses*

547 Of all authors who participated in the survey, 31.58% had preregistered for the first time
548 (i.e., 68.42% had previous preregistration experience). Overall, 94.74% of participants intended to
549 preregister in the future (i.e., their scores were above 3.5 on a scale from 0 = “very unlikely” to 7
550 = “very likely”), of which 73.68% selected “very likely”.

551 Overall, the authors rated the PRP-QUANT Template favourably. They mostly agreed that
552 it facilitated creating their preregistration (*Mean* = 1.63, *Median* = 2, *SD* = 1.01, *IQR* = 1, *range* =
553 3), that it made them think of details that were important for planning their study (*Mean* = 1.79,
554 *Median* = 2, *SD* = 0.92, *IQR* = 1, *range* = 3), and that it helped them include all relevant information
555 (*Mean* = 2.06, *Median* = 2, *SD* = 0.87, *IQR* = 0.75, *range* = 3). They mostly disagreed with the
556 statements that the template items should be reduced (*Mean* = -0.47, *Median* = -1, *SD* = 1.84, *IQR*
557 = 3, *range* = 6), that they (i.e., the authors) had difficulty understanding what they were supposed
558 to fill in on some items (*Mean* = -1.16, *Median* = -2, *SD* = 1.83, *IQR* = 3.5, *range* = 5), and that

559 they would have preferred to write the proposal in a normal report format (continuous text) rather
560 than a structured format ($Mean = -1.37$, $Median = -2$, $SD = 1.64$, $IQR = 2$, $range = 5$).

561 In the open comments, most participants expressed satisfaction with using the template.
562 Some suggestions for improvements were made, each voiced by one participant, respectively. For
563 example, it was suggested to provide a front page with a link to all different subsections to make it
564 easier for authors and reviewers to navigate the document, shorten the template and reduce
565 redundancies, revise the structure of the Word template, and query the abstract in one item rather
566 than subdividing it.

567 ***Reviewer feedback***

568 Reviewers also perceived various advantages of the PRP-QUANT Template. They
569 described that the structure of the template facilitated their evaluation ($Mean = 1.07$, $Median = 1$,
570 $SD = 1.25$, $IQR = 2$, $range = 4$), that it helped them find the information ($Mean = 1.03$, $Median =$
571 1 , $SD = 1.4$, $IQR = 2$, $range = 4$), that the structure of the template helped them focus their attention
572 on the relevant sections ($Mean = 1.21$, $Median = 1$, $SD = 1.29$, $IQR = 2$, $range = 5$), and that it
573 helped them evaluate the completeness of the information ($Mean = 1.14$, $Median = 2$, $SD = 1.46$,
574 $IQR = 2$, $range = 5$). Correspondingly, they would not have preferred to review the proposal in a
575 normal report format of continuous text rather than in a structured format ($Mean = -0.83$, $Median$
576 $= -1$, $SD = 1.56$, $IQR = 2$, $range = 5$), they did not feel that the layout of the template hindered their
577 reading flow ($Mean = -0.9$, $Median = -1$, $SD = 1.8$, $IQR = 3$, $range = 5$), or that the template
578 contained too many elements irrelevant to reviewing the proposal ($Mean = -1.1$, $Median = -2$, SD
579 $= 1.59$, $IQR = 2$, $range = 5$).

580 In the open text field, reviewers commented that they found the template generally helpful
581 and provided some suggestions for improvements. For example, it was suggested to add an item
582 about scientific and thematic relevance, or the possibility of including scripts and results from data
583 analysis (e.g., power analyses). This is already possible by using the PRP-QUANT Template in R
584 Markdown or JupyterLab. Furthermore, it was commented that the template could be more concise,
585 and that the items A2 “Objectives and Research questions” and I2 “Objectives and Research
586 question(s)” seemed redundant (however, since A2 is part of the abstract, these items ask for
587 different depths of information). One person recommended dividing the template into two sections
588 to facilitate reviewing: authors could elaborate in the first section everything relevant to reviewing
589 and then give all the relevant technical information in the second section.

590

Discussion

591 We conducted two studies to evaluate the usability of the PRP-QUANT Template and
592 identify areas for improvement. Furthermore, we wanted to find out whether psychological
593 researchers plan to use the template in the future to create their preregistrations and examine which
594 variables might be important for this intention formation.

595 Usability of the template rated high, with suggestions for improvements

596 We assessed the usability of the PRP-QUANT Template in study 1 by conducting a
597 simulation trial in which we asked psychological researchers to think of one of their studies to
598 complete selected parts of the template, and in study 2 by surveying authors and reviewers of
599 preregistrations that were part of a call for online studies. For this evaluation, we referred to the
600 four aspects of usability defined by Shackel [17]: learnability, flexibility, effectiveness, and
601 attitude.

