

Why free choices take longer than forced choices: evidence from response threshold manipulations

Christoph Naefgen¹  · Michael Dambacher² · Markus Janczyk¹

Received: 19 October 2016 / Accepted: 3 July 2017 / Published online: 3 August 2017
© Springer-Verlag GmbH Germany 2017

Abstract Response times (RTs) for free choice tasks are usually longer than those for forced choice tasks. We examined the cause for this difference in a study with intermixed free and forced choice trials, and adopted the rationale of sequential sampling frameworks to test two alternative accounts: Longer RTs in free choices are caused (1) by lower rates of information accumulation, or (2) by additional cognitive processes that delay the start of information accumulation. In three experiments, we made these accounts empirically discriminable by manipulating decision thresholds via the frequency of catch trials (Exp. 1) or via inducing time pressure (Exp. 2 and 3). Our results supported the second account, suggesting a temporal delay of information accumulation in free choice tasks, while the accumulation rate remains comparable. We propose that response choice in both tasks relies on information accumulation towards a specific goal. While in forced choice tasks, this goal is externally determined by the stimulus, in free choice tasks, it needs to be generated internally, which requires additional time.

Introduction

In 1980, the New Wave band Devo claimed that “freedom of choice is what [we] got” and that “freedom from choice is what [we] want”. Indeed, it appears that a *lack* of freedom is what we want to speed up our decisions: an increase

in choice options can slow down decisions, which has been shown in situations ranging from complex decision-making contexts (e.g., Hanoch, Wood, Barnes, Liu, & Rice, 2011) to minimalist laboratory experimental setups (e.g., Merkel, 1885). A specific example of these latter setups are comparisons between so-called forced choice and free choice tasks (Berlyne, 1957).

Forced choice and free choice tasks and their use in research

In the simplest version of forced and free choice tasks (see, e.g., Berlyne, 1957), participants have two response options (e.g., a left and a right key) and are confronted with three different stimuli (e.g., letters or color patches). Participants are instructed to respond to two of these stimuli with prescribed responses (e.g., red → left key press; blue → right key press)—the *forced choice task*. In case of the third stimulus (e.g., white), in contrast, they can choose “freely” from the two response options—the *free choice task*.¹

Notably, and of particular importance to the present study, the vast majority of studies comparing forced and free choice tasks report shorter response times (RTs) in forced choice compared with free choice tasks (e.g., Berlyne, 1957; Janczyk, Nolden, & Jolicoeur, 2015). It is the purpose of the present study to elucidate where this RT difference results from.

One interpretation of the RT difference is that both tasks differ in terms of their underlying response/action selection systems or processes. In this vein, free and forced choice

✉ Christoph Naefgen
christoph.naefgen@uni-tuebingen.de

¹ Department of Psychology, Eberhard Karls University of Tübingen, Schleichstraße 4, 72076 Tübingen, Germany

² University of Konstanz, Constance, Germany

¹ It should be noted that this freedom of choice is often constrained to some degree by instructions such as “choose both response options about equally often”.

tasks have often been used to operationalize qualitatively different *self-generated* (or intentional, internally generated, intention-based, voluntary) and *externally triggered* (or stimulus-based) actions (e.g., Brass & Haggard, 2008; Herwig, Prinz, & Waszak, 2007; Keller et al., 2006; Passingham, Bengtsson, & Lau, 2010; Waszak et al., 2005). Evidence for such a distinction comes, for example, from research on learning and using associations between bodily movements and their environmental consequences (i.e., their action effects), a field that was inspired by Ideomotor Theory (e.g., Greenwald, 1970; Harleß, 1861; Shin, Proctor, & Capaldi, 2010; Stock & Stock, 2004). In particular, when specific bodily movements are consistently followed by an auditory stimulus as an action effect (e.g., left key → low-pitch tone, right key → high-pitch tone), results from some studies suggested that associations between the movements and the effects are only learned in free choice tasks, that is, in an intention-based action control mode (Herwig et al., 2007; see also Gaschler & Nattkemper, 2012; Herwig & Waszak, 2009, 2012; Pfister, Kiesel, & Melcher, 2010).

This claim is, however, controversial. For example, Pfister, Kiesel, and Hoffmann (2011) reported learning of action effects even in forced choice tasks, and many other studies observed clear evidence for a role of action effects for performance in forced choice tasks (e.g., Janczyk, Pfister, Crognale, & Kunde, 2012; Janczyk, Pfister, Hommel, & Kunde, 2014; Janczyk, Pfister, & Kunde, 2012; Janczyk, Skirde, Weigelt, & Kunde, 2009; Kühn, Elsner, Prinz, & Brass, 2009, Exp. 3; Kunde, 2001; Kunde, Pfister, & Janczyk, 2012; Wolfensteller & Ruge, 2011). Furthermore, studies using the response-effect (R-E) compatibility paradigm (Kunde, 2001) reported R-E compatibility effects of the same size in forced and free choice tasks (e.g., Janczyk, Durst, & Ulrich, 2017), and the size of dual-task interference is also comparable for both tasks (Janczyk, Nolden & Jolicoeur, 2015). In addition, Janczyk, Dambacher, Bieleke, and Gollwitzer (2015) used the Psychological Refractory Period (PRP) paradigm in combination with the locus of slack logic (Schweickert, 1978; see also Janczyk, 2013, 2017, or Miller & Reynolds, 2003, for applications) to identify the source of the RT difference within the stream of processing. Based on Gollwitzer's (1999) implementation intention account, they argued for a perceptual locus, and indeed reported evidence in support of this idea in their study. Essentially, their observations suggest that the RT difference actually results from facilitated perceptual processing of forced choice stimuli.

In light of the evidence summarized in the last paragraph and the importance of forced and free choice tasks in contemporary research, we argue that *effect or goal state anticipation* drives response selection in both forced and free choice tasks, but that for the latter task, the effect must

be self-generated, which comes with additional demands. Here, we investigate further whether both tasks and their RT difference can be described within a common theoretical framework. This will help understanding the sources of the RT difference between the tasks.

A sequential sampling account of the RT difference

Sequential sampling models offer tools to delineate the source(s) of the RT difference between free and forced choice tasks. These approaches assume that evidence for one or the other response is (noisily) accumulated until a decision threshold is reached and the corresponding response is initiated (for an overview, see Ratcliff, Smith, Brown, & McKoon, 2016). The best-known model of this class is the drift-diffusion model, proposed by Ratcliff (1978). While many sophisticated models from this family feature high complexity, the present study focuses on a very simple model with three parameters reminiscent of the features in Grice's (1968) variable criterion model: (1) the decision thresholds that must be reached to count as a decision and to initiate emission of a response, (2) the non-accumulation time reflecting all the time before and after the accumulation time proper (i.e., early perceptual processing, motor execution, and perhaps other additional processes), and (3) the drift rate reflecting the strength of evidence for one particular response, and thus, the amount of evidence for each response added at each time-step.² With higher drift rates, for example, a threshold is on average reached earlier resulting in shorter RTs and fewer errors (which occur when the incorrect threshold is reached, e.g., due to the noise in the accumulation process). Furthermore, lowering the threshold (using a more liberal criterion) yields shorter RTs but more errors (because the chance of reaching the incorrect threshold increases), and augmenting the thresholds (using a more conservative criterion) yields longer RTs and fewer errors. Importantly, the exact kind of evidence that is accumulated is not further specified within this model. In a simple two-alternative forced choice task, one may think of an individual stimulus as the immediate cause of evidence accumulation into one or the other direction, but as already noted in the previous

² Mattler and Palmer (2012) also used a sequential sampling approach to investigate how priming affects performance and choices in both types of tasks. They observed that while masked primes always influence forced choice RTs, free choices are not influenced when the stimuli (prime and target) are of arbitrary shape. They also specified an accumulator model to explain the data, with the notable assumption of rapidly shrinking threshold separations after onset of a free choice stimulus. In their paper, they conclude that forced choice priming is a result of the integration of the automatic processing of primes and evidence from the stimulus while free choice priming is based on the integration of "external stimulation by the prime and internal response tendencies" (p.359).

section, it is also conceivable that an anticipated effect or goal state is the source of evidence being accumulated.

