1 **Manuscript: Journal of Experimental Psychology: Learning, Memory, and Cognition**

2 **Research Article**

3 Fast or Frugal, but not both: Decision Heuristics under Time Pressure

4 **Authors:**

5 Sebastian Bobadilla-Suarez[a][b][#] & Bradley C. Love[a][b]

6 [a] Department of Experimental Psychology, University College London, 26 Bedford Way,

7 London, UK, WC1H 0AP

8 [b] The Alan Turing Institute, 96 Euston Road, London, UK, NW1 2DB

9 [#] To whom correspondence should be addressed.

10 Email addresses:

11 sebastian.suarez.12@ucl.ac.uk, b.love@ucl.ac.uk

12 **5 color figures, 1 b&w figure**

13

14

15

16

17

18

Abstract

20 Heuristics are simple, yet effective, strategies that people use to make decisions. Because

21 heuristics do not require all available information, they are thought to be easy to implement

22 and to not tax limited cognitive resources, which has led heuristics to be characterized as

23 fast-and-frugal. We question this monolithic conception of heuristics by contrasting the

24 cognitive demands of two popular heuristics, Tallying and Take-the-Best. We contend that

25 heuristics that are frugal in terms of information usage may not always be fast because of

26 the attentional control required to implement this focus in certain contexts. In support of

27 this hypothesis, we find that Take-the-Best, while being more frugal in terms of

28 information usage, is slower to implement and fares worse under time pressure

29 manipulations than Tallying. This effect is then reversed when search costs for Take-the-

30 Best are reduced by changing the format of the stimuli. These findings suggest that

31 heuristics are heterogeneous and should be unpacked according to their cognitive demands

32 to determine the circumstances a heuristic best applies.

40 Some decision procedures are slow and information demanding whereas others are "fast

41 and frugal" (Gigerenzer & Gaissmaier, 2011), though for a dissenting opinion on the

42 prevalence of heuristic use see Newell, Weston, and Shanks (2003), and Newell (2005).

43 Consider a scenario in which a child suddenly crosses the street to get his ball. The driver

44 has less than a second to evaluate the situation and decide whether to press hard on the

45 brakes without swerving or press on the brakes and swerve onto the sidewalk. The former

46 option risks hitting the child and the latter option risks hitting other pedestrians. The

47 optimal decision depends on cues such as the speed of the vehicle, the distance from the car

48 to other people, the car's stopping distance, the number of people on the sidewalk, the

49 driver's ability, etc. What is the best way to integrate all this information quickly?

50 Representing and integrating all this information in an optimal manner may be impossible

51 or too time-consuming.

52 Alternatives to the computationally intensive decision strategies are commonly referred to

53 as "fast and frugal" heuristics (Todd & Gigerenzer, 2000). Heuristics are fast in that they

54 can be applied quickly and frugal in the sense that they use less information to make a

55 decision than more complex procedures that selectively weigh all information sources.

56 Despite the fact that heuristics use less information from the environment, in practice they

57 can perform very well, often surpassing regression approaches for certain decision

58 problems (Czerlinski, Gigerenzer, & Goldstein, 1999). Heuristics have been described as

59 "efficient cognitive processes, conscious or unconscious, that ignore part of the

60 information" (Gigerenzer & Gaissmaier, 2011) and as employing "a minimum of time,

61 knowledge, and computation to make adaptive choices in real environments." (Todd &

62 Gigerenzer, 2000). This characterization of heuristics differs from earlier accounts that cast

63 heuristics as imperfect approximations of rational decision procedures (Tversky &

64 Kahneman, 1974).

65 When complex decision strategies, such as multiple linear regression, cannot be

66 implemented because of resource constraints, such as lack of time, people might use

67 alternative strategies like heuristics (Todd & Gigerenzer, 2000). In this study, we

68 considered whether two popular heuristics differ from one another in cognitive processing

69 requirements as reflected by timing constraints. Specifically, we compared two popular

70 heuristics, Tallying (TAL) and Take-the-Best (TTB), which we introduce by way of

71 example below. These heuristics follow from previous work, such as the lexicographic

72 heuristics that only take into account the most discriminating attribute value (Fishburn,

73 1967; Tversky, 1969), the majority of confirming dimensions heuristic (Russo & Dosher,

74 1983), and the equal weights strategy (Dawes, 1979).

75 Suppose one wants to predict whether China or India will have higher gross domestic

76 product (GDP) growth based on their productive capabilities, natural resource wealth, and

77 the diversity in their exports, etc. The TAL heuristic chooses the country that bests the

78 other across the most measures. TAL does not selectively weigh cues as linear regression

79 does, but instead merely counts the number of cues favoring one alternative over the other

80 (Gigerenzer & Gaissmaier, 2011). On the other hand, the TTB heuristic chooses based on

81 the most predictive cue and only considers the next best cue when there is a tie. TTB

82 implies that cues are rank ordered in terms of their predictive validity in determining the

83 criterion (e.g., in predicting China or India). TTB sequentially searches until a

84 discriminating cue is found and, thus, may reach a decision after only considering the first

85 best cue. These two heuristics can both be effective in practice, but can differ in their

86 choices as shown in an example trial of our first experiment in Figure 1.

87 [Figure 1 here]

88 We hypothesize that these two heuristics differ in their cognitive processing requirements

89 such that what is fast and what is frugal is contingent on the cognitive processes invoked by

90 the environment. In Experiment 1, we predict that TTB will fare worse under time pressure

91 than TAL, whereas we predict the opposite pattern in Experiment 2. Although one might

92 expect TTB to be fast given that it samples very little information (Bröder & Gaissmaier,

93 2007; Khader et al., 2011), the cognitive demands in Experiment 1 should be high for TTB

94 users because the stimulus format invites an effortful sequential search procedure. In

95 contrast, the stimulus format in Experiment 2 reduces this search burden while making it

96 more difficult for TAL users to perform rapid summation operations.

97 Our experimental procedures are intended to expand the scope of inquiry by deviating from

98 the original formulation of TTB (Gigerenzer & Todd, 1999) in which decisions were made

99 from memory for environments; instead our method invites participants to make inference

100 from givens in the environment. Although memory demands are an important aspect of

101 heuristic application, our studies focus more on the attentional demands of applying

102 heuristics. Related previous efforts have noted that one subtle complexity of TTB is that it

103 requires a hierarchy of cue validities (Dougherty, Franco-Watkins, & Thomas, 2008; Juslin

104 & Persson, 2002) and that non-compensatory strategies such as TAL can be affected by cue

105 salience (Platzer & Bröder, 2012).

106   When compared to algorithms that are more computationally intensive such as linear

107   regression, TAL and TTB have various algorithmic aspects in common. Their most salient

108   similarity is that they both disregard covariance structure among cues (Parpart, Jones &

109   Love, under revision). Both heuristics also disregard relative cue weight magnitudes.

110   Despite these similarities, these two heuristics may differ in their cognitive demands. For

111   instance, TTB implements a search for the best discriminating cue which has been argued

112   to take time (Bröder & Gaissmaier, 2007) and a certain level of cognitive control due to

113   selective attention to the relevant cue, as well as inhibition of the irrelevant ones. Such

114   sequential control processes are thought to be effortful, serial in nature, and time-

115   consuming (Posner & Presti, 1987). In effect, TTB has the prescription that people will

116   embark on a serial search, which in our first experiment is a visually guided search. Such

117   visually guided searches are a common domain for research on top-down attentional control

118   mechanisms (Mozer & Baldwin, 2008; Wolfe, Cave, & Franzel, 1989).

119   TAL's cognitive requirements may be quite different than those used for TTB. We argue

120   that TAL requires people's ability to do quick summations over stimuli. Related research

121   on numerosity has shown that people can be very fast at doing these types of operations

122   without explicit counting (Feigenson, Dehaene, & Spelke, 2004). Given the right

123   representation format of cues, a TAL decision problem could be reduced to a low-level

124   perceptual categorization problem (e.g., Palmer, Huk, & Shadlen, 2005). However, if the

125   representation format is not suitable for such operations, as we manipulate in Experiment 2,

126   we predict that TAL's performance should be reduced.

127 In Experiment 1, we predicted that compliance with the TAL heuristic would be higher

128 than for TTB under time pressure conditions even though TAL considers more aspects of

129 the stimulus. In Experiment 2, we attempted to reverse this effect by eliminating search

130 costs for TTB and altering the stimulus format in a manner that obstructs quick summation

131 operations that favor TAL performance. Together, these two studies aim to establish that

132 heuristics should be understood not only in terms of how they perform in various

133 information environments, but also in terms of the cognitive processes they engage.

134 **Experiment 1**

135 In Experiment 1, we test the hypothesis that the TTB heuristic will not always lead to rapid

136 decisions because of the search costs and attentional control it can demand. For certain

137 stimulus formats, it should be possible for TAL to be faster than TTB. For example, color-

138 coded stimulus values should allow for rapid perceptual integration of cue values, making

139 TAL faster than TTB. Furthermore, TAL use should be unaffected by the randomization of

140 cue position across trials because TAL treats all cues identically and summates. In contrast,

141 TTB's search requirements should be increased by this randomization and should not

142 benefit from the color coding. Thus, we predict that TAL should be faster than TTB under

143 these conditions because of the basic resource requirements of each heuristic.

144 **Method**

145 *Participants.* Participants (206 total, 95 female) were recruited on Amazon Mechanical

146 Turk, an online study platform commonly used in psychological studies with good results

147 (Crump, McDonnell, & Gureckis, 2013). Participants were restricted to the United States of

148 America and assigned either to the TAL condition (107 total, 58 female) or the TTB

149   condition (99 total, 39 female). They received $2.50 to complete a 40 min. (approx.)

150   learning and decision making task and the best participant was offered a $20 bonus in each

151   condition. The average age of participants was 37.9 years ($SD = 12.56$). The study was

152   approved by the local UCL ethics committee.

