

Predictability Reduces Event-file Retrieval

Philip Schmalbrock¹, Bernhard Hommel², Alexander Münchau³, Christian Beste⁴
& Christian Frings¹

¹Cognitive Psychology, Department of Psychology, University of Trier

²University Neuropsychology Center, Faculty of Medicine, TU Dresden, Germany

³Institute of Systems Motor Science, University of Lübeck, Germany

⁴Cognitive Neurophysiology, Department of Child and Adolescent Psychiatry, Faculty of Medicine, TU Dresden, Germany

Number of Pages: 45 – Number of Figures: 7 – Number of Tables: 0

Words Main Text: 7912 – Words in Abstract: 195

Authors Notes

The Deutsche Forschungsgemeinschaft (DFG) supported the research reported in this article (FR2133/15-1). Correspondence concerning this article should be addressed to Philip Schmalbrock, University of Trier, Universitätsring 15, DE-54296 Trier; Email: schmalbrock@uni-trier.de, Twitter: PSchmalbrock_TR

Abstract

There is growing consensus that stimulus-response bindings (event-files) play a central role in human action control. Here, we investigated how the integration and the retrieval of event-files are affected by the predictability of stimulus components of event-files. We used the distractor-response binding paradigm, in which nominally task-irrelevant distractors are repeated or alternated from a prime to a probe display. The typical outcome of these kinds of tasks is that the effects of distractor repetition and response repetition interact: Performance is worse if the distractor repeats but the response does not, or vice versa. This partial-repetition effect was reduced when the distractor was highly predictable (Experiment 1). Separate manipulations of distractor predictability in the prime and probe trial revealed that this pattern was only replicated if the probe distractors were predictable (Experiment 2b, 3), but not if prime distractors were predictable (Experiment 2a). This suggests that stimulus predictability does not affect the integration of distractor information into event-files, but the retrieval of these files when one or more of the integrated features are repeated. We take our findings to support theoretical claims that integration and retrieval of event-files might differ concerning their sensitivity to top-down factors.

Keywords: S-R binding, predictability, curiosity

Statement of Significance

Humans can learn to adapt the different processes that make up their behavior based on previous experiences. One of these processes is the binding of stimulus and response features into so-called event-files. We here show that predictability can actually alter this process by reducing the availability of these event-files in later episodes (i.e., retrieval).

Introduction

Even tasks as simple as grasping a cup of coffee require a plethora of different processes that culminate in executing the corresponding action. These comprise, amongst others, perception of a stimulus in one's environment, inferring its meaning, considering possible consequences, and making a decision about computing, executing, and monitoring the required action, all of which is accomplished remarkably rapidly and supposedly readily but, in fact, entails complex brain activity (see also Moravec, 1998, p. 15). One important ingredient of this scenario is the binding of stimulus and response features (S-R binding) enabling us to act in the world in goal-directed ways (Frings et al., 2020).

The *Theory of Event Coding* (TEC, Hommel et al., 2001) assumes that stimulus and response features are *integrated* into so-called *event-files* that interconnect stimulus (e.g., stimulus color or shape) and response features (e.g., the direction of a response or the effector; Hommel, 2004) into one combined representation (cf., Treisman & Gelade, 1980). Ample evidence shows that previously created event-files are *retrieved* if at least one ingredient is repeated in a later episode (Hommel, 1998, 2004). How this retrieval affects performance depends on the relationship between the retrieved and the present combination of stimulus and response features: while performance (sometimes) improves in trials where all features of an event-file repeat (benefits), performance is impaired if some but not all features are repeated (costs; Henson et al., 2014; Hommel, 1998). This is likely due to the feature codes contained in the event-file conflicting with the features and feature combinations of the present episode (Frings & Moeller, 2012; Geissler et al., 2021; Hommel, 2004). This pattern of costs (and benefits) is referred to as *S-R binding effects*¹.

¹ Note, that we use the term 'S-R binding effects' to reference the whole effect pattern that integration and retrieval contribute to. Other studies use the terms binding and integration interchangeably.

Recent theorizing in the *Binding and Retrieval in Action Control* framework (BRAC; Frings et al., 2020) emphasizes that integration of features and the retrieval of features should be separated (cf., Hommel, 2022). In previous studies, it was often hard to disentangle whether experimental factors that successfully modify the size of such S-R binding effects were actually affecting feature integration proper, or the retrieval of integrated features, or both of these processes. There already exist several studies that demonstrate that disentangling both processes is possible and necessary (task instructions: Hommel et al., 2014; Memelink & Hommel, 2013; Mocke et al., 2020; temporal expectation: Schmalbrock & Frings, 2021; or perceptual salience: Schmalbrock et al., 2021). The present study aimed at disentangling these two processes with respect to S-R binding effects including task-irrelevant feature codes. There is good evidence that participants do not only integrate nominally task-relevant information about stimuli and responses into event-files but also consider task-irrelevant information, like distractors in flanker-like setups (so-called distractor-response binding, DRB: Frings et al., 2007). This shows that irrelevant information is not only integrated but also retrieved. Here, we investigated whether the integration and retrieval processes involved in the S-R binding effects of irrelevant stimuli could be disentangled by manipulating stimulus predictability.

Predictability and S-R Binding

Research from adjacent fields demonstrates that