Assuming that effect or goal anticipation in the case of a free choice must happen endogenously (without the stimulus entirely determining the goal as in the case of forced choice trials), at least two scenarios can explain the RT difference between forced and free choice tasks within the framework described above: (1) Accumulation starts at the same time in both tasks, but the evidence driving the accumulation process towards one of the response thresholds is weaker in free choice tasks, and thus, the drift rate is lower (see Fig. 1, left panel). (2) Longer RTs in free choice tasks can also result when drift rates are the same in both tasks, but additional time is needed before (or after) the start of accumulation in free choice tasks (for a more thorough description of the consequences of different onsets of information accumulation, see Bausenhart, Rolke, Seibold, & Ulrich, 2010). In this case, the additional delay would be reflected in the non-accumulation time (see Fig. 1, right panel).

Even though on the global level, both accounts predict longer RTs in free than in forced choice tasks, there is a way to empirically distinguish them by manipulating the decision thresholds. Under the assumption of different non-accumulation times but equal drift rates, the RT difference between free and forced choice tasks should be independent of the actual threshold (see Fig. 1, right panel) and, therefore, of the same size under liberal and conservative criteria. Thus, task type and the manipulation of the decision thresholds should combine additively, because gathering the required additional information takes the same amount of time when both types of task have the same speed of information acquisition. To use a metaphor: If two horses in a horse race run at the same speed, but one horse starts 5 m closer to the

goal than the other horse, the distance between the two horses when they cross the finishing line will not change, even if the goal is moved closer to or farther away from the starting point of the race. In contrast, if there is a difference in the drift rates between the two tasks, the RT difference should become smaller the lower the threshold and bigger the higher the threshold is (see Fig. 1, left panel). In other words, task type and the manipulation should statistically interact. In the horse race metaphor, this means that one horse is faster than the other, but they start in the same position. Over the course of the race, the distance between the two horses would increase. If the race is short (liberal criterion), there is less time for the distance to increase, whereas distance can increase more in a longer race (conservative criterion), resulting in a larger difference.

Two previously established methods of manipulating decision thresholds are the amount of catch-trials in an experimental block and time pressure. Catch-trials are trials in which no stimulus appears at the time when a stimulus would normally appear. Participants are instructed not to react to this absence of a stimulus. In general, the more catch-trials there are, the longer the reaction will take (e.g., Gordon, 1967; Näätänen, 1972). It has been theorized that this is because a higher amount of catch-trials leads to a decreased stimulus expectancy, which in turn leads to a higher and thus more conservative decision threshold (e.g., Brysbaert, 1994; Grice, Nullmeyer, & Spiker, 1982; Seibold, Bausenhart, Rolke, & Ulrich, 2011). Another manipulation of the threshold is to vary the time available for responding, that is, varying the time pressure. Increasing time pressure has been repeatedly theorized and empirically shown to lower the decision criterion (e.g., Diederich, 1997; Dror, Basola, & Bussemeyer, 1999; Forstmann et al., 2008; Ratcliff & McKoon, 2008).

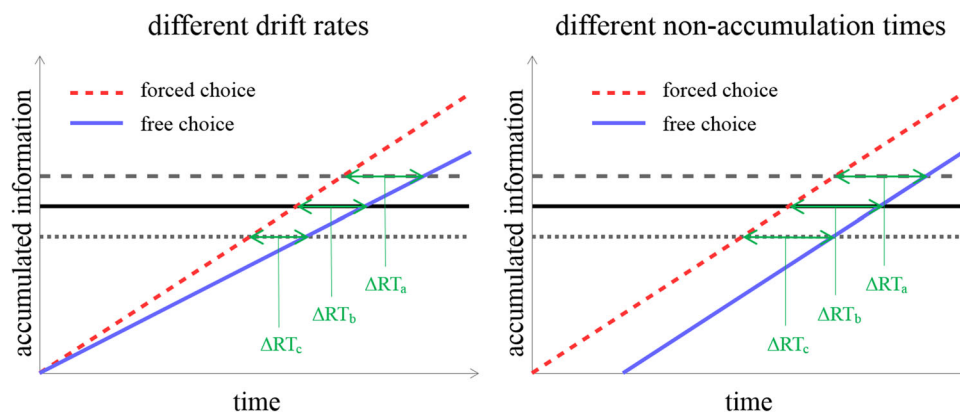
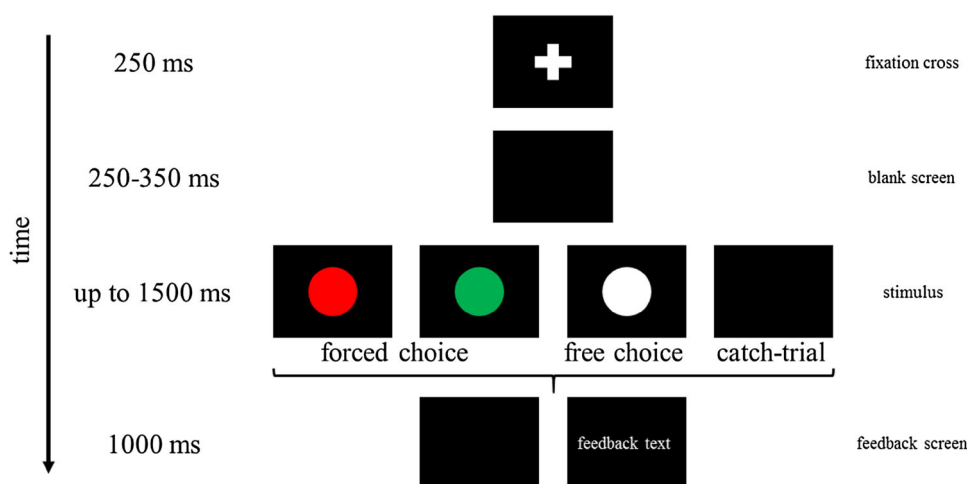


Fig. 1 Two accounts of the mean RT difference between free and forced choice tasks. The *continuous black line* represents a medium decision threshold. The *dashed grey line* indicates an increased decision threshold and the *dotted grey line* indicates a lowered threshold. Under the “drift rate” account, the RT difference between forced and free choice tasks becomes smaller with lower thresholds

($\Delta RT_a > \Delta RT_b > \Delta RT_c$). In other words, task and threshold manipulation interact with each other. In contrast, with differences in non-accumulation times, the RT difference remains the same irrespective of the threshold ($\Delta RT_a = \Delta RT_b = \Delta RT_c$) and, therefore, reflects an additive relation between task and threshold manipulation

Fig. 2 Time course of a trial in Experiment 1. No feedback text was given if no error occurred. Feedback text was displayed in German and described the type of error made (“Wrong key!”, “Too slow!”, “No stimulus was given!”)



The present experiments

The aim of the present study was to investigate the RT difference between forced and free choice tasks and to distinguish between the two accounts introduced in the previous section. In Experiment 1, we varied the amount of catch-trials to manipulate thresholds (Näätänen, 1972; Seibold et al., 2011). In Experiments 2 and 3, we manipulated the response deadline (thus inducing time pressure) to manipulate the thresholds (Ratcliff & McKoon, 2008).