153   ***Design and Materials.*** As a between participants manipulation, participants were explicitly

154   instructed to use either TAL or TTB in a two-alternative forced choice task (i.e., choosing

155   which country will have higher GDP). Each participant first completed a *practice* phase (72

156   trials) followed by a *test* phase, consisting of two blocks of trials. Whether the test block

157   was self-paced (72 trials) or speeded (72 trials) was counterbalanced across trials. Trial

158   order was randomized separately for each participant.

159   Within participants, the same 72 trials were used across these three (practice, self-paced

160   test, speeded test) experimental segments. Trials were designed such that one response was

161   consistent with TAL and the other with TTB. In other words, the heuristics disagreed on

162   every trial, which allowed for discriminating heuristic use. Perfect performance is

163   achievable for both heuristics because we measured compliance with an instructed heuristic

164   (compliance with heuristic is understood as % correct). Conversely, expected performance

165   under random responding was at chance (50%). The presentation order of the cues was

166   randomized on a trial-by-trial basis.

167   The 72 trials for each experimental segment consisted of three trial types in terms of three

168   difficulty levels. The 72 trials were equally partitioned into the three trial types resulting in

169   24 trials for each difficulty level. Difficulty was defined differently for each heuristic. For

170   TTB, difficulty level is referred to as Q1 (cue 1), Q2 (cue 2) and Q3 (cue 3), in order of

171   ascending difficulty. Q1 trials represent trials where retrieving the value for the best cue

172   was sufficient to respond in compliance with TTB. Similarly, Q2 trials require retrieving

173   the value for the second-best cue and Q3 trials require retrieving the value for the third-best

174   cue. For TAL, difficulty level is referred to as $\Delta 3$ (delta 3), $\Delta 2$ (delta 2), and $\Delta 1$ (delta 1), in

175   order of ascending difficulty. The $\Delta 3$ trial types represented trials where one option was a

176   better choice by a difference of three cue values, $\Delta 2$ trials represented trials where one

177   option was a better choice by a difference of two cue values, and the $\Delta 1$ trial types

178   represented trials where one option was a better choice by a difference of only one cue

179   value. Controlling for difficulty level in each heuristic was also a way to verify that indeed

180   participants were still trying to implement the respective heuristic in each experimental

181   condition. For both the TAL and TTB condition, trials were randomly sampled from the

182   same trial space (see Table S1 and Table S2 in section B of Supplemental Material

183   available online for more information on the exact sampling procedure).

184   The seven cues shown on each trial were economic statistics that would predict whether a

185   developing country would achieve higher Gross Domestic Product (GDP) levels the

186   following year when compared to another developing country (see section A of

187   Supplemental Material available online for a list of the statistics). These statistics were

188   artificially created and do not correspond to any real-world data in keeping with our focus

189   on heuristic compliance as opposed to real-world performance. We chose this domain in the

190   hopes that it would be familiar in a general sense while also unlikely to invite prior

191   knowledge to guide decisions (i.e., not strongly interfere with the instructed heuristic use).

192   **Procedure.** Subjects were shown the seven cues on each trial and asked to choose which

193   country would have higher GDP the following year. The two options were "Country A" or

194 "Country B". Participants saw seven economic statistics with a value for each country on

195 each trial (see Figure 1 for an example of stimulus presentation). Each statistic was framed

196 as a comparison between two options (i.e., two developing countries) where a green

197 checkmark was presented if superior to the other option, a red cross was presented if

198 inferior to the other option, and a black equal's sign was presented for ties between

199 countries on a given statistic. A list which ranked the statistics in order of importance, with

200 randomized order between participants, was presented on every trial in the *practice* phase

201 but not in the test phase. Before starting the *practice* phase, participants were provided with

202 detailed instructions on how to use the corresponding heuristic for the condition they had

203 been assigned to (either TAL or TTB).

204 In the *practice* phase, the participant would see the seven statistics with a value for each

205 country as well as the list which ranked the statistics in order of importance. After making a

206 response, immediate feedback was provided on the screen. If the participant responded in

207 accordance with the heuristic, the screen would show "Good! You understood the rule!", in

208 addition to an explanation of why the choice was correct according to the heuristic. If the

209 participant did not respond in accordance with the heuristic, the screen would show "Bad.

210 You did not understand the rule.", in addition to an explanation of why the choice was

211 incorrect according to the heuristic. The statistics with the values for each country were left

212 on screen during feedback but the list of ranked statistics was not. The list was not

213 presented during feedback to discourage reliance on the list and incentivize better

214 engagement with the task. Participants were not directly questioned about their knowledge

215 of the cue validities. The feedback was presented until the participant decided to move on

216 to the next trial followed by a presentation of a blank white screen for 500 milliseconds.

217   Subsequently, the participant would enter either a block with time pressure or a block

218   without time pressure of the *test* phase. The list which ranked the statistics in order of

219   importance was removed for the *test* phase. For the block without time pressure, the

220   participant was asked to answer in accordance with the heuristic that had been practiced on

221   previously, only this time without feedback. The inter-trial interval was 1500 milliseconds

222   with the first second showing "Thank You!" followed by a presentation of a blank white

223   screen for 500 milliseconds. For the block with time pressure, the participant was also

224   asked to answer in accordance with the heuristic practiced on previously and without

225   feedback. Then, instructions stating that the participant only had 2000 milliseconds to

226   respond on each trial were provided. The inter-trial interval was also 1500 milliseconds

227   with the first second showing "Thank You!" followed by a presentation of a blank white

228   screen for 500 milliseconds except when the participant reached the 2000 milliseconds

229   deadline. If the participant reached the 2000 milliseconds deadline, the screen would

230   explain why it was important for them to adhere to the imposed deadline. This screen was

231   presented for 10000 milliseconds followed by a presentation of a blank white screen for

232   500 milliseconds.

233   **Results**

234   [Figure 2 here]

235   *Exclusion Criteria.* Exclusion criteria were based on the responses made on the second half

236   of the *practice* phase (36 trials) and number of missed responses when under time pressure.

237   Participants with performance under 90% in the second half of the *practice* phase or over

238   16 missed responses when under time pressure were considered outliers and excluded from

239   all subsequent analysis (26 in the TTB condition, 1 in the TAL condition). Including the

240   exclusion criteria from the *practice* phase, this resulted in a total of 26.26% of participants

241   who were excluded from all further analyses in the TTB condition and 1% of participants

242   excluded from the TAL condition. Analyses were also run without any exclusion[1] and this

243   did not change any conclusions from the analyses shown below. All analyses that follow

244   use people who passed exclusion ($n = 179$) to provide a more stringent evaluation of our

245   predictions. All response time analyses were calculated with median response times. For a

246   presentation of results from the *practice* phase, please refer to section D in Supplemental

247   Material available online.

248   ***Test Phase.*** TTB participants had lower compliance than TAL, mostly because they were

249   affected more by the time pressure manipulation (see panel A in Figure 2). Proportion of

250   compliance was analyzed using a 2x2x3 mixed-design ANOVA with a between-subject

251   factor of heuristic (TAL, TTB), a within-participants factor of time pressure (present or

252   absent), and a within-participants factor of trial difficulty (3 levels of trial difficulty). Main

253   effects were observed for heuristic, $F(1, 177) = 285.89$, $p < .001$, $\eta^2 = .62$, time pressure,

254   $F(1, 177) = 352.62$, $p < .001$, $\eta^2 = .67$, and trial difficulty, $F(2, 354) = 13.85$, $p < .001$, $\eta^2 =$

255   .07, as well as an interaction between heuristic and time pressure, $F(1, 177) = 329.37$, $p <$

256   .001, $\eta^2 = .65$, an interaction between heuristic and trial difficulty, $F(2, 354) = 4.53$, $p =$

257   .011, $\eta^2 = .03$, an interaction between time pressure and trial difficulty, $F(2, 354) = 18.66$, p

258   $< .001$, $\eta^2 = .10$, and the three-way interaction between heuristic, time pressure, and trial

259   difficulty, $F(2, 354) = 14.87$, $p < .001$, $\eta^2 = .08$. These results directly support our

260   hypothesis by showing that TTB has lower compliance than TAL, especially under time

261 pressure. The three-way interaction highlights the asymmetric effect of the time pressure

262 manipulation on both heuristics' difficulty levels.

263 TTB participants responded more slowly than TAL participants and were markedly slower

264 on difficulty trials (see panels D, E, and F in Figure 2). Response times were analyzed

265 using a 2x2x3 mixed-design ANOVA with a between-subject factor of heuristic (TAL,

266 TTB), a within-participants factor of time pressure (present or absent) and with a within-

267 participants factor of trial difficulty (3 levels of trial difficulty). Individual response times

268 were calculated as median response times. Main effects were observed for heuristic, $F(1,$

269 $177) = 382.41, p < .001, \eta^2 = .68$, time pressure, $F(1, 177) = 370.92, p < .001, \eta^2 = .68$, and

270 trial difficulty, $F(2, 354) = 470.60, p < .001, \eta^2 = .73$, as well as an interaction between

271 heuristic and time pressure, $F(1, 177) = 313.50, p < .001, \eta^2 = .64$, an interaction between

272 heuristic and trial difficulty, $F(2, 354) = 248.39, p < .001, \eta^2 = .58$, an interaction between

273 time pressure and trial difficulty, $F(2, 354) = 285.34, p < .001, \eta^2 = .62$, and the three-way

274 interaction between heuristic, time pressure, and trial difficulty, $F(2, 354) = 265.66, p <$

275 $.001, \eta^2 = .60$. The main effect of heuristic shows that TTB is a slower heuristic than TAL

276 and directly supports our main hypothesis. The interaction between heuristic and time

277 pressure suggests that TAL can accommodate to time pressure more readily than TTB can.

278 The three-way interaction shows that trial difficulty still shows differences between

279 heuristics in both blocks, with and without time pressure. This gives reassurance that

280 participants were still engaged and attempting to implement the respective heuristic even

281 for blocks with time pressure.