602 *Learnability and flexibility*

603 In both study 1 and study 2, participants indicated a good understanding of the template.
604 They did not find it unnecessarily complex and instead tended towards finding it easy to use.
605 Researchers from many different research areas participated – many came from experimental and
606 cognitive psychology, but also from social psychology, clinical psychology, educational
607 psychology, organisational psychology, developmental psychology, research methods, general
608 psychology, neuroscience, neuropsychology, and differential psychology. Overall, the participants
609 indicated that the PRP-QUANT Template fit their research areas quite well and covered the most
610 important aspects of their studies. This suggests that both the learnability and flexibility of the
611 template are adequate and that the template seems to capture the main points of various sub-
612 disciplines. Still, some participants struggled to understand complex terms and differentiate
613 between similar items, which could be improved, for instance, by providing examples.

614 *Effectiveness*

615 Various indicators point to the effectiveness of the PRP-QUANT Template. Participants felt
616 that most of the template items were important for preregistering their studies, with the most highly
617 rated items being AP6 “Statistical models”, M12 “Measured variables, manipulated variables,
618 covariates”, M14 “Study Procedures”, M10 “Type of study and study design”, M13 “Study
619 Materials”, and M1 “Time point of registration”. More than three-quarters of participants’
620 responses to the template items matched the requested information moderately or well, with nearly
621 half of the responses fitting well. When looking at the individual items, it appears that most had a
622 moderate to good response fit. For some items, participants indicated that they were not applicable
623 to them, which is consistent with how the PRP-QUANT Template was designed (i.e., with many
624 different items, of which only the applicable ones should be answered).

625 There was a higher proportion of poor answers for the items M13 “Study Materials”, M14
626 “Study Procedures” AP3 “Data preprocessing”, AP5 “Descriptive statistics”, AP6 “Statistical
627 models” and AP7 “Inference criteria”. These items could be prime candidates for revision.
628 However, it must be noted that these are also items that require very elaborate responses. Since
629 study 1 imposed quite high demands on the participants while no compensation was given, it may
630 be assumed that this could be the lower end of the scale of possible response quality. Nevertheless,
631 the participants had various suggestions on how these and other items could be improved, which
632 could be included in a new version of the template.

633 *Attitudes*

634 Participants’ attitudes towards the PRP-QUANT Template were also rather positive. They
635 were satisfied with using the template and found it effective for helping them create a
636 preregistration, which was also true for the authors of preregistrations in study 2. Compared to their
637 favourite preregistration template, the PRP-QUANT Template convinced the participants to about
638 the same extent. They also indicated an average probability of over 60% that they would use it in
639 the future to create their preregistrations or recommend it to a colleague. However, while the
640 template was not considered unnecessarily complex, participants commented on its length on
641 several occasions. Accordingly, many of the participants’ suggestions were aimed at requesting the
642 information in a more condensed form, which could possibly be considered in a new version of the
643 template.

644 In study 2, authors of preregistrations indicated that the template made them think of
645 elements important for planning their study and that it helped them include all relevant information.
646 Correspondingly, reviewers felt that the structure of the template facilitated their evaluation, helped

647 them find all relevant information and focus their attention, and helped them evaluate the
648 completeness of the information.

649 In summary, the usability of the PRP-QUANT Template was found to be high in our studies.
650 Nevertheless, we identified some possibilities for improvements. The next step is to implement
651 these in a new version of the template.

652 **Prediction of intention primarily through expected performance gains**

653 Besides evaluating the usability of the PRP-QUANT Template, study 1 examined
654 researchers' intention to use it in the future, as well as possible influences on that intention. Based
655 on the UTAUT [18,19], we expected that the intention to use the template is influenced by
656 performance expectancy (moderated by age), effort expectancy (moderated by age and experience),
657 and social influence (moderated by age, experience, and voluntariness of use).

658 Our results show that participants' average intention to use the template in the future to
659 create their preregistrations was rather high (61.47%). Descriptively, all UTAUT variables
660 indicated a positive perception of the template, that is, all scale means were above zero, where zero
661 indicated a neutral opinion and positive values indicated a positive opinion. This suggests that
662 participants tended to believe that using the template would help them attain gains in performance
663 (*performance expectancy*) and that the template would be easy to use (*effort expectancy*).
664 Additionally, participants felt that, while preregistration is voluntary (*voluntariness*), others would
665 approve of them preregistering (*social influence*) and that organisational and technical
666 infrastructures exist that support the preregistration process (*facilitating conditions*).