Experiment 1

In Experiment 1, participants worked on forced and free choice tasks that were randomly intermingled. We expected longer RTs in the free than in the forced choice task (Berlyne, 1957). The critical manipulation was the proportion of catch-trials within a block (0%, 25%, 50%, or 75%), in which no stimulus appeared and thus no response was to be given. If task type and the catch-trial manipulation affect RTs additively, this would support the idea of comparable drift rates but longer non-accumulation times in free choice tasks. In contrast, if both interact in a way that the RT difference increases with the amount of catch-trials, this would favor an account in terms of different drift rates.

Methods

Participants

Thirty-two persons from the Tübingen area participated (mean (M) age = 24 years; standard deviation (SD) = 3 years; 27 female; one unknown value for age) for monetary compensation or course credit. All participants reported normal or corrected-to-normal vision,

were naïve regarding the underlying hypotheses, and provided written informed consent prior to data collection.

Apparatus and stimuli

Stimulus presentation and response collection were done via a standard PC connected to a 17-inch CRT monitor. Stimuli were red, green, and white circles, presented against a black background. Manual responses were collected with the two CTRL keys on a standard keyboard placed on the table in front of the participants.

Tasks and procedure

The task was either to give a predefined response to two of the possible colors (forced choice task: red and green stimuli), or to freely choose one of the two possible responses to the third color (free choice task: white stimulus). On catch-trials, where no stimulus appeared, the participants were instructed not to respond at all. Prior to each block, participants were informed about the percentage of catch-trials in this block. A trial began with the presentation of a small fixation cross (250 ms; see Fig. 2). Following a blank screen (250–350 ms), the stimulus appeared and remained on screen until the response was made. A trial was terminated if no response was given within 1500 ms after stimulus onset. General errors (i.e., no response in non-catch-trials within the time limit of 1500 ms and responses before stimulus appearance) and erroneous responses (response in a catch-trial or wrong key in forced choice trials) triggered respective feedback (1000 ms). The next trial started after an inter-trial interval (ITI) of 1000 ms. Eight blocks of 120 trials (all three stimuli appeared equally often in the normal non-catch-trials) were administered. The amount of catch-trials was varied across four block types (0%, 25%, 50%, and 75%). The

first four blocks (one of each type) were ordered by a Latin Square, and the order of the next four blocks was the reverse of the first four blocks.

Participants were tested individually in one single session of about 45 min. Written instructions emphasized speed as well as accuracy and, for the free choice trials, an even distribution of left and right responses as well as the avoidance of patterns in maintaining this distribution. The mappings of stimuli and responses in the forced choice task and the order of blocks were counterbalanced across participants. The data of participants whose free choice responses showed a strong bias towards one response option (>80% of choices) were discarded and new data were collected from new participants with the same block sequence (three participants in this experiment).

Design and analyses

The experimental manipulations resulted in two independent variables of interest, namely (1) task type (forced choice vs. free choice) and (2) block type (0% vs. 25% vs. 50% vs. 75% catch-trials). Trials with general errors were discarded. For RT analyses, only correct responses were considered (note that no erroneous responses can be made in free choice tasks). Trials with RTs deviating more than 2.5 SDs from the participants' mean per condition were excluded. Data were then submitted to a 2×4 Analysis of Variance (ANOVA) with repeated measures on task type and block type. Percentages of errors (PEs) were only analyzed for the forced choice task with an ANOVA with block type as repeated measures factor. The choice rates in the free choice task were analyzed similarly as a function of block type. p values were Greenhouse–Geisser adjusted when the assumption of sphericity was violated. In these cases, the respective ϵ is reported.

Results

Participants chose the left response button in the free choice task about 48.7% of the time in the 0%, 48.0% in the 25%, 47.2% in the 50%, and 50.6% in the 75% catch-trials blocks. These differences were not significant, $F(3,93) = 1.13$, $p = .331$, $\eta p^2 = .04$, $\epsilon = .69$.

Mean correct RTs (2.5% excluded as outliers) are shown in Fig. 3 and are summarized in Table 1. As expected, responses in the forced choice task were faster than in the free choice task, $F(1,31) = 55.50$, $p < .001$, $\eta p^2 = .64$, and responses slowed down with an increasing amount of catch-trials in a block, $F(3,93) = 102.59$, $p < .001$, $\eta p^2 = .77$, $\epsilon = .63$. This latter result suggests that the manipulation worked as intended and increased the decision thresholds. Most importantly, there was a

significant interaction between block type and task type, $F(3,93) = 3.15$, $p = .048$, $\eta p^2 = .09$, $\epsilon = .69$. A closer look at Fig. 3, however, suggests that this interaction is driven by the smaller RT difference in the 0% catch-trials blocks compared to the other blocks, and arguably, the 0% blocks differ in an important aspect from the other blocks: While in the 0% condition, participants knew that a response is always required, in the other blocks, the additional demand of distinguishing normal from catch-trials was imposed. The drift rate account, however, predicts an increasing RT difference across all levels of increasing decision thresholds (with growing differences as the amount of catch-trials increases). Therefore, we re-analyzed the data but omitted the 0% catch-trial blocks. Again the two main effects were significant as expected, task type: $F(1,31) = 47.33$, $p < .001$, $\eta p^2 = .60$, and block type: $F(2,62) = 90.47$, $p < .001$, $\eta p^2 = .74$, $\epsilon = .87$. Clearly, however, their interaction was not significant, $F(2,62) = 0.36$, $p = .697$, $\eta p^2 = .01$.

PEs in the forced choice task (i.e., wrong response keys pressed) increased with the amount of catch-trials (see Table 1), $F(3,93) = 8.41$, $p = .001$, $\eta p^2 = .21$, $\epsilon = .57$. Finally, there was a negligible amount (<0.1%) of catch-trials in which a response was given.

Discussion

The results of Experiment 1 are not in line with the drift rate account, but more compatible with differences in the non-accumulation time. Yet, they entail several aspects that complicate a straightforward interpretation.

First, RTs increased with increasing amount of catch-trials, an observation that complies with the intended manipulation of increasing thresholds (see also Näätänen, 1972; Seibold et al., 2011). At the same time, though, more errors were made in the forced choice task as well. This is unexpected, since increasing thresholds should make errors less likely.³ We will get back to this in the General Discussion.

Second, we replicated the common observation of longer RTs in the free than in the forced choice task, and task type interacted with the amount of catch-trials in the initial analyses, which included the 0% catch trials condition. Straightforwardly, this would argue against the account of equal drift rates with the differences arising from different non-accumulation times. Yet, different drift rates should result in increasing RT differences across all amounts of catch-trials, and, clearly, this was not the case. Rather, the interaction was driven by a smaller RT

³ A similar observation with PEs increasing descriptively with the amount of catch-trials can be seen in the condition with low intensity stimuli in the study by Seibold et al. (2011; see their Fig. 4).