282 [Figure 3 here]

283   *Model-Based Analysis.* For the TTB participants, three models were fit to determine how

284   many cues participants tended to successfully incorporate under self-paced and speeded

285   conditions. The models are reduced versions of TTB. Only three models were considered;

286   TTB first-cue (only considers the first cue presented on all trials), TTB first-two-cues (only

287   considers the first and second cue presented on all trials), and TTB-first-three-cues (only

288   considers the first, second and third cue presented on all trials). As shown in Figure 3, the

289   number of cues successfully incorporated tended to be higher under self-paced conditions,

290   suggesting that the TTB procedure broke down under time pressure. Fisher's exact test

291   found that the model contingencies differed across self-paced and speeded conditions ($p <$

292   .001).

293   Only models with three or fewer cues were considered because only the first three cues

294   were needed to comply with TTB in the study design. Which model best fit each participant

295   was determined by comparing the agreement in choices between human and model. For the

296   one and two cues models, agreement was scored as .5 (i.e., the expectation for a random

297   guess) for trials that exceeded the cues encoded by the model because TTB guesses when

298   there is no discriminating cue. All three models are parameter-free and, thus, are readily

299   comparable. For the speeded blocks, the agreement in choices between human and model

300   were significantly above chance (.5) for the best cue model, $t(72) = 2.81$, $p = .006$, but not

301   for the two best cues model, $t(72) = 1.68$, $p = .097$, or three best cues model, $t(72) = 1.21$, p

302    $= .229$. Additional model fits confirmed that participants were using the cues themselves, as

303   opposed to their positions (see Figure S1 of section C in Supplemental Material available

304   online).

305   **Discussion**

306 Experiment 1 strongly supported our a priori hypothesis: TTB, however frugal, can be slow

307 given its requirements with respect to attentional control and search costs. This can be seen

308 by the lower proportion of compliance of the TTB heuristic under time pressure, whereas

309 the TAL heuristic maintains the same proportion of compliance for both the self-paced and

310 time pressure conditions. We assume that this effect is due to both the trial-by-trial

311 randomization of the cue positions and the format of the values shown (e.g. red crosses and

312 green checkmarks). These results highlight the importance of considering not only the

313 structure of the environment, but also the cognitive demands of heuristics when pairing

314 heuristics with tasks. Experiment 2 attempts to reverse the TAL advantage observed in

315 Experiment 1 by using a stimulus format that favors TTB by reducing search costs (helpful

316 for TTB) and the possibility of perceptual summation (harmful for TAL).

317 **Experiment 2**

318 In Experiment 1, TAL's advantage over TTB was striking – the less frugal (in terms of

319 cues) heuristic was faster. We predicted this pattern of results by considering the cognitive

320 processes each heuristic requires. In Experiment 2, we attempted to build on this result by

321 creating conditions that should favor TTB over TAL. In Experiment 2, the search costs for

322 TTB are reduced by ordering stimulus cues by their validity. Unlike Experiment 1, the cues

323 are no longer color-coded, which prevents TAL from utilizing perceptual summation

324 operations. Mirroring Experiment 1, the format choice in Experiment 2 should strongly

325 favor TAL over TTB.

326 **Method**

327   *Participants.* Participants (194 total, 128 female) were also recruited on Amazon

328   Mechanical Turk. Participants were restricted to the United States of America and assigned

329   either to the TAL condition (97 total, 63 female) or the TTB condition (97 total, 65 female).

330   They received $2.50 to complete a 40 min. (approx.) learning and decision making task and

331   the best participant was offered a $20 bonus in each condition. The average age of

332   participants was 41.5 years ($SD = 12.34$). The study was approved by the local UCL ethics

333   committee.

334   *Design and Materials.* The design and materials for the experiment were similar to

335   Experiment 1 with some critical differences. Rather than order cues randomly on each trial,

336   cues were always ordered by their validity. Although the stimuli consisted of the same

337   economic statistics as in Experiment 1, they were no longer color-coded (see Figure 4).

338   Instead, pairs of positive and negative adjectives (e.g., better and worse) were used (see

339   section A of Supplemental Material available online for a list of the adjectives). The word

340   "equal" was presented for ties.

341   [Figure 4 here]

342   *Procedure.* As in Experiment 1, participants were shown the seven cues on each trial and

343   asked to choose which country ("Country A" or "Country B") would have higher GDP the

344   following year. Participants saw seven economic statistics with a value for each country on

345   each trial (see Figure 4 for an example of stimulus presentation).

346   **Results**

347   [Figure 5 here]

348   *Exclusion Criteria.* Exclusion criteria were the same as in Experiment 1. Participants with

349   performance under 90% in the second half of the practice phase or over 16 missed

350   responses when under time pressure were considered outliers and excluded from all

351   subsequent analysis (8 in the TTB condition, 13 in the TAL condition). Including the

352   exclusion criteria from the practice phase, this resulted in a total of 8.25% of participants

353   who were excluded from all further analyses in the TTB condition and 13.4% of

354   participants excluded from the TAL condition. Analyses were also run without any

355   exclusion[2] and this did not change any conclusions from the analyses shown below. All

356   analyses that follow use people who passed exclusion ($n = 173$) to provide a more stringent

357   evaluation of our predictions. All response time analyses were calculated with median

358   response times.

359   *Test Phase.* TAL participants had lower compliance than TTB, mostly because they were

360   affected more by the time pressure manipulation (see panel A in Figure 5). Proportion of

361   compliance was analyzed using a 2x2x3 mixed-design ANOVA with a between-subject

362   factor of heuristic (TAL, TTB), a within-participants factor of time pressure (present or

363   absent), and a within-participants factor of trial difficulty (3 levels of trial difficulty). Main

364   effects were observed for heuristic, $F(1, 171) = 134.83$, $p < .001$, $\eta^2 = .44$, time pressure,

365    $F(1, 171) = 244.74$, $p < .001$, $\eta^2 = .59$, and trial difficulty, $F(2, 342) = 21.51$, $p < .001$, $\eta^2 =$

366   .11, as well as an interaction between heuristic and time pressure, $F(1, 171) = 135.95$, $p <$

367   .001, $\eta^2 = .44$, an interaction between heuristic and trial difficulty, $F(2, 342) = 109.75$, $p =$

368   .011, $\eta^2 = .39$, an interaction between time pressure and trial difficulty, $F(2, 342) = 19.89$, p

369    $< .001$, $\eta^2 = .10$, and the three-way interaction between heuristic, time pressure, and trial

370   difficulty, $F(2, 354) = 29.42$, $p < .001$, $\eta^2 = .15$. These results directly support our

371 hypothesis by showing that when search costs are eliminated and the representation of

372 stimuli values obstruct quick summations, then TAL has lower compliance than TTB,

373 especially under time pressure. The three-way interaction highlights the asymmetric effect

374 of the time pressure manipulation on both heuristics' difficulty levels.

375 TAL participants responded more slowly than TTB participants and were markedly slower

376 on difficult trials (see panels D, E, and F in Figure 5). Response times were analyzed using

377 a 2x2x3 mixed-design ANOVA with a between-subject factor of heuristic (TAL, TTB), a

378 within-participants factor of time pressure (present or absent) and with a within-participants

379 factor of trial difficulty (3 levels of trial difficulty). Individual response times were

380 calculated as median response times. Main effects were observed for heuristic, $F(1, 171) =$

381 $171.41, p < .001, \eta^2 = .50$, time pressure, $F(1, 171) = 291.79, p < .001, \eta^2 = .63$, and trial

382 difficulty, $F(2, 342) = 110.18, p < .001, \eta^2 = .39$, as well as an interaction between heuristic

383 and time pressure, $F(1, 171) = 186.14, p < .001, \eta^2 = .52$, an interaction between heuristic

384 and trial difficulty, $F(2, 342) = 15.80, p < .001, \eta^2 = .09$, an interaction between time

385 pressure and trial difficulty, $F(2, 342) = 40.50, p < .001, \eta^2 = .19$, and the three-way

386 interaction between heuristic, time pressure, and trial difficulty, $F(2, 342) = 22.11, p <$

387 $.001, \eta^2 = .11$. The main effect of heuristic shows that TAL is a slower heuristic than TTB

388 and directly supports our main hypothesis. For this experiment, the interaction between

389 heuristic and time pressure suggests that TTB can accommodate to time pressure more

390 readily than TAL can. The three-way interaction shows that trial difficulty finds differences

391 between heuristics in both blocks, with and without time pressure, which provides

392 reassurance that participants were engaged and attempting to implement the respective

393 heuristic even for blocks with time pressure.

394   [Figure 6 here]

395   *Model-Based Analysis.* As in Experiment 1, three models were fit for the TTB condition to

396   determine how many cues participants tended to successfully incorporate under self-paced

397   and speeded conditions. As shown in Figure 6, the number of cues successfully

398   incorporated was not different under self-paced conditions, suggesting that the TTB

399   procedure was unaffected by time pressure in this experiment. Fisher's exact test found that

400   the model contingencies did not differ across self-paced and speeded conditions ($p = .840$).

401   Model agreement with human behavior followed the same rationale as in Experiment 1 (see

402   above).

403   **Discussion**

404   In accord with our hypothesis, Experiment 2 reversed the effect found in Experiment 1 –

405   TAL was more strongly affected by time pressure than TTB. TAL had lower compliance,

406   particularly under time pressure. One interesting side note is that TTB compliance was

407   better for difficult TTB trials, though these trials were also slower which indicates a speed-

408   accuracy tradeoff. Overall, Experiment 2's results reinforce the conclusions from

409   Experiment 1, namely that heuristics are not monolithic and need to be decomposed into

410   their constituent cognitive processes to understand and predict their performance under

411   various task conditions.