667 As hypothesised, all predictors of the UTAUT combined were able to significantly predict
668 researchers' intention to use the PRP-QUANT Template in the future. However, of the individual

669 predictors, only performance expectancy significantly predicted intention. This suggests that the
670 expectation that the template will be useful for one's own research has the strongest influence on
671 whether researchers plan to use it. Highlighting this benefit could therefore help raise researchers'
672 awareness and adoption of the template in the future.

673 **Limitations**

674 The implementation of study 1 had some limitations. Responding to the template
675 represented a considerable amount of effort for the participants, for which they were not
676 compensated. It may be assumed that the quality of the responses reported here represents the lower
677 bound, as researchers would likely put much more effort into creating their actual preregistrations.
678 This suggests that the fit of the answers might be even better in a real deployment of the template.
679 In the future, this could be tested by examining the preregistrations of the authors we surveyed in
680 study 2, as they created their preregistrations to apply for a high external incentive and therefore
681 likely spent more time on their preregistrations. However, this does not undermine the participants'
682 suggestions for improvements that can be used to revise the template.

683 The format and method of responding to the items in study 1 were also constrained. We
684 presented participants with the PRP-QUANT Template in a table format and queried the template
685 items one after another in our online questionnaire, where participants could not skip back and
686 forth. Again, it can be assumed that satisfaction with the template would probably be even higher
687 if researchers could freely choose between all available formats (e.g., table, text, online form, R
688 Markdown, and Jupyter Notebook) and be able to switch flexibly between items. In addition, if
689 they were using the template to preregister a study outside the present usability test, they would
690 probably invest much more time and would not have to complete responding to the items in one

691 session. In line with this assumption, the template was evaluated very positively in study 2, both
692 by authors and reviewers.

693 Regarding our hypotheses tests, it must be noted that our a priori power analysis was
694 calculated with respect to the overall model, not the individual predictors. It could be that other
695 predictors besides performance expectancy are important for predicting the intention to use the
696 template but were not detectable with our sample size. In addition, there was multicollinearity in
697 our model, which we could improve by centring, but not eliminate completely.

698 Lastly, on a more general note, there has been some challenge to the validity of the UTAUT
699 in a recent meta-analysis [36]. Nevertheless, the robustness of the UTAUT and its main effects has
700 been repeatedly validated by research (e.g., see [19,37,38]). The authors of the meta-analysis argue
701 that there is not one specification of UTAUT that applies to all contexts, but that the ability of the
702 theory to predict the use of a new system is context dependent. Besides low power, this might have
703 also contributed to some of the predictors not being significant in our model.

704 **Future research**

705 Our studies examined the usability of the PRP-QUANT Template and identified its
706 strengths and areas for improvement. These can now be used to create an empirically founded
707 revision.

708 In the future, usability studies should be used to continually adapt the template to the needs
709 of the community. Other templates could also benefit from such usability assessments.
710 Additionally, studies could not only focus on the templates themselves, but also on preregistrations
711 created with the respective templates, following the approach of [14] and [15]. This would allow

712 an empirical investigation of the effectiveness of the templates in reducing researchers' degrees of
713 freedom.

714 **Conclusion**

715 In two studies, we identified both strengths and areas for improvement in the PRP-QUANT
716 Template. We obtained insights into learnability, flexibility, effectiveness, and attitudes, as well as
717 participants' comments and suggestions regarding the template. These can now serve as the basis
718 for an empirically informed revision. Moreover, we demonstrated that performance expectancy, as
719 well as all variables of the UTAUT combined, significantly predicted psychological researchers'
720 intention to use the template in the future. Overall, participants were likely to use the template or
721 recommend it to a colleague, which indicates that the template is being well received.

722 **Acknowledgments**

723 We would like to thank the members of the Joint Psychological Societies Preregistration Task
724 Force for their feedback on the items of study 1, our student assistants for coding, and T. Dauber
725 and J. Pauquet for proofreading.