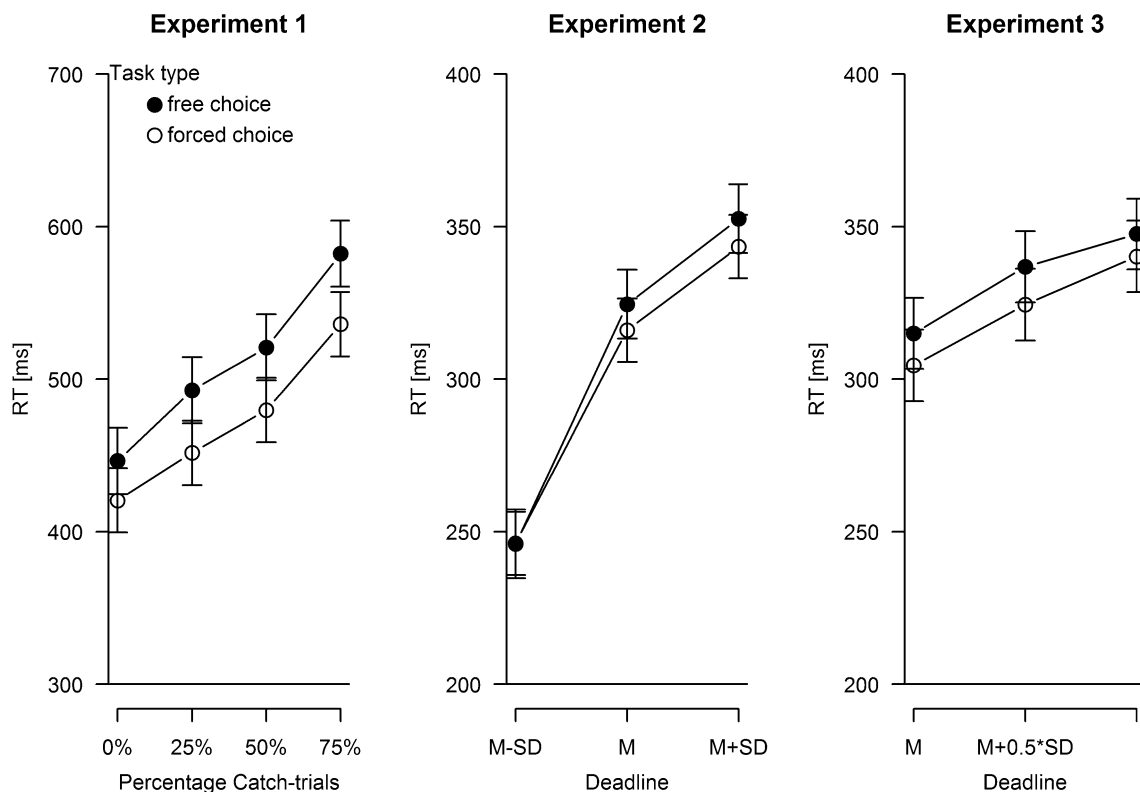


Fig. 3 Mean correct RTs in milliseconds (ms) from all three experiments as a function of task type and block type. Error bars are 95% within-subject confidence intervals calculated for the difference between free and forced choice tasks collapsed across block types (see Pfister & Janczyk, 2013)

Table 1 Ms (and SDs) of response times (RTs) in ms and percentages of errors (PEs) in forced choice tasks of Experiments 1–3 as a function of block type and trial type

	Experiment 1 Catch-trials (%)				Experiment 2 Deadlines			Experiment 3 Deadlines		
Task type	0%	25%	50%	75%	-1	0	1	0	.5	1.5
Free choice										
RT	446 (66)	493 (77)	521 (83)	582 (106)	246 (39)	325 (39)	353 (36)	315 (36)	337 (39)	348 (40)
Forced choice										
RT	421 (66)	452 (73)	480 (79)	536 (93)	246 (46)	316 (35)	343 (35)	305 (35)	324 (41)	340 (42)
PE	4.6 (3.3)	4.9 (4.2)	4.8 (5.1)	8.2 (7.6)	29.6 (10.2)	14.5 (6.1)	10.5 (5.7)	17.4 (8.2)	14.4 (7.5)	12.5 (6.5)

Deadline conditions are denoted in SD steps from *M* (i.e., $M + x \cdot SD$ with x denoting the block type in the table)

difference in the 0% catch-trial blocks, perhaps reflecting the absence of the additional demand of distinguishing normal from catch-trials in the other blocks. When considering only the comparable blocks with catch-trials, task type and block type combined additively.

Tentatively, we, therefore, take the results as support for the account of different non-accumulation times between

the tasks (see Fig. 1, right panel). At the same time, we wish to avoid premature conclusions based on this single experiment. Accordingly, in the following experiments, we sought for converging evidence and employed time pressure as a different means of manipulating response thresholds.

Experiment 2

Experiment 2 followed the same logic as Experiment 1, but time pressure was used to manipulate response thresholds. To individually adjust time limits, the mean and the standard deviation of participants' RTs in free and forced choice tasks were determined first. Subsequently, the same tasks were presented with three different levels of time pressure that were announced prior to each block and were varied block-wise.

Methods

Participants

Thirty-six persons from the Tübingen area participated (M age = 23 years; SD = 4 years; 31 female) for monetary compensation or course credit. All participants reported normal or corrected-to-normal vision, were naïve regarding the underlying hypotheses, and provided written informed consent prior to data collection.

Stimuli and procedure

Stimuli were adopted from Experiment 1. The task was largely the same, except that there were no catch-trials, and the blank screen interval between the fixation cross and the stimulus' appearance was fixed to 250 ms. At the beginning, two pre-experimental blocks with a response window of 1500 ms assessed mean RTs of each participant separately for free and forced choices. The respective M s and their SD s were then used to calculate three different response deadlines separately for the free and forced choice tasks: long ($M + SD$), medium (M), and short ($M - SD$). Then three experimental blocks, one of each deadline condition, followed. The order of these blocks and the S-R mapping within the forced choice task were fully counterbalanced. After these three blocks, another three blocks in reverse order followed. At the beginning of every block, the time limit of the task type with the shorter deadline (determined in the first two blocks, see above) was announced to the participants. After each block, participants were informed about how long their responses took on average (averaged across both free and forced choice trials). The same exclusion criterion as in Experiment 1 was used, and data from one participant were discarded and replaced by a new data set in the same condition.

Design and analyses

The experimental manipulations resulted in two independent variables of interest, namely (1) task type (forced choice vs. free choice) and (2) block type ($M - SD$ vs.

M vs. $M + SD$ response deadline). Trials with general errors were discarded. For RT analyses, only correct responses were considered (note that no erroneous responses can be made in free choice tasks), and trials with RTs deviating more than 2.5 SD s from the participants' mean per condition were excluded as outliers from analyses. Data from the experimental blocks were then submitted to a 2×3 ANOVA with repeated measures on task type and block type. Error data were only analyzed for the forced choice task by means of an ANOVA with repeated measures on block type. The choice rates in the free choice task were analyzed similarly, but included the pre-experimental blocks. p values were Greenhouse–Geisser adjusted when the assumption of sphericity was violated. In these cases, the respective ϵ is reported.

Results

Participants chose the left response button in the free choice task about 54.0% of the time in the $M + SD$ blocks, 56.1% in the M blocks, 58.3% in the $M - SD$ blocks, and 51.1% in the pre-experimental blocks, and the main effect of block type was significant, $F(3,105) = 5.65$, $p = .004$, $\eta p^2 = .14$, $\epsilon = .72$. In the pre-experimental blocks, mean RTs were 423 ms in the forced choice task and 444 ms in the free choice task, $F(1,35) = 18.16$, $p < .001$, $\eta p^2 = .34$.

Mean correct RTs (1.3% excluded as outliers) are shown in Fig. 3 (middle panel) and are summarized in Table 1. As expected, there was a main effect of block type on RTs, $F(2,70) = 363.32$, $p < .001$, $\eta p^2 = .91$, $\epsilon = .75$, with higher time pressure induced by shorter response deadline resulting in shorter RTs, as well as a main effect of task type, $F(1,35) = 5.72$, $p = .022$, $\eta p^2 = .14$, with longer RTs in the free choice task compared to the forced choice task. The interaction between block type and task type was also significant, $F(2,70) = 4.50$, $p = .021$, $\eta p^2 = .11$, $\epsilon = .81$. Inspection of the RTs revealed virtually no RT difference between both tasks in the high time-pressure ($M - SD$) block, which may point to a large proportion of fast guesses in this condition. Indeed, the PEs in this block ranged from 9.6% to 46.6%, that is, close to chance level. Thus, we performed a median split based on error rates in this condition (with mean PEs in the $M - SD$ condition of 22% and 38% for the below- and above-median groups, respectively), and ran an ANOVA that included this grouping variable. This ANOVA yielded an almost significant interaction between block type, task type, and the grouping variable, $F(2,68) = 2.98$, $p = .068$, $\eta p^2 = .08$, $\epsilon = .84$, and we continued to analyze both groups separately. As expected, for the participants with the above-median PEs, the interaction of task type and block type was significant, $F(2,34) = 6.41$, $p = .004$, $\eta p^2 = .27$. In

contrast, for the other group of participants with lower PEs—and thus a performance not as close to chance level—the interaction was far from significance, $F(2,34) = 0.23$, $p = .799$, $\eta^2 = .01$.