412   **General Discussion**

413   Heuristics are often contrasted collectively with other decision procedures, which could

414   give the impression that heuristics form a uniform class. Instead, we argued that heuristics

415   are best understood in terms of their constituent cognitive processes. When decomposed

416   into these processes, it becomes possible to predict when people will be successful in

417   applying given the demands of the task.

418   In Experiment 1, stimulus format was chosen to suite the demands of TAL at the expense

419   of TTB. We found that the more frugal heuristic (i.e., requiring less information), TTB, was

420   slower than the less frugal heuristic, TAL. We hypothesized this advantage was due to the

421   search costs incurred by TTB (cue order was randomized) and the ability of TAL to take

422   advantage of fast perceptual summation operations (cue values were color-coded). In

423   Experiment 2, these advantages and disadvantages were reversed by ordering cues by their

424   validity and removing color-coding. As predicted, we observed an advantage of TTB over

425   TAL.

426   Our results mirror previous results using different experimental procedures. As in

427   Experiment 1, previous work finds that the time requirements of TTB are dependent on the

428   number of cues retrieved from memory (Bergert & Nosofsky, 2007; Bröder & Gaissmaier,

429   2007; Khader et al., 2011; Lee & Cummins, 2004; Newell & Shanks, 2003). These results

430   dovetail with the notion that TTB's implementation involves greater complexity than its

431   algorithmic description suggests (Dougherty et al., 2008; Juslin & Persson, 2002). The

432   findings of Experiment 1 also resemble results of another study where more cue

433   information was processed quicker, as long as the added information would increase

434   coherence among cues (Glöckner and Betsch, 2012). Conversely, participants difficulties

435   with TAL in Experiment 2 mirror previous research exploring when people choose to adopt

436   a compensatory strategy (Bröder, 2003; Bröder & Schiffer, 2006; Platzer & Bröder, 2012;

437   Platzer, Bröder, & Heck, 2014). A surprising result in Experiment 2 was that proportion of

438   compliance for TTB varied with difficulty in the opposite direction of that which was

439 predicted. However, the tradeoff observed between speed and accuracy limits the

440 interpretation of this result (e.g. Ratcliff & McKoon, 2008).

441 Our results support the view that heuristics should be unpacked in terms of the

442 psychological processes they rely upon. One important area for continued research is

443 understanding how the representational format of cues and their values affect performance

444 (Bröder & Schiffer, 2006; Gigerenzer & Hoffrage, 1995). Ideally, this research would

445 expand to consider other heuristics and to conditions under which participants must choose

446 which strategy to adopt for a given environment (Rieskamp & Otto, 2006). One possibility

447 is that participants are adaptive and choose the heuristic procedure that minimizes the

448 cognitive demands given the stimulus format. The set of candidate solution procedures

449 could extend beyond heuristic procedures to include gist representations (Peters et al.,

450 2009; Reyna, 2008) and parallel constraint satisfaction approaches (Glockner & Betsch,

451 2008). Alternatively, there is also recent work suggesting that heuristics can be viewed as a

452 special case of Bayesian inference (Parpart, Jones & Love, under revision).

453 In contrast to findings suggesting that participants tend to default to strategies similar to

454 TTB under time pressure (Ben Zur & Breznitz, 1981; Payne, Bettman, & Johnson, 1993),

455 we found in Experiment 1 that compliance with TTB was dramatically reduced by time

456 pressure, whereas TAL was largely unaffected. As discussed, our choice of stimuli played

457 an important role in determining these results, as observed when the stimuli were changed

458 in Experiment 2. Additionally, previous research has found that people tend to use

459 strategies more like TTB when the decision is largely made from memory (Bröder &

460 Schiffer, 2003). In contrast, people tend to use strategies more like TAL when the stimulus

461 conveys key information about the options (Bröder & Schiffer, 2003). Our findings, which

462 follow from the predictions of our attentional hypothesis, suggest a possible explanation for

463 previous results. Likewise, our account may help explain the costs of learning TTB.

464 An important limitation of these two studies is that they differ from past research in that we

465 directly instructed subjects on which strategy to use, either TAL or TTB, and how to use

466 them. For the same reasons as another study (Khader et al., 2011), we preferred to instruct

467 participants to use these heuristics –instead of studying their spontaneous use– as a

468 preventive measure for strategy switching. This detail provides the experimenters with

469 more control over the task structure but it limits both comparison with other studies that do

470 not employ instructed strategy use and the interpretability of these results for real-world

471 applications.

472 In conclusion, more frugal heuristics will not necessarily be faster to implement than less

473 frugal ones. Similarly, less frugal strategies can be fast given the right stimuli format. To

474 understand how heuristics will perform across situations, the cognitive mechanisms that

475 underlie heuristics need to be specified.

476 In light of these results, one important challenge is matching heuristics to decision

477 environments according to multiple factors, such as stimulus format, time available, and

478 cognitive resources. One practical implication of this line of research is to inform those

479 who aim to best train people, such as emergency responders, to appropriately match

480 decision heuristics to the information and cognitive demands of the task. Although there is

481 much more yet to do, our results make clear that heuristics are not universally "fast and

482 frugal" but in cases may be "fast or frugal".

483 **Author Contributions**

484 BCL developed the study concept. All authors contributed to the study design. Testing and

485 data collection were performed by SBS. SBS performed the data analysis and interpretation

486 under the supervision of BCL. SBS drafted the manuscript, and BCL provided critical

487 revisions. All authors approved the final version of the manuscript for submission.

499 **Supplemental Material and Open Access Data**

500 For additional supporting information and open access to the data please visit:

501 https://osf.io/xnbz4/

502 **Notes**

1. We also conducted all analyses with participants that had 100% performance in the second half of the *practice* phase (95 participants in the TAL condition and 25 in the TTB condition) and this did not change any conclusions presented here.

2. We also conducted all analyses with participants that had 100% performance in the second half of the *practice* phase (44 participants in the TAL condition and 48 in the TTB condition) and this did not change any conclusions presented here.

References

Ben Zur, H., & Breznitz, S. J. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, *47*(2), 89–104. http://doi.org/10.1016/0001-6918(81)90001-9

Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *33*(1), 107–129. http://doi.org/10.1037/0278-7393.33.1.107

Bröder, A. (2003). Decision making with the "adaptive toolbox": influence of environmental structure, intelligence, and working memory load. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *29*(4), 611–625. http://doi.org/10.1037/0278-7393.29.4.611

Bröder, A., & Gaissmaier, W. (2007). Sequential processing of cues in memory-based multiattribute decisions. *Psychonomic Bulletin & Review*, *14*(5), 895–900. http://doi.org/10.3758/BF03194118

Bröder, A., & Schiffer, S. (2003). Take the best versus simultaneous feature matching: probabilistic inferences from memory and effects of representation format. *Journal of*

Experimental Psychology. *General*, *132*(2), 277–293. http://doi.org/10.1037/0096-
3445.132.2.277

Bröder, A., & Schiffer, S. (2006). Stimulus format and working memory in fast and frugal
strategy selection. *Journal of Behavioral Decision Making*, *19*(4), 361–380.
http://doi.org/10.1002/bdm.533

Crump, M. J. C., McDonnell, J. V, & Gureckis, T. M. (2013). Evaluating Amazon's
Mechanical Turk as a tool for experimental behavioral research. *PloS One*, *8*(3),
e57410.

Czerlinski, J., Gigerenzer, G., & Goldstein, D. G. (1999). How good are simple heuristics?
In *Simple heuristics that make us smart* (pp. 97–118). New York, NY, US: Oxford
University Press.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making.
American Psychologist, *34*(7), 571.

Dougherty, M. R., Franco-Watkins, A. M., & Thomas, R. (2008). Psychological
plausibility of the theory of probabilistic mental models and the fast and frugal
heuristics. *Psychological Review*, *115*(1), 199–213. http://doi.org/10.1037/0033-
295X.115.1.199

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in
Cognitive Sciences*, *8*(7), 307–314. http://doi.org/10.1016/j.tics.2004.05.002

Fishburn, P. C. (1967). Methods of estimating additive utilities. *Management Science*,
13(7), 435–453.

Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of*

547   *Psychology*, *62*, 451–482. http://doi.org/10.1146/annurev-psych-120709-145346

548   Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without

549   instruction: frequency formats. *Psychological Review*, *102*(4), 684.

550   Gigerenzer, G., & Todd, P. M. (1999). Fast and frugal heuristics: The adaptive toolbox. In

551   *Simple heuristics that make us smart* (pp. 3–34). Oxford University Press.

552   Glockner, A., & Betsch, T. (2008). Multiple-Reason Decision Making Based on Automatic

553   Processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,

554   *34*(5), 1055. http://doi.org/10.1037/0278-7393.34.5.1055

555   Juslin, P., & Persson, M. (2002). PROBabilities from EXemplars (PROBEX): A "lazy"

556   algorithm for probabilistic inference from generic knowledge. *Cognitive Science*,

557   *26*(5), 563–607. http://doi.org/10.1016/S0364-0213(02)00083-6

558   Kahneman, D., Slovic, P., & Tversky, A. (1974). Judgment under uncertainty: heuristics

559   and biases. *Science*, *185*(4157), 1124–1131.

560   http://doi.org/10.1126/science.185.4157.1124

561   Khader, P. H., Pachur, T., Meier, S., Bien, S., Jost, K., & Rösler, F. (2011). Memory-based

562   Decision-making with Heuristics: Evidence for a Controlled Activation of Memory

563   Representations. *Journal of Cognitive Neuroscience*, *23*(11), 3540–3554.

564   http://doi.org/10.1162/jocn_a_00059

565   Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making:

566   unifying the "take the best" and the "rational" models. *Psychonomic Bulletin &*

567   *Review*, *11*(2), 343–352. http://doi.org/10.3758/BF03196581

568   Mandler, G., & Shebo, B. J. (1982). Subitizing: an analysis of its component processes.

569   *Journal of Experimental Psychology: General*, *111*(1), 1.