726 **Funding**

727 We received no funding for these studies.

728

729

- 730 **References**
- 731 1. Munafò MR *et al.* 2017 A manifesto for reproducible science. *Nat. Hum. Behav.* 1, 0021.
732 (doi:10.1038/s41562-016-0021)
- 733 2. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. 2018 The preregistration revolution. *Proc.*
734 *Natl. Acad. Sci.* 115, 2600–2606. (doi:10.1073/pnas.1708274114)
- 735 3. Parsons S *et al.* 2022 A community-sourced glossary of open scholarship terms. *Nat. Hum.*
736 *Behav.* 6, 312–318. (doi:10.1038/s41562-021-01269-4)
- 737 4. Kaplan RM, Irvin VL. 2015 Likelihood of Null Effects of Large NHLBI Clinical Trials Has
738 Increased over Time. *PLOS ONE* 10, e0132382. (doi:10.1371/journal.pone.0132382)
- 739 5. Swaen GG, Teggeler O, van Amelsvoort LG. 2001 False positive outcomes and design
740 characteristics in occupational cancer epidemiology studies. *Int. J. Epidemiol.* 30, 948–954.
741 (doi:10.1093/ije/30.5.948)
- 742 6. Protzko J *et al.* 2020 High Replicability of Newly-Discovered Social-behavioral Findings is
743 Achievable [Preprint]. (doi:10.31234/osf.io/n2a9x)
- 744 7. Spitzer L, Mueller S. 2023 Registered report: Survey on attitudes and experiences regarding
745 preregistration in psychological research. *PLOS ONE* 18, e0281086.
746 (doi:10.1371/journal.pone.0281086)
- 747 8. van 't Veer AE, Giner-Sorolla R. 2016 Pre-registration in social psychology—A discussion and
748 suggested template. *J. Exp. Soc. Psychol.* 67, 2–12. (doi:10.1016/j.jesp.2016.03.004)

- 749 9. Beyer F *et al.* 2021 A fMRI pre-registration template. PsychArchives.
750 (doi:10.23668/PSYCHARCHIVES.5121)
- 751 10. Brandt MJ *et al.* 2014 The Replication Recipe: What makes for a convincing replication? *J.*
752 *Exp. Soc. Psychol.* 50, 217–224. (doi:10.1016/j.jesp.2013.10.005)
- 753 11. Crüwell S, Evans NJ. 2021 Preregistration in diverse contexts: a preregistration template for
754 the application of cognitive models. *R. Soc. Open Sci.* 8, 210155. (doi:10.1098/rsos.210155)
- 755 12. Van den Akker OR *et al.* 2021 Preregistration of secondary data analysis: A template and
756 tutorial. *Meta-Psychol.* 5. (doi:10.15626/MP.2020.2625)
- 757 13. Bosnjak M, Fiebach CJ, Mellor D, Mueller S, O'Connor DB, Oswald FL, Sokol RI. 2022 A
758 template for preregistration of quantitative research in psychology: Report of the joint
759 psychological societies preregistration task force. *Am. Psychol.* 77, 602–615.
760 (doi:10.1037/amp0000879)
- 761 14. Bakker M, Veldkamp CLS, van Assen MALM, Cromptvoets EAV, Ong HH, Nosek BA,
762 Soderberg CK, Mellor D, Wicherts JM. 2020 Ensuring the quality and specificity of
763 preregistrations. *PLOS Biol.* 18, e3000937. (doi:10.1371/journal.pbio.3000937)
- 764 15. Heirene R, LaPlante D, Louderback ER, Keen B, Bakker M, Serafimovska A, Gainsbury SM.
765 2021 Preregistration specificity & adherence: A review of preregistered gambling studies &
766 cross-disciplinary comparison [Preprint]. (doi:10.31234/osf.io/nj4es)
- 767 16. International Organization for Standardization (ISO). 2018 ISO 9241-11:2018: Ergonomics of
768 human-system interaction — Part 11: Usability: Definitions and concepts. Available from:
769 <https://www.iso.org/standard/63500.html>