The PEs in the forced choice task increased with shorter response deadline, $F(2,70) = 52.16$, $p < .001$, $\eta^2 = .60$, $\varepsilon = .63$ (see Table 1).

Discussion

In this experiment, we manipulated the thresholds by inducing time pressure with a response deadline. First, and as expected, RTs were shorter the more time pressure was induced in a block, and also the PEs (in the forced choice task) increased accordingly. This pattern suggests that the time pressure manipulation worked as intended. Second, the initial analysis revealed a significant interaction of task and block type. Taking into account PEs, however, post hoc analyses indicated that this interaction likely resulted from a substantial proportion of fast guesses in the high time-pressure condition, which undermines the validity of the measured performance. When considering only the half of participants with below-median PEs, the interaction vanished, and results are compatible with our tentative proposal from Experiment 1, favoring an account in terms of comparable drift rates but different non-accumulation times. In addition, the RT difference in the other two blocks remained constant, whereas the drift rate account would predict an increase of the RT difference in the longer deadline.

To further validate our conclusion that longer RTs in free than in forced choices are due to differences in non-accumulation times rather than in drift rates, we ran Experiment 3. This experiment was essentially a repetition of Experiment 2, but with less severe time pressure to avoid the high error rates that supposedly resulted from fast guesses.

Experiment 3

Experiment 3 used the same setup as Experiment 2 except that we used response deadlines of M , $M + 0.5*SD$, and $M + 1.5*SD$ to avoid fast guesses as in the very short time limit in Experiment 2. We expected an additive combination of task type and block type in the present experiment.

Methods

Thirty-six persons from the Tübingen area participated (M age = 23 years, $SD = 4$ years; 29 female) for monetary compensation or course credit. All participants reported normal or corrected-to-normal vision, were naïve regarding

the underlying hypotheses, and provided written informed consent prior to data collection. This experiment was identical to Experiment 2 in all regards with the exception of the time limits, which were set at M , $M + 0.5*SD$, and $M + 1.5*SD$.

Results

Participants chose the left response button in the free choice task in about 55.6% of the time in the M blocks, 55.0% in the $M + 0.5*SD$ blocks, 55.6% in the $M + 1.5*SD$ blocks, and 52.3% in the pre-experimental blocks. These differences were not significant, $F(3,105) = 1.80$, $p = .164$, $\eta^2 = .05$, $\varepsilon = .82$. In the pre-experimental blocks, mean RTs were 398 ms in the forced choice condition and 417 ms in the free choice condition, $F(1,35) = 15.00$, $p < .001$, $\eta^2 = .30$.

Mean correct RTs (1.3% excluded as outliers) are shown in Fig. 3 (right panel) and are summarized in Table 1. As expected, there was a main effect of block type, $F(2,70) = 98.41$, $p < .001$, $\eta^2 = .74$, $\varepsilon = .96$, with shorter response deadlines resulting in shorter RTs, as well as a main effect of task type, $F(1,35) = 12.92$, $p = .001$, $\eta^2 = .27$, with longer RTs in free choice tasks compared with the forced choice tasks. The interaction between block type and task type was not significant, $F(2,70) = 2.07$, $p = .133$, $\eta^2 = .06$.

The PEs in the forced choice task decreased with increasing response deadlines, $F(2,70) = 11.93$, $p < .001$, $\eta^2 = .25$, $\varepsilon = .88$ (see Table 1).

Discussion

In Experiment 3, we observed no significant interaction between block type and task type and, if anything, the numerical decrease of the RT effect with longer deadlines was in a direction incompatible with the drift rate account (see Fig. 1, left panel). Rather, the results are in line with predictions of different non-accumulation times between the tasks.

General Discussion

Three experiments were run to elucidate the source of the RT difference between forced and free choice tasks. We used the sequential sampling framework to derive two hypotheses (see Fig. 1): First, the difference can arise from differences in the speed of evidence accumulation with drift rates being smaller for free choice tasks (see Fig. 1, left panel). Second, the difference can arise from differences in the non-accumulation time with a later onset (but a similar rate) of accumulation in the case of free choice

tasks (see Fig. 1, right panel). To distinguish these two accounts, we manipulated the response thresholds by varying the amount of catch-trials per block in Experiment 1 (Näätänen, 1972; Seibold et al., 2011) and by inducing time pressure via response deadlines in Experiments 2 and 3 (Ratcliff & McKoon, 2008).

Summary of results

First, in all experiments, forced choice stimuli were responded to faster than free choice stimuli. Second, the manipulations of catch-trials and response deadlines effectively changed the overall level of RTs as expected. However, evidence about the nature of interactions between these manipulations and task type was rather mixed. In Experiments 1 and 2, we observed significant interactions, which seemingly argue against non-accumulation time differences. In Experiment 1, though, this interaction was attributable to the block without any catch-trials, thus without an additional demand of distinguishing normal and catch-trials. In Experiment 2, no RT difference between forced and free choice tasks was evident in the high time-pressure condition, and RTs were only about 250 ms. We suspected a large proportion of fast guesses in this case, and indeed, only the participants with above-median PEs yielded a significant interaction. In the group with below-median PEs, and thus a performance not as close to chance level, the interaction vanished. Admittedly, excluding trials or reducing the number of participants lowers the statistical power for detecting an interaction. However, as a further aspect, the drift rate account predicts increasing RT differences with increasing thresholds, and this was not even descriptively the case. The clearest evidence against the drift rate account, however, comes from Experiment 3. This experiment was a repetition of Experiment 2 without a very high level of time-pressure. In this experiment, no interaction was observed and the results are compatible with the predictions derived from assuming differences in the non-accumulation time.

Overall, it seems that the drift rate account received little if any support from these results. In contrast, we did not observe evidence against the idea that there is a difference in non-accumulation times between free and forced choice tasks. Therefore, we suggest that the RT difference between free and forced choice tasks is at least partly caused by additional processes subsumed in the non-accumulation times of free choice tasks.

Limitations

One odd result in Experiment 1 is that PEs increased with increasing proportions of catch-trials. Because a higher PE, especially together with a longer RT, is compatible with a

lower drift rate, a possible explanation would be that the manipulation in Experiment 1 targeted the drift rates instead of decision thresholds. This would have broader implications for every argumentation that requires the assumption or concludes that the amount of catch-trials influences (only) the decision thresholds (e.g., in Brysbaert, 1994; Grice et al., 1982; Seibold et al., 2011). Should the manipulation through catch-trials target the drift rates instead of or additionally to the decision thresholds, this of course complicates the interpretation of the results of Experiment 1. The mean RTs and PEs in Experiments 2 and 3, though, were in line with our assumptions about the manipulation of decision thresholds as PEs increased, while RTs became shorter with shorter response deadlines and thus increasing time pressure. It should be noted, though, that also for time-pressure manipulations and speed–accuracy trade-off instructions, concerns have been raised that not (only) decision thresholds but also other parameters such as the drift rate change (e.g., Arnold, Bröder, & Bayen, 2015; Dambacher & Hübner, 2015; Rae, Heathcote, Donkin, Averall, & Brown, 2014; Rinkenauer, Osman, Ulrich, Müller-Gethmann, & Mattes, 2004).