570   Mozer, M., & Baldwin, D. (2008). Experience-Guided Search: A Theory of Attentional

571   Control. *Advances in Neural Information Processing Systems 20*, 1033–1040.

572   http://doi.org/10.1037/e527342012-076

573   Newell, B. R. (2005). Re-visions of rationality? *Trends in Cognitive Sciences*, *9*(1), 11–15.

574   http://doi.org/10.1016/j.tics.2004.11.005

575   Newell, B. R., & Shanks, D. R. (2003). Take the best or look at the rest? Factors

576   influencing "one-reason" decision making. *Journal of Experimental Psychology*.

577   *Learning, Memory, and Cognition*, *29*(1), 53–65. http://doi.org/10.1037/0278-

578   7393.29.1.53

579   Newell, B. R., Weston, N. J., & Shanks, D. R. (2003). Empirical tests of a fast-and-frugal

580   heuristic: Not everyone "takes-the-best." *Organizational Behavior and Human*

581   *Decision Processes*, *91*(1), 82–96. http://doi.org/10.1016/S0749-5978(02)00525-3

582   Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the

583   speed and accuracy of a perceptual decision. *Journal of Vision*, *5*(5), 376–404.

584   http://doi.org/10.1167/5.5.1

585   Parpart, P., Jones, M., & Love, B.C. Heuristics as extreme bayesian priors. Manuscript

586   submitted for publication.

587   Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). The adaptive decision maker. *The*

588   *Adaptive Decision Maker*, *45*(7), 352. http://doi.org/10.1057/jors.1994.133

589   Peters, E., Dieckmann, N. F., Västfjäll, D., Mertz, C. K., Slovic, P., & Hibbard, J. H.

590   (2009). Bringing meaning to numbers: the impact of evaluative categories on

591   decisions. *Journal of Experimental Psychology. Applied*, *15*(3), 213–227.

592   http://doi.org/10.1037/a0016978

593   Platzer, C., & Bröder, A. (2012). Most people do not ignore salient invalid cues in memory-

594   based decisions. *Psychonomic Bulletin & Review*, *19*(4), 654–661.

595   http://doi.org/10.3758/s13423-012-0248-4

596   Platzer, C., Bröder, A., & Heck, D. W. (2014). Deciding with the eye: How the visually

597   manipulated accessibility of information in memory influences decision behavior.

598   *Memory & Cognition*, *42*(4), 595–608.

599   Posner, M. I., & Presti, D. E. (1987). Selective attention and cognitive control. *Trends in*

600   *Neurosciences*, *10*(1), 13–17. http://doi.org/10.1016/0166-2236(87)90116-0

601   Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: theory and data for two-

602   choice decision tasks. *Neural Computation*, *20*(4), 873–922.

603   Reyna, V. F. (2008). A theory of medical decision making and health: fuzzy trace theory.

604   *Medical Decision Making*.

605   Rieskamp, J., & Otto, P. E. (2006). SSL: a theory of how people learn to select strategies.

606   *Journal of Experimental Psychology: General*, *135*(2), 207.

607   Russo, J. E., & Dosher, B. A. (1983). Strategies for multiattribute binary choice. *Journal of*

608   *Experimental Psychology: Learning, Memory, and Cognition*, *9*(4), 676–696.

609   http://doi.org/10.1037/0278-7393.9.4.676

610   Todd, P. M., & Gigerenzer, G. (2000). Précis of Simple heuristics that make us smart. *The*

611   *Behavioral and Brain Sciences*, *23*(5), 727-741-780.

612   http://doi.org/10.1017/S0140525X00003447

613   Tversky, A. (1969). Intransitivity of preferences. *Psychological Review, 76*(1), 31.

614   Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the

615   feature integration model for visual search. *Journal of Experimental Psychology*.

616   *Human Perception and Performance, 15*(3), 419–433. http://doi.org/2527952

617   **Figure 1.**

618   Example trial for the *practice* phase of Experiment 1. Participants were required to choose

619   the country that would have higher gross domestic product (GDP) for the following year

620   depending on the values of the economic statistics presented. Participants were assigned

621   either to the Tallying (TAL) or the Take-the-Best (TTB) condition and asked to respond

622   according to what the respective heuristic would predict. In this example trial, TAL would

623   choose Country B since four cues have superior values (green checkmarks) whereas

624   Country A is only superior on one cue. In contrast, TTB would choose Country A since the

625   value for the best-discriminating cue, which in this case is the third most predictive cue

626   (i.e., "Increased employment opportunities"), is superior to Country B.

627   **Figure 2.**

628   Main results after applying exclusion criteria for Experiment 1 ($n = 179$). The figure shows

629   the proportion of compliance with (a) both heuristics for blocks without time pressure (blue

630   bars) and blocks with time pressure (green bars), (b) the proportion of compliance in the

631   TAL condition for blocks with and without time pressure displayed by degrees of difficulty

632   (Deltas), and (c) the proportion of compliance in the TTB condition for blocks with and

633   without time pressure displayed by degrees of difficulty (Qs). Shows the response times

634   for (d) both heuristics in both blocks with and without time pressure, (e) the response times

635   in the TAL condition for blocks with and without time pressure displayed by degrees of

636   difficulty (Deltas), and (f) the response times in the TTB condition for blocks with and

637   without time pressure displayed by degrees of difficulty (Qs). For all panels, error bars are

638   95% within-participants confidence intervals.

639   **Figure 3.**

640   Model-based analyses reveal cue usage for Take-the-Best (TTB) for Experiment 1. The

641   differences in how many cues were used in blocks without time pressure (blue bars) and

642   with time pressure (red bars) for the TTB condition. The vertical axis shows the proportion

643   of participants that were best described by one of the three models. Absent time pressure

644   (blue bars), most participants' decisions were best fit by the full model using the best three

645   cues, whereas with time pressure (red bars) some participants appeared to rely on fewer

646   cues. Error bars are 95% confidence intervals.

647   **Figure 4.**

648   Example trial for the *test* phase of Experiment 2. Participants were required to choose the

649   country that would have higher GDP for the following year depending on the values of the

650   economic statistics presented. Participants were assigned either to the Tallying (TAL) or

651   the Take-the-Best (TTB) condition and asked to respond according to what the respective

652   heuristic would predict. In this example trial, TAL would choose Country A since four cues

653   have superior values (positive adjectives) whereas Country B is only superior on three cues.

654   In contrast, TTB would choose Country B since the value for the best-discriminating cue,

655   which in this case is also the most predictive cue (i.e., "Increased employment

656   opportunities"), is superior to Country A.

657   **Figure 5.**

658   Main results after applying exclusion criteria for Experiment 2 (n = 173). The figure shows

659   the proportion of compliance with (a) both heuristics for blocks without time pressure (blue

660   bars) and blocks with time pressure (green bars), (b) the proportion of compliance in the

661   TAL condition for blocks with and without time pressure displayed by degrees of difficulty

662   (Deltas), and (c) the proportion of compliance in the TTB condition for blocks with and

663   without time pressure displayed by degrees of difficulty (Qs). Shows the response times

664   for (d) both heuristics in both blocks with and without time pressure, (e) the response times

665   in the TAL condition for blocks with and without time pressure displayed by degrees of

666   difficulty (Deltas), and (f) the response times in the TTB condition for blocks with and

667   without time pressure displayed by degrees of difficulty (Qs). For all panels, error bars are

668   95% within-participants confidence intervals.

669   **Figure 6.**

670   Model-based analyses reveal cue usage for Take-the-Best (TTB). The differences in how

671   many cues were used in blocks without time pressure (blue bars) and with time pressure

672   (red bars) for the TTB condition. The vertical axis shows the proportion of participants that

673   were best described by one of the three models. Most participants' decisions were best fit

674   by the full model using the best three cues for both speeded and self-paced blocks. Error

675   bars are 95% confidence intervals.

676

Figure 1

Country A	Stats	Country B
✓	Increased employment opportunities	✗
=	Competitiveness in medium enterprises	=
✗	Public investment in infrastructure	✓
✗	Development of civic participation	✓
✗	Increased life expectancy for women	✓
=	Price stability in cheap basic goods	=
✗	Decreased rates of infectious diseases	✓

Stats ranking:
1. Competitiveness in medium enterprises
2. Price stability in cheap basic goods
3. Increased employment opportunities
4. Public investment in infrastructure
5. Decreased rates of infectious diseases
6. Increased life expectancy for women
7. Development of civic participation