- 770 17. Shackel B. 2009 Usability – Context, framework, definition, design and evaluation. *Interact.*
771 *Comput.* 21, 339–346. (doi:10.1016/j.intcom.2009.04.007)
- 772 18. Venkatesh, Morris, Davis, Davis. 2003 User Acceptance of Information Technology: Toward
773 a Unified View. *MIS Q.* 27, 425–478. (doi:10.2307/30036540)
- 774 19. Venkatesh V, Thong JYL, Xu X. 2016 Unified Theory of Acceptance and Use of Technology:
775 A Synthesis and the Road Ahead. *J. Assoc. Inf. Syst.* 17, 328–376. Available from:
776 <https://ssrn.com/abstract=2800121>
- 777 20. Spitzer L, Mueller S, Bosnjak M. 2021 Preregistration: Testing the Usability of the
778 Psychological Research Preregistration-Quantitative (PRP-QUANT) Template.
779 PsychArchives. (doi:10.23668/PSYCHARCHIVES.4636)
- 780 21. Leiner DJ. 2019 *SoSci Survey [Computer software]*. See <https://www.soscisurvey.de>.
- 781 22. Behr D, Meitinger K, Braun M, Kaczmirek L. 2017 Web probing – implementing probing
782 techniques from cognitive interviewing in web surveys with the goal to assess the validity of
783 survey questions. Mannheim, GESIS – Leibniz-Institute for the Social Sciences (GESIS –
784 Survey Guidelines). (doi:10.15465/gesis-sg_en_023)
- 785 23. R Core Team. 2021 R: A language and environment for statistical computing (version 4.2.2)
786 [Computer software]. R Foundation for Statistical Computing.
- 787 24. Wei T, Simko V. 2021 *R package ‘corrplot’: Visualization of a Correlation Matrix (Version*
788 *0.92) [Computer software]*. See <https://github.com/taiyun/corrplot>.
- 789 25. Behrendt S. 2014 *lm.beta: Add Standardized Regression Coefficients to lm-Objects (Version*
790 *1.7-2) [Computer software]*. See <https://CRAN.R-project.org/package=lm.beta>.

- 791 26. Hebbali A. 2020 *olsrr: Tools for Building OLS Regression Models (Version 0.5.3) [Computer*
792 *software]*. See <https://CRAN.R-project.org/package=olsrr>.
- 793 27. Revelle W. 2022 *psych: Procedures for Psychological, Psychometric, and Personality*
794 *Research (Version 2.2.9) [Computer Software]*. Evanston, Illinois: Northwestern University.
795 See <https://CRAN.R-project.org/package=psych>.
- 796 28. Neuwirth E. 2022 *RColorBrewer: ColorBrewer Palettes (Version 1.1-3) [Computer software]*.
797 See <https://CRAN.R-project.org/package=RColorBrewer>.
- 798 29. Wickham H, Bryan J. 2022 *readxl: Read Excel Files (Version 1.4.2) [Computer software]*. See
799 <https://CRAN.R-project.org/package=readxl>.
- 800 30. Hope RM. 2022 *Rmisc: Ryan Miscellaneous (Version 1.5.1) [Computer software]*. See
801 <https://CRAN.R-project.org/package=Rmisc>.
- 802 31. Wickham H *et al.* 2019 Welcome to the tidyverse (Version 2.0.0) [Computer software]. *J. Open*
803 *Source Softw.* 4, 1686. (doi:10.21105/joss.01686)
- 804 32. Ooms J. 2021 *writexl: Export Data Frames to Excel 'xlsx' Format (Version 1.4.2) [Computer*
805 *software]*. See <https://CRAN.R-project.org/package=writexl>.
- 806 33. Spitzer L, Bosnjak M, Mueller S. 2023 Dataset for: Testing the Usability of the Psychological
807 Research Preregistration-Quantitative (PRP-QUANT) Template. PsychArchives.
808 (doi:10.23668/PSYCHARCHIVES.12885)
- 809 34. Spitzer L, Bosnjak M, Mueller S. 2023 Code for: Testing the Usability of the Psychological
810 Research Preregistration-Quantitative (PRP-QUANT) Template. PsychArchives.
811 (doi:10.23668/PSYCHARCHIVES.12884)

- 812 35. Williams MD, Rana NP, Dwivedi YK. 2015 The unified theory of acceptance and use of
813 technology (UTAUT): a literature review. *J. Enterp. Inf. Manag.* 28, 443–488.
814 (doi:10.1108/JEIM-09-2014-0088)
- 815 36. Blut M, Chong A, Tsiga Z, Venkatesh V. 2021 Meta-Analysis of the Unified Theory of
816 Acceptance and Use of Technology (UTAUT): Challenging its Validity and Charting a
817 Research Agenda in the Red Ocean. *J. Assoc. Inf. Syst. Forthcom.* Available from:
818 <https://ssrn.com/abstract=3834872>
- 819 37. Jadil Y, Rana NP, Dwivedi YK. 2021 A meta-analysis of the UTAUT model in the mobile
820 banking literature: The moderating role of sample size and culture. *J. Bus. Res.* 132, 354–372.
821 (doi:10.1016/j.jbusres.2021.04.052)
- 822 38. Khechine H, Lakhel S, Ndjambou P. 2016 A meta-analysis of the UTAUT model: Eleven years
823 later: A meta-analysis of the UTAUT model: Eleven years later. *Can. J. Adm. Sci. Rev. Can.*
824 *Sci. Adm.* 33, 138–152. (doi:10.1002/cjas.1381)