To check which parameters our manipulations affected, we extracted the parameters for the forced choice trials with EZ (Wagenmakers, Van Der Maas, & Grasman, 2007) and compared the parameter estimates across the block types for each experiment (see the Appendix for a summary of these analyses). Notably, in Experiments 1 and 2, the drift rates were, indeed, affected by the manipulations (i.e., smaller drift rates for conditions which should only have higher decision thresholds in Experiment 1 and larger drift rates for conditions which should only have higher decision thresholds in Experiment 2). No such effect was observed for Experiment 3, which was also the one with the most straightforward and clear data pattern in RTs and PEs. Furthermore, in Experiment 1, there was no significant effect on the extracted response thresholds, while for Experiments 2 and 3, the threshold pattern matched our theoretical assumptions. Finally, in all three experiments, there was a significant influence on the non-accumulation times, which increased with the decision thresholds.

While the results from Experiments 1 and 2 must be interpreted with some caution, it is unclear whether the effects in drift rates are due to trade-offs in parameter estimation itself. If not, the previous studies using similar manipulations may suffer from the same limitations, which potentially has broader implications for other research fields. Importantly, the results of Experiment 3 revealed no drift rate effect.

Another potential limitation is that our conclusion is based on retaining the null-hypotheses of the critical (2×3) interaction effects. To facilitate interpretation of the results, we ran a power analysis using GPower (Faul,

Table 2 Percentage of left responses in free choice trials, both unconditional (column overall) and conditional on the previous trial being a free or forced choice task and left or right responses

	Overall (%)	Trial <i>n</i> -1 forced choice, left (%)	Trial <i>n</i> -1 forced choice, right (%)	Trial <i>n</i> -1 free choice, left (%)	Trial <i>n</i> -1 free choice, right (%)
Experiment 1	51.2	43.5	58.1	58.6	41.8
Experiment 2	45.5	39.8	48.1	57.6	44.2
Experiment 3	45.6	42.0	47.1	61.3	37.8

Erdfelder, Lang, & Buchner, 2007). To achieve a power of $1-\beta = .8$ with $\alpha = .05$ and $\rho = .3$ to detect a medium sized effect, the calculated required sample size was $n = 20$.⁴

Potential additional processes

If we accept that RT differences in free and forced choices are due to differences in the non-accumulation time and assume that there are one or more additional cognitive processes involved in free choice task performance: what is known about them? They are most likely not or only minimally influenced in their duration by manipulations of stimulus features, because in Experiment 3 of Janczyk, Dambacher, Bieleke & Gollwitzer (2015), stimulus brightness only affected forced choice RTs but not free choice RTs (both task types were intermixed in the same blocks). As there is an alternative explanation of the latter result (that participants only ruled out the presence of a forced choice stimulus instead of identifying free choice stimuli), this should be seen as a tentative conclusion. We discuss candidates for the additional processes in the following.

(1) *Memory processes triggered one or the other response (trial history bias)*: Part of the premise of free choice tasks is that participants are asked to respond roughly with the same amount with each response option and without a clear pattern, essentially asking the participants to act as (pseudo-)random number generators for the experiment. The breadth of the literature on random number generation alone suggests that this task is not trivial and can be approached in many different ways. Participants either really generate random numbers or they try to generate patterns that ‘feel’ random but are, in fact, not. Various biases in human random number generation are known (see also Heuer, Janczyk, & Kunde, 2010, for an overview). Examples are a lack of symmetrical response sequences, a lack of long runs of the same response, or a balancing of responses across short sequences (Bar-Hillel & Wagenaar, 1991). Both negative and positive recency effects (i.e., lowered and heightened chances of repetitions)

can be observed under different circumstances (Ayton & Fischer, 2004). We suggest that investigating what strategies, if any, are used to generate the pattern of decisions in free choice tasks could provide insight into the processes that are subsumed in the non-accumulation time. To shed some light on whether a free choice is affected by the immediate history of responses in the preceding trials, we ran a post hoc analysis on choice frequencies. In particular, when comparing the percentage of left responses with the percentages of the same type conditional on the previous response (left or right) and type of task (free or forced choice), there were significant differences for all three experiments (see Table 2),⁵ Experiment 1: $F(4,124) = 6.11$, $p = .005$, $\eta^2 = .16$, $\varepsilon = .451$; Experiment 2: $F(4,140) = 8.16$, $p = .002$, $\eta^2 = .19$, $\varepsilon = .404$; Experiment 3: $F(4,140) = 16.39$, $p < .001$, $\eta^2 = .32$, $\varepsilon = .438$. Interestingly, the resulting pattern of choice frequencies bears similarities to reports in the task switching literature where a response repetition benefit (in RTs) is only observed when the task repeats but not when the task switches (e.g., Kleinsorge, 1999; Rogers & Monsell, 1995). Whether or not the present result of fewer response repetitions following a switch from a forced choice to a free choice extends this effect is open to future research.

In sum, these observations point to the idea that the responses in the immediately preceding trials were considered on a current free choice trial. In other words, participants seem to use systematic strategies to decide what response to give, which takes time and adds to the RTs in free choice tasks.

(2) *Endogenous generation of stimulus/effect representations*: Free choice tasks as used in this study are usually intermixed with forced choice tasks. In the introduction, we stated that the exact basis of evidence accumulation is not fully specified in diffusion models. First, after having realized to be in a free choice trial, participants may endogenously generate a representation of one of the two forced choice stimuli and evidence is then accumulated for internal representations of these stimuli that are associated with one or the other responses. Second, according to

⁴ To correct the effect size entered into GPower, we used the method described by Rasch, Friese, Hofmann, and Naumann (2010).

⁵ We thank one of the reviewers for this suggestion.

Ideomotor Theory (e.g., Greenwald, 1970; Harleß, 1861; Shin et al., 2010), bodily movements are always addressed via an anticipation of the sensorial consequences of these movements, that is, their action effects (see also Janczyk, 2016; Janczyk et al., 2017; Kunde, 2001). Importantly, the possible action effects (depressed left/right response keys, visual and proprioceptive feedback from moving a left/right finger) are the same in forced and free choice tasks. A difference, however, is that for forced choice tasks, the stimulus determines the desired action effect, while in free choice tasks, this state must be generated again endogenously. Either way, such processes take time and would, therefore, be compatible with the results of this study.

The present data do not allow distinguishing between these two possibilities, and we do not claim that our list is exhaustive. For example, it is also possible that motor execution takes longer in free choice compared to forced choice tasks. Furthermore, these accounts are not mutually exclusive. It may well be that the choice of an effect is first driven by response history and then the action effect is endogenously generated, and thus both processes contribute to the non-accumulation time.

Modeling free choice data

The present results can be used as constraints for future formal models of free choice behavior that assess their parameters more directly. We are currently aware of only one direct application of a sequential sampling model to data from priming experiments in free and forced choice tasks (Mattler & Palmer, 2012). The most important outcome of the experiments in this study was that the response in the free choice task was biased by a (subliminal) stimulus-preceding prime. In the model, the activity of two accumulator nodes mutually inhibits the response unit of the other accumulator. This inhibition may account for potential response–response conflicts in free choices, which slow down responses (Berlyne, 1957). Furthermore, when a free choice stimulus appears, an exponential drop of the decision thresholds is assumed. While this model fits the priming data, the

threshold drop can generally be seen critically, because it is not assumed for the forced choice task, but it is assumed to start immediately after the appearance of the stimulus, implying some sort of stimulus identification. Nevertheless, we believe that the approach by Mattler and Palmer (2012) is a valuable step towards the identification of similarities and differences between forced and free choice tasks.