Figure 2   Click here to download Figure figure2.png ⬇

Proportion of Compliance & Response Times

Figure 3

Number of Cues used in Take-the-Best
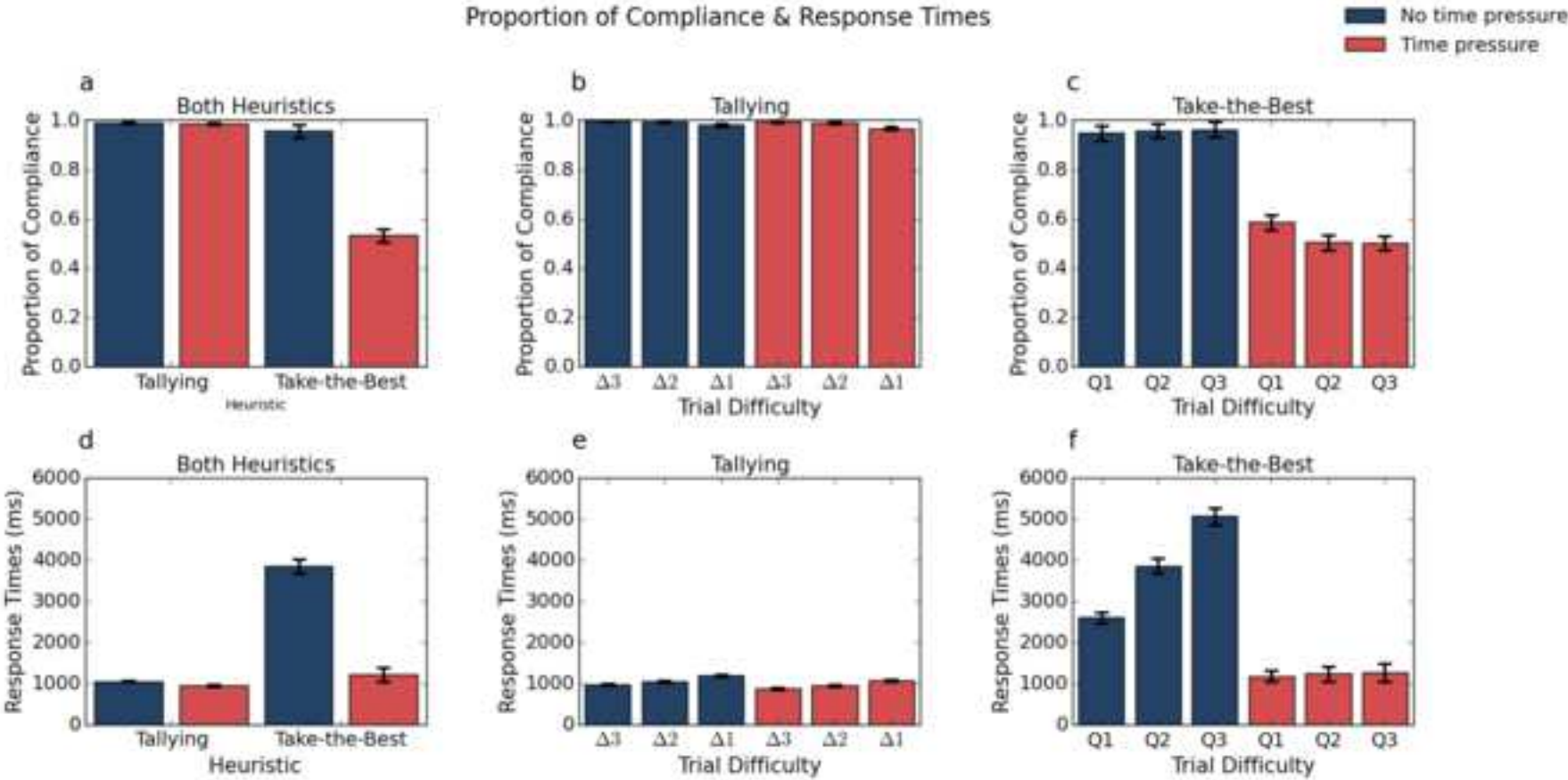
Figure 4
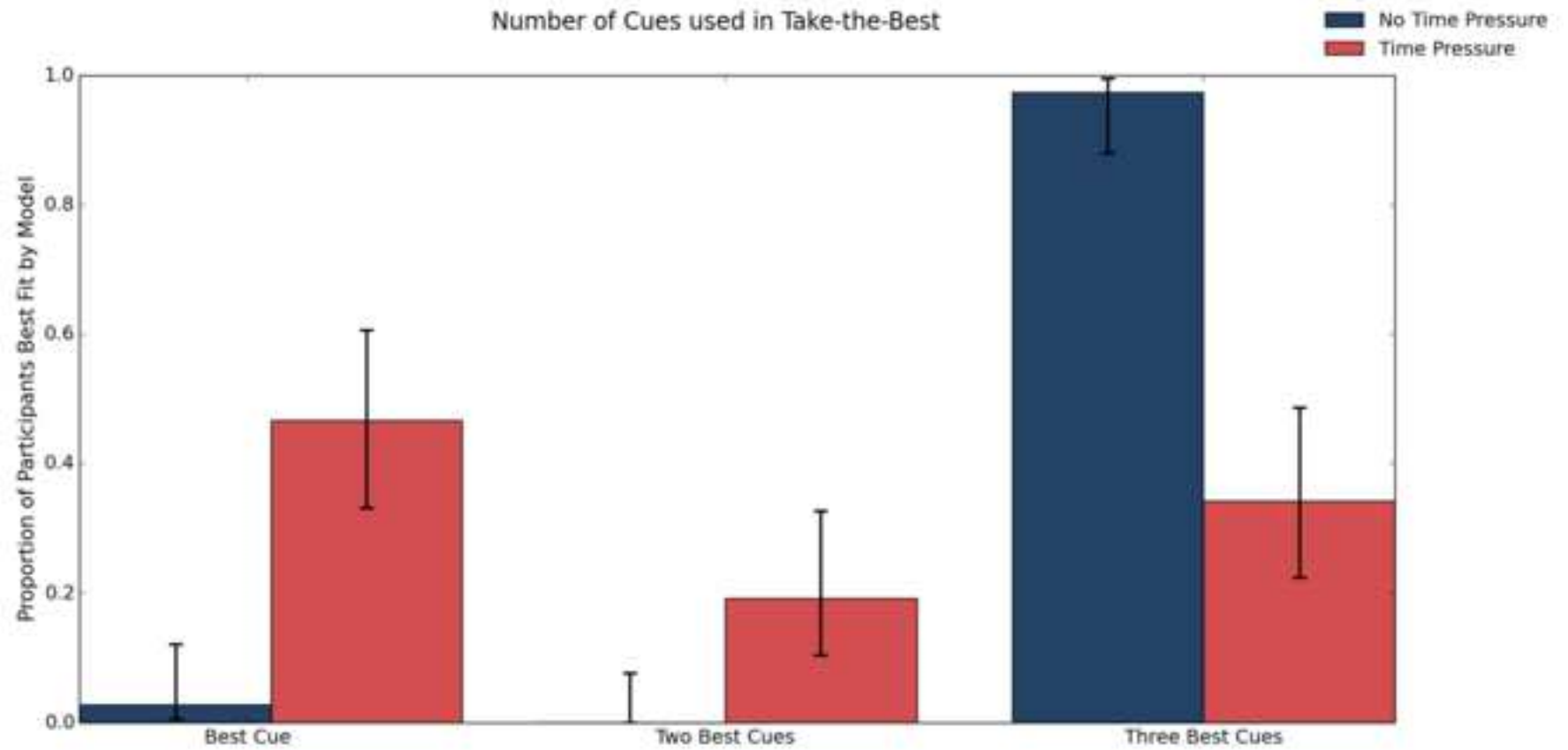
Country A	Stats	Country B
lesser	Increased employment opportunities	greater
higher	Competitiveness in medium enterprises	lower
larger	Public investment in infrastructure	smaller
stronger	Development of civic participation	weaker
inferior	Increased life expectancy for women	superior
less stable	Price stability in cheap basic goods	more stable
better	Decreased rates of infectious diseases	worse