Conclusion

Applying a framework borrowed from diffusion models, we observed no evidence that the mean RT difference between free and forced choice tasks is attributable to a higher drift rate in forced than in free choice tasks. Our results are rather compatible with a delay of the information accumulation process in free compared to forced choice tasks. Future work should aim at identifying the nature of this delay and the concurrent processes in more detail.

Acknowledgements Work of MJ is supported by the Institutional Strategy of the University of Tübingen (Deutsche Forschungsgemeinschaft/German Research Foundation ZUK 63). Furthermore, this work was supported by Grant JA2307/1-2 from the Deutsche Forschungsgemeinschaft/German Research Council awarded to MJ. We thank Dirk Vorberg and an anonymous reviewer for very helpful and constructive comments on a previous version of this manuscript.

Appendix

In this appendix, we report the results from a diffusion model analysis on the forced choice data from Experiments 1–3. EZ (Wagenmakers et al., 2007) was used to extract parameters for every participant and relevant experimental condition (i.e., excluding the 0% catch-trial condition of Experiment 1 and excluding the pre-experimental blocks of Experiments 2 and 3) for each experiment. Tables 3, 4, 5 summarize the resulting means and standard deviations for drift rates, response thresholds, and non-accumulation times. The

Table 3 Extracted parameter means and standard deviations in parentheses per experimental condition for Experiment 1. $n = 32$

	Response thresholds	Drift rates	Non-accumulation times
Block type	$F(2,62) = 2.40, p = .109, \eta p^2 = .07, \varepsilon = .846$	$F(2,62) = 9.90, p < .001, \eta p^2 = .24$	$F(2,62) = 33.21, p < .001, \eta p^2 = .52, \varepsilon = .811$
25% catch-trials	3.45 (0.69)	0.0097 (0.0024)	294 (47)
50% catch-trials	3.71 (0.85)	0.0092 (0.0024)	300 (56)
75% catch-trials	3.47 (0.90)	0.0081 (0.0027)	352 (51)

In nine cases, edge corrections were applied, because there were no errors in one or more conditions

Table 4 Extracted parameter means and standard deviations in parentheses per experimental condition for Experiment 2. $n = 36$

	Response thresholds	Drift rates	Non-accumulation times
Block type	$F(2,70) = 123.37, p < .001, \eta p^2 = .78, \varepsilon = .755$	$F(2,70) = 62.90, p < .001, \eta p^2 = .64, \varepsilon = .660$	$F(2,70) = 146.79, p < .001, \eta p^2 = .81, \varepsilon = .760$
$M - SD$	1.67 (0.20)	0.0057 (0.0035)	172 (50)
M	1.94 (0.22)	0.0096 (0.0024)	238 (38)
$M + SD$	2.25 (0.27)	0.0100 (0.0021)	253 (30)

Table 5 Extracted parameter means and standard deviations in parentheses per experimental condition for Experiment 3. $n = 36$

	Response thresholds	Drift rates	Non-accumulation times
Block type	$F(2,70) = 73.10, p < .001, \eta p^2 = .68$	$F(2,70) = 2.20, p = .118, \eta p^2 = .06$	$F(2,70) = 24.46, p < .001, \eta p^2 = .41$
M	1.83 (0.25)	0.0091 (0.0030)	231 (32)
$M + 0.5*SD$	1.98 (0.33)	0.0097 (0.0027)	245 (33)
$M + 1.5*SD$	2.15 (0.30)	0.0095 (0.0023)	252 (35)

parameters were submitted to ANOVAs with block type as a repeated-measures factor. Greenhouse–Geisser corrected p values are reported when the sphericity assumption was violated (in this case, the respective ε is reported, as well). In case a participant made no mistakes in a given condition, the edge correction proposed by Wagenmakers et al. (2007) was performed, in which, essentially, the sum of errors is changed from zero errors to half an error.

References

- Arnold, N. R., Bröder, A., & Bayen, U. J. (2015). Empirical validation of the diffusion model for recognition memory and a comparison of parameter-estimation methods. *Psychological Research, 79*(5), 882–898. doi:10.1007/s00426-014-0608-y.
- Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition, 32*(8), 1369–1378. doi:10.3758/BF03206327.
- Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics, 12*(4), 428–454. doi:10.1016/0196-8858(91)90029-I.
- Bausenhart, K. M., Rolke, B., Seibold, V. C., & Ulrich, R. (2010). Temporal preparation influences the dynamics of information processing: Evidence for early onset of information accumulation. *Vision Research, 50*(11), 1025–1034. doi:10.1016/j.visres.2010.03.011.
- Berlyne, D. E. (1957). Conflict and choice time. *British Journal of Psychology, 48*(2), 106–118. doi:10.1111/j.2044-8295.1957.tb00606.x.
- Brass, M., & Haggard, P. (2008). The what, when, whether model of intentional action. *The Neuroscientist, 14*(4), 319–325. doi:10.1177/1073858408317417.
- Brybaert, M. (1994). Behavioral estimates of interhemispheric transmission time and the signal detection method: A reappraisal. *Perception & Psychophysics, 56*(4), 479–490. doi:10.3758/BF03206739.
- Dambacher, M., & Hübner, R. (2015). Time pressure affects the efficiency of perceptual processing in decisions under conflict. *Psychological Research, 79*(1), 83–94.
- Diederich, A. (1997). Dynamic stochastic models for decision making under time constraints. *Journal of Mathematical Psychology, 41*(3), 260–274. doi:10.1006/jmps.1997.1167.
- Dror, I. E., Basola, B., & Busemeyer, J. R. (1999). Decision making under time pressure: An independent test of sequential sampling models. *Memory & Cognition, 27*(4), 713–725. doi:10.3758/BF03211564.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
- Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences, 105*(45), 17538–17542. doi:10.1073/pnas.0805903105.
- Gaschler, R., & Nattkemper, D. (2012). Instructed task demands and utilization of action effect anticipation. *Frontiers in Cognition, 3*, 578. doi:10.3389/fpsyg.2012.00578.
- Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist, 54*(7), 493–503.
- Gordon, I. E. (1967). Stimulus probability and simple reaction time. *Nature, 215*(5103), 895–896. doi:10.1038/215895a0.
- Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review, 77*(2), 73–99. doi:10.1037/h0028689.
- Grice, R. (1968). Stimulus intensity and response evocation. *Psychological Review, 75*(5), 359–373. doi:10.1037/h0026287.
- Grice, G. R., Nullmeyer, R., & Spiker, V. A. (1982). Human reaction time: Toward a general theory. *Journal of Experimental Psychology: General, 111*(1), 135–153. doi:10.1037/0096-3445.111.1.135.
- Hanoch, Y., Wood, S., Barnes, A., Liu, P.-J., & Rice, T. (2011). Choosing the right medicare prescription drug plan: The effect of age, strategy selection, and choice set size. *Health Psychology, 30*(6), 719–727. doi:10.1037/a0023951.
- Harleß, E. (1861). Der Apparat des Willens. *Zeitschrift für Philosophie und philosophische Kritik, 38*, 50–73.
- Hervig, A., Prinz, W., & Waszak, F. (2007). Two modes of sensorimotor integration in intention-based and stimulus-based actions. *The Quarterly Journal of Experimental Psychology, 60*(11), 1540–1554. doi:10.1080/17470210601119134.