Figure 5
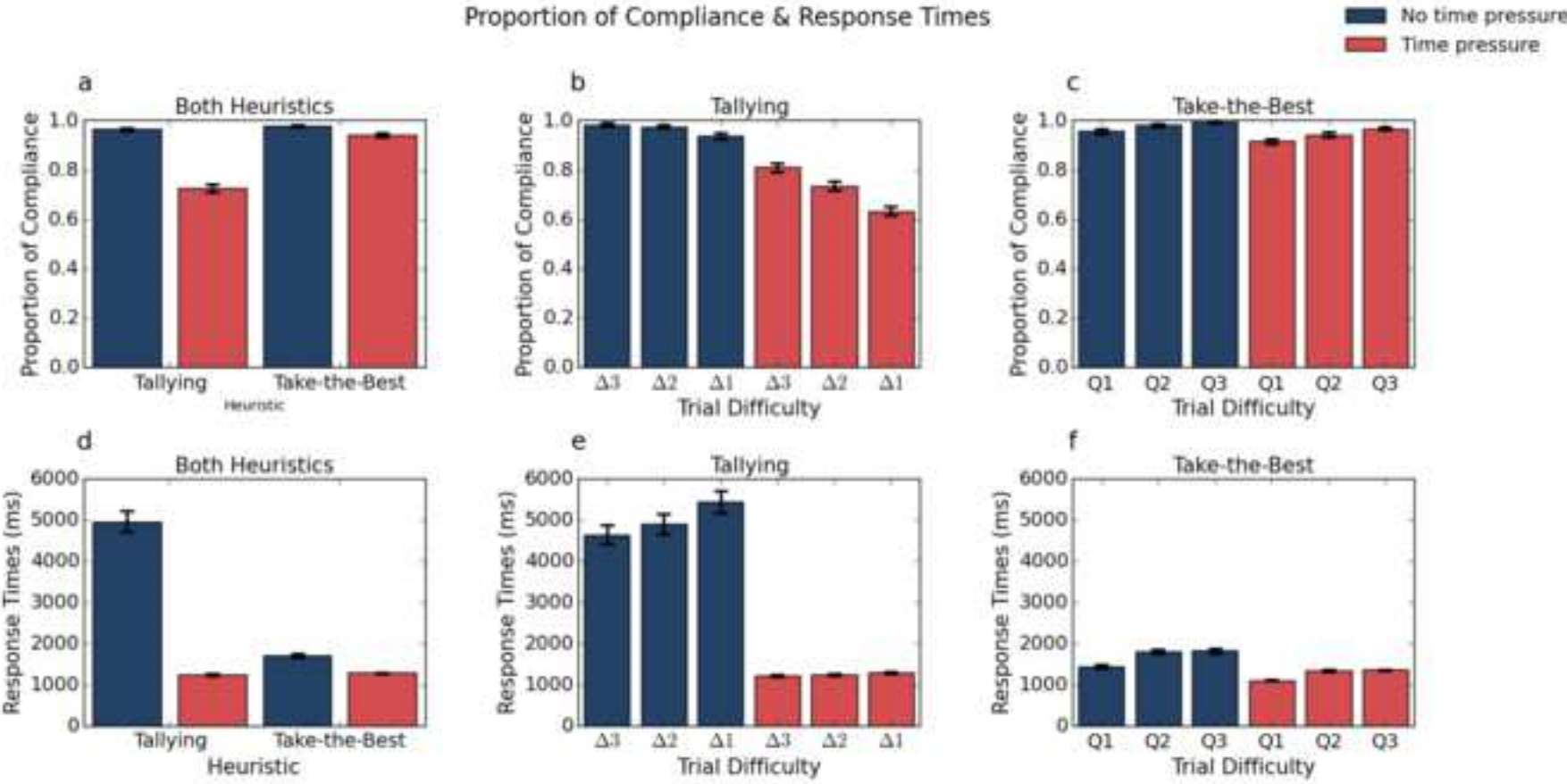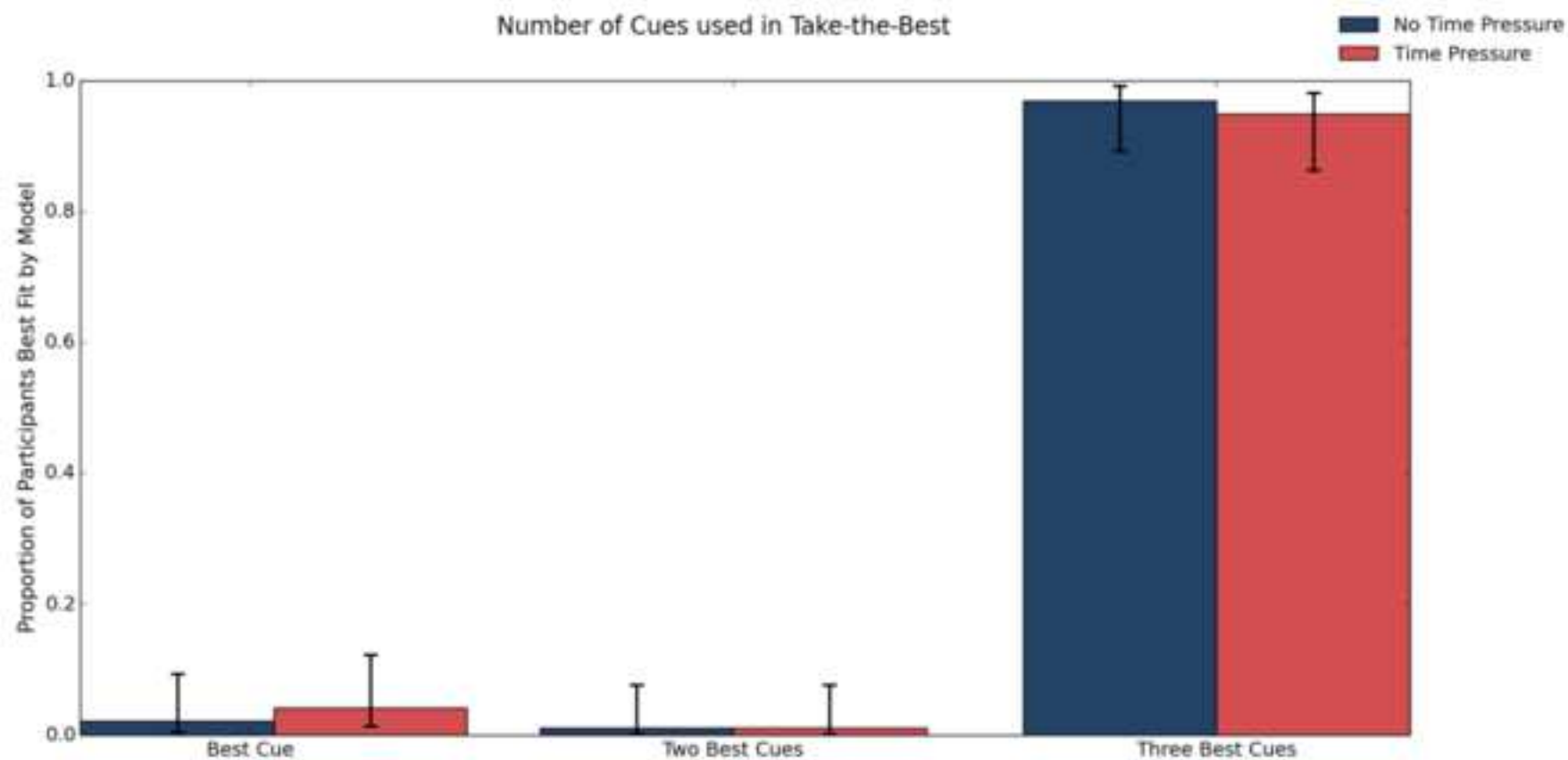
Proportion of Compliance & Response Times

Figure 6

Number of Cues used in Take-the-Best

Supplemental Material A: *List of Cues (Statistics for Developing Countries)*

"Competitiveness in medium enterprises",

"Price stability in cheap basic goods",

"Increased employment opportunities",

"Public investment in infrastructure",

"Decreased rates of infectious diseases",

"Increased life expectancy for women",

"Development of civic participation"

Supplemental Material B: *Stimuli Sampling Procedure*

Classifying trials based on Take-the-Best difficulty results in Q1, Q2, Q3, Q4, Q5 trials (i.e. trials where looking at the first cue is sufficient, looking at the second is sufficient, etc.). Because only Q1, Q2 & Q3 trials were used, this reduces the space to 468 trials.

Classifying trials based on Tallying difficulty results in trials with $\Delta 1$, $\Delta 2$, $\Delta 3$, $\Delta 4$ & $\Delta 5$ difference between the amount of negatively-valued (red cross) and positively-valued (green checkmark) cues. Since only $\Delta 1$, $\Delta 2$ & $\Delta 3$ trials were used this reduces the space to 462 trials. Cross-tabulating Tallying & Take-the-Best trial spaces results in a space of 452 trials (Table B1).

Table B1
Cross Tabulation between TAL & TTB Trial Types

		TAL		
		Δ3	*Δ2*	*Δ1*
	Q1	40	100	180
TTB	*Q2*	10	30	60
	Q3	2	8	20

In addition to balanced sampling across the two different difficulty classifications, the amount of non-discriminating cues can also affect difficulty in both experimental conditions, (whether participants use TAL or TTB). Trials range from having seven cues with discriminating power down to only three cues with discriminating power. We tried to sample trials with the most diversity possible without repetition of stimuli (Table B2).

Table B2
Cross Tabulation between TAL & TTB Trial Types Sampled

		TAL			
		Δ3	*Δ2*	*Δ1*	*Total*
	Q1	6	3	3	12
TTB	*Q2*	5	5	2	12
	Q3	1	4	7	12
	Total	12	12	12	36

This results in a total of 36 trials that can be partitioned into 3 blocks of Q1, Q2, Q3 trials or equivalently as 3 blocks of Δ3, Δ2, or Δ1 trials respectively. Now adding the trials where the signs are reversed -for right-left presentation instead of left-right presentation of the stimuli- gives a total of 72 trials.

Globally, all 72 trials can also be partitioned into trials where all cues can discriminate between options (n=8), trials where six cues can discriminate between options (n=8), trials where five cues can discriminate between options (n=28), trials where four cues can discriminate between options (n=16) and trials where three cues can discriminate between options (n=12). This can be seen in Table B3.

Table B3
Cross Tabulations of Trial Difficulty with Amount of Non-discriminating Cues for TAL & TTB

Tallying						**Take-the-Best**				
# of Equals Signs	Δ3	Δ2	Δ1	**Total**		**# of Equals Signs**	Q1	Q2	Q3	**Total**
0	6	0	2	6		0	8	0	0	6
1	0	8	0	8		1	4	4	0	8
2	18	0	10	28		2	8	12	8	28
3	0	16	0	16		3	2	6	8	16
4	0	0	12	12		4	2	2	8	12
Total	24	24	24	72		**Total**	24	24	24	72

Subset models for the time pressure phase. In addition to comparing subjects' vector of responses to what the respective heuristic predicted for each condition, we also compared these responses to what these heuristics would predict for subsets of the cues. These models were applied exclusively to the time pressure phase as shown in Figure C1. The models were classified in a two by two factorial design (heuristic model x model class), giving four model classes in total for each experimental condition. The first factor consisted of modeling subject responses for TAL subsets ($n = 128$) and for TTB subsets ($n = 13700$) for both conditions to see whether TAL subsets could better account for the data in the TTB condition or vice versa. The second factor (model class) was constructed based on the way cues were defined, either featural or positional. The featural models represent tracking cues based on their names (e.g. "Public investment in infrastructure") and the positional models represent tracking just the positions of the cues (e.g. just looking at the first three cues). A 2x2x2 ANOVA (experimental condition x heuristic model x model class) for five-fold cross-validated predictive accuracy of the models showed a main effect of model class $F(1,177) = 172.86$, $p < .001$, $\eta^2 = .49$, a main effect of heuristic $F(1,177) = 305.49$, $p < .001$, $\eta^2 = .63$, and of their interaction $F(1,177) = 326.32$, $p < .001$, $\eta^2 = 65$. There was also an effect of experimental condition $F(1,177) = 774.59$, $p < .001$, $\eta^2 = .81$, its interaction with heuristic model $F(1,177) = 233.36$, $p < .001$, $\eta^2 = .57$, and the three way interaction of model class with heuristic model and experimental condition $F(1,177) = 79.15$, $p < .001$, $\eta^2 = .31$.
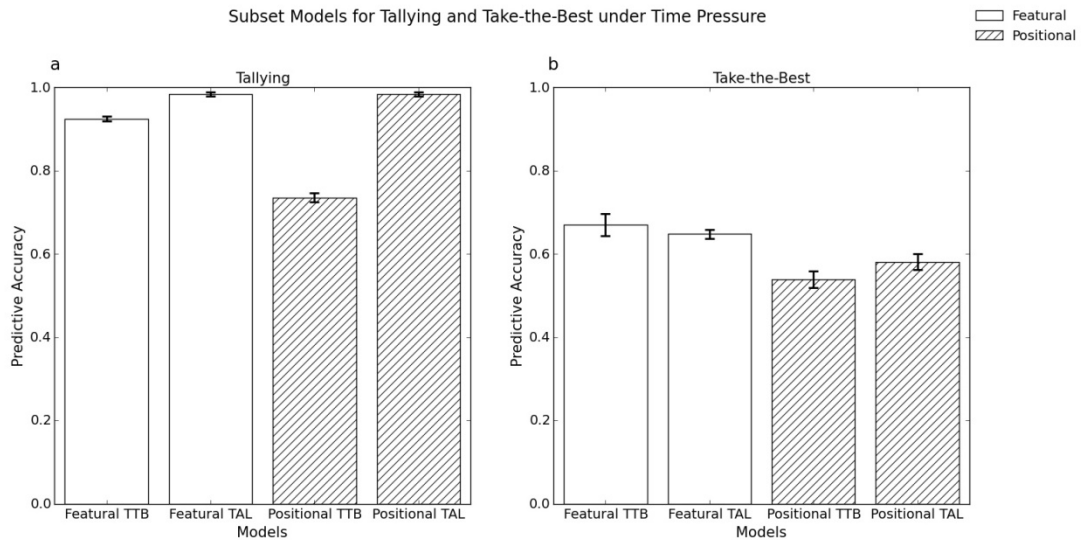

Fig. C1.

Subset models under time pressure. On the left is (a) mean predictive accuracy from five-fold cross validation for the four classes of subset models for TAL in the time pressure phase and on the right is (b) the same for TTB. Error bars are within-participant confidence intervals.

It is important to note that distinguishing between Positional TAL models and Featural TAL models was not possible due to equivalent predictions of both classes of models when subjects were using the full set of cues. We conclude this section with an exposition of the actual models considered for this analysis.

Number of TAL Subset Models:

$|TAL| = 2^n = 128$, where n is the total number of cues ($n = 7$).

Number of TTB Subset Models:
$|TTB| = \sum_{k=1}^{n=7} \binom{n}{k} k! = 13700$

Doubling the number of models to distinguish between positional and featural models, and accounting for duplicates, gives a total of 27639 subset models per subject.

Supplemental Material D: *Practice phase analysis for Experiments 1 and 2*

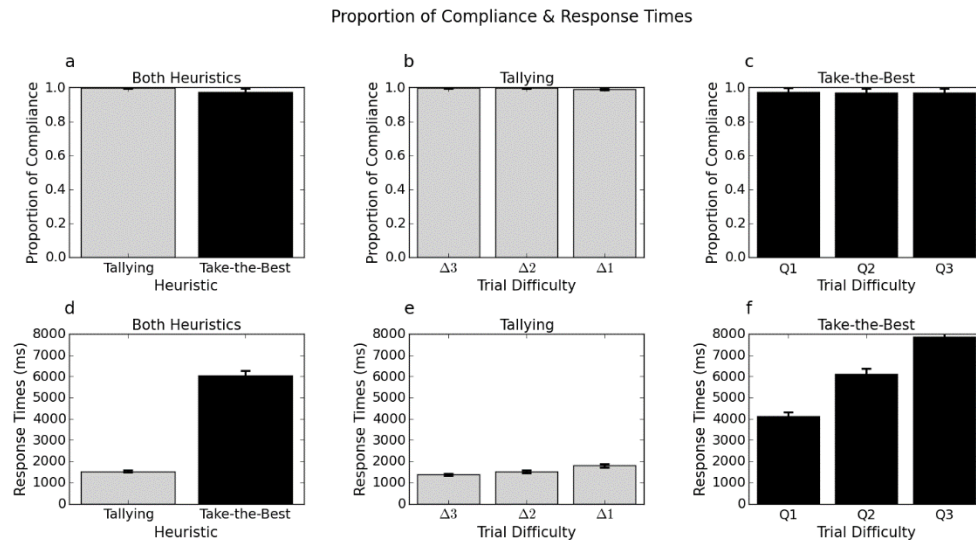Proportion of Compliance & Response Times

Fig. D1. Results for practice phase of Experiment 1 after applying exclusion criteria (n = 179). The figure shows the proportion of compliance with (a) both heuristics with TAL in gray and TTB in black, (b) the proportion of compliance in the TAL condition displayed by degrees of difficulty (Deltas), and (c) the proportion of compliance in the TTB condition displayed by degrees of difficulty (Qs). Shows the response times for (d) both heuristics, (e) the response times in the TAL condition displayed by degrees of difficulty (Deltas), and (f) the response times in the TTB condition displayed by degrees of difficulty (Qs). For all panels, error bars are 95% within-participants confidence intervals.

Experiment 1. Proportion of compliance was analyzed using a 2x3 mixed-design ANOVA with a between-subject factor of heuristic (TAL, TTB) and a within-participants factor of trial difficulty (3 levels of trial difficulty). Main effects were observed for heuristic, $F(1, 177) = 74.29$, $p < .001$, $\eta 2 = .296$. However, we did not observe effects of trial difficulty $F(2, 354) = 1.40$, $p = .248$, $\eta 2 = .008$, nor of the interaction between trial difficulty and heuristic, $F(2, 354) = .767$, $p = .465$, $\eta 2 = .004$. These results suggest that participants had a slightly harder time learning TTB although both heuristics were close to ceiling (see panels a of Figure D1). Interestingly, trial difficulty was not a relevant factor (with respect to compliance) during the practice phase (see panels b and c of Figure D1).

Response times were analyzed using a 2x3 mixed-design ANOVA with a between-subject factor of heuristic (TAL, TTB) and with a within-participants factor of trial difficulty (3 levels of trial difficulty). Main effects were observed for heuristic, $F(1, 177) = 456.69$, $p < .001$, $\eta 2 = .721$ and trial difficulty, $F(2, 354) = 398.59$, $p < .001$, $\eta 2 = .692$, as well as for the interaction of heuristic and trial difficulty, $F(2, 354) = 253.75$, $p < .001$, $\eta 2 = .589$. These results confirm the fact that TTB was harder to implement in this experiment given the slower response times (see panel d of Figure D1). Furthermore, trial difficulty did not affect response times for TAL (panel e of Figure D1) but did affect response times for TTB (panel f of figure D1).
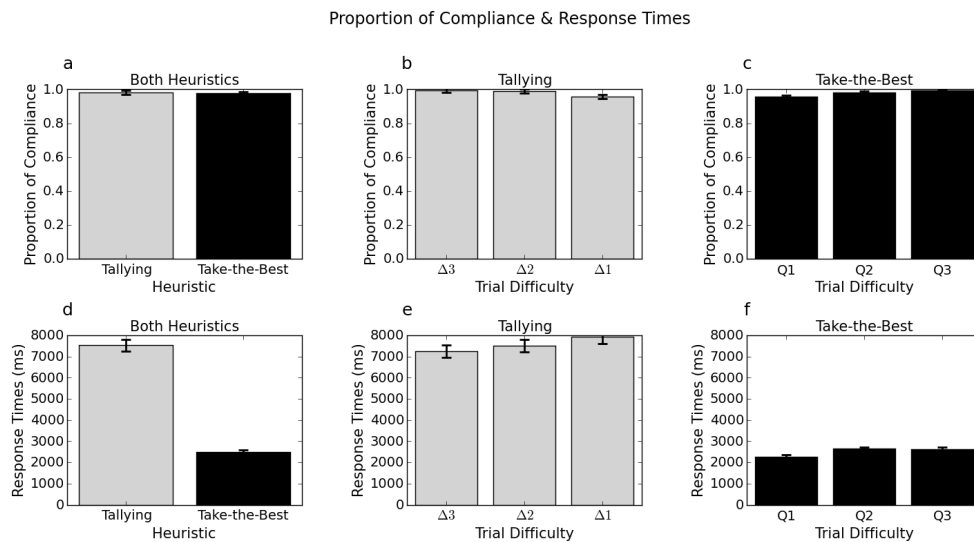

Fig. D2. Results for practice phase of Experiment 2 after applying exclusion criteria (n = 173). The figure shows the proportion of compliance with (a) both heuristics with TAL in gray and TTB in black, (b) the proportion of compliance in the TAL condition displayed by degrees of difficulty (Deltas), and (c) the proportion of compliance in the TTB condition displayed by degrees of difficulty (Qs). Shows the response times for (d) both heuristics, (e) the response times in the TAL condition displayed by degrees of difficulty (Deltas), and (f) the response times in the TTB condition displayed by degrees of difficulty (Qs). For all panels, error bars are 95% within-participants confidence intervals.

Experiment 2. Proportion of compliance was analyzed using a 2x3 mixed-design ANOVA with a between-subject factor of heuristic (TAL, TTB) and a within-participants factor of trial difficulty (3 levels of trial difficulty). Main effects were not observed for heuristic, $F(1, 171) = .295$, $p = .588$, $\eta 2 = .002$. However, we did observe effects of trial difficulty $F(2, 342) = 5.87$, $p = .003$, $\eta 2 = .033$, and of the interaction between trial difficulty and heuristic, $F(2, 342) = 60.59$, $p < .001$, $\eta 2 = .26$. These results resemble the fact that participants adequately learned the use of both heuristics and that trial difficulty affected their learning (see panels a, b and c of Figure D2). The fact that TTB shows better compliance with harder trials is unexpected but is also in line with the results seen during the test phase.

Response times were analyzed using a 2x3 mixed-design ANOVA with a between-subject factor of heuristic (TAL, TTB) and with a within-participants factor of trial difficulty (3 levels of trial difficulty). Main effects were observed for heuristic, $F(1, 171) = 309.58$, $p < .001$, $\eta 2 = .644$ and trial difficulty, $F(2, 342) = 21.33$, $p < .001$, $\eta 2 = .111$, as well as for the interaction of heuristic and trial difficulty, $F(2, 342) = 3.94$, $p = .02$, $\eta 2 = .022$. These results confirm the fact that TAL was harder to implement in this experiment given the slower response times (see panel d of Figure D2).