- Herwig, A., & Waszak, F. (2009). Intention and attention in ideomotor learning. *The Quarterly Journal of Experimental Psychology*, 62(2), 219–227. doi:10.1080/17470210802373290.
- Herwig, A., & Waszak, F. (2012). Action-effect bindings and ideomotor learning in intention- and stimulus-based actions. *Frontiers in Psychology*, 3, 444. doi:10.3389/fpsyg.2012.00444.
- Heuer, H., Janczyk, M., & Kunde, W. (2010). Random noun generation in younger and older adults. *The Quarterly Journal of Experimental Psychology*, 63(3), 465–478. doi:10.1080/17470210902974138.
- Janczyk, M. (2013). Level-2 perspective taking entails two processes: Evidence from PRP experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 1878–1887. doi:10.1037/a0033336.
- Janczyk, M. (2016). Die Rolle von Handlungszielen bei der Entstehung von Doppelaufgabenkosten. *Psychologische Rundschau*, 67, 237–249. doi:10.1026/0033-3042/a000324.
- Janczyk, M. (2017). A common capacity limitation for response and item selection in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/xlm0000408.
- Janczyk, M., Dambacher, M., Bieleke, M., & Gollwitzer, P. M. (2015). The benefit of no choice: goal-directed plans enhance perceptual processing. *Psychological Research*, 79(2), 206–220. doi:10.1007/s00426-014-0549-5.
- Janczyk, M., Durst, M., & Ulrich, R. (2017). Action selection by temporally distal goal states. *Psychonomic Bulletin & Review*, 24, 467–473. doi:10.3758/s13423-016-1096-4.
- Janczyk, M., Nolden, S., & Jolicoeur, P. (2015). No differences in dual-task costs between forced- and free-choice tasks. *Psychological Research*, 79(3), 463–477. doi:10.1007/s00426-014-0580-6.
- Janczyk, M., Pfister, R., Crognale, M. A., & Kunde, W. (2012). Effective rotations: Action effects determine the interplay of mental and manual rotations. *Journal of Experimental Psychology: General*, 141(3), 489–501. doi:10.1037/a0026997.
- Janczyk, M., Pfister, R., Hommel, B., & Kunde, W. (2014). Who is talking in backward crosstalk? Disentangling response- from goal-conflict in dual-task performance. *Cognition*, 132(1), 30–43. doi:10.1016/j.cognition.2014.03.001.
- Janczyk, M., Pfister, R., & Kunde, W. (2012). On the persistence of tool-based compatibility effects. *Journal of Psychology*, 220(1), 16–22. doi:10.1027/2151-2604/a000086.
- Janczyk, M., Skirde, S., Weigelt, M., & Kunde, W. (2009). Visual and tactile action effects determine bimanual coordination performance. *Human Movement Science*, 28(4), 437–449. doi:10.1016/j.humov.2009.02.006.
- Keller, P. E., Wascher, E., Prinz, W., Waszak, F., Koch, I., & Rosenbaum, D. A. (2006). Differences between intention-based and stimulus-based actions. *Journal of Psychophysiology*, 20(1), 9–20. doi:10.1027/0269-8803.20.1.9.
- Kleinsorge, T. (1999). Response repetition benefits and costs. *Acta Psychologica*, 103(3), 295–310. doi:10.1016/S0001-6918(99)00047-5.
- Kühn, S., Elsner, B., Prinz, W., & Brass, M. (2009). Busy doing nothing: Evidence for nonaction-effect binding. *Psychonomic Bulletin & Review*, 16(3), 542–549. doi:10.3758/PBR.16.3.542.
- Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance*, 27(2), 387–394. doi:10.1037/0096-1523.27.2.387.
- Kunde, W., Pfister, R., & Janczyk, M. (2012). The locus of tool-transformation costs. *Journal of Experimental Psychology: Human Perception and Performance*, 38(3), 703–714. doi:10.1037/a0026315.
- Mattler, U., & Palmer, S. (2012). Time course of free-choice priming effects explained by a simple accumulator model. *Cognition*, 123(3), 347–360. doi:10.1016/j.cognition.2012.03.002.
- Merkel, J. (1885). Die zeitlichen Verhältnisse der Willenstätigkeit. *Philosophische Studien*, 2, 73–127.
- Miller, J., & Reynolds, A. (2003). The locus of redundant-targets and nontargets effects: evidence from the psychological refractory period paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, 29(6), 1126–1142. doi:10.1037/0096-1523.29.6.1126.
- Näätänen, R. (1972). Time uncertainty and occurrence uncertainty of the stimulus in a simple reaction time task. *Acta Psychologica*, 36(6), 492–503. doi:10.1016/0001-6918(72)90029-7.
- Passingham, R. E., Bengtsson, S. L., & Lau, H. C. (2010). Medial frontal cortex: From self-generated action to reflection on one's own performance. *Trends in Cognitive Sciences*, 14(1), 16–21. doi:10.1016/j.tics.2009.11.001.
- Pfister, R., & Janczyk, M. (2013). Confidence intervals for two sample means: Calculation, interpretation, and a few simple rules. *Advances in Cognitive Psychology*, 9(2), 74–80. doi:10.2478/v10053-008-0133-x.
- Pfister, R., Kiesel, A., & Hoffmann, J. (2011). Learning at any rate: Action-effect learning for stimulus-based actions. *Psychological Research*, 75(1), 61–65. doi:10.1007/s00426-010-0288-1.
- Pfister, R., Kiesel, A., & Melcher, T. (2010). Adaptive control of ideomotor effect anticipations. *Acta Psychologica*, 135(3), 316–322. doi:10.1016/j.actpsy.2010.08.006.
- Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. (2014). The Hare and the Tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(5), 1226–1243. doi:10.1037/a0036801.
- Rasch, B., Frieze, M., Hofmann, W., & Naumann, E. (2010). *G*Power-Ergänzungen*. Retrieved from http://quantitative-methoden.de/Dateien/Auflage3/Band_II/Kapitel_7_GPower_Ergaenzungen_A3.pdf.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. doi:10.1037/0033-295X.85.2.59.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. doi:10.1162/neco.2008.12-06-420.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281. doi:10.1016/j.tics.2016.01.007.
- Rinkenauer, G., Osman, A., Ulrich, R., Müller-Gethmann, H., & Mattes, S. (2004). On the locus of speed-accuracy trade-off in reaction time: Inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General*, 133(2), 261–282. doi:10.1037/0096-3445.133.2.261.
- Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124(2), 207–231. doi:10.1037/0096-3445.124.2.207.
- Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology*, 18(2), 105–139. doi:10.1016/0022-2496(78)90059-7.
- Seibold, V. C., Bausenhardt, K. M., Rolke, B., & Ulrich, R. (2011). Does temporal preparation increase the rate of sensory information accumulation? *Acta Psychologica*, 137(1), 56–64. doi:10.1016/j.actpsy.2011.02.006.
- Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin*, 136(6), 943–974. doi:10.1037/a0020541.

- Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research*, *68*(2–3), 176–188. doi:[10.1007/s00426-003-0154-5](https://doi.org/10.1007/s00426-003-0154-5).
- Wagenmakers, E. J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*(1), 3–22. doi:[10.3758/BF03194023](https://doi.org/10.3758/BF03194023).
- Waszak, F., Wascher, E., Keller, P., Koch, I., Aschersleben, G., Rosenbaum, D. A., & Prinz, W. (2005). Intention-based and stimulus-based mechanisms in action selection. *Experimental Brain Research*, *162*(3), 346–356. doi:[10.1007/s00221-004-2183-8](https://doi.org/10.1007/s00221-004-2183-8).
- Wolfensteller, U., & Ruge, H. (2011). On the timescale of stimulus-based action–effect learning. *The Quarterly Journal of Experimental Psychology*, *64*(7), 1273–1289. doi:[10.1080/17470218.2010.546417](https://doi.org/10.1080/17470218.2010.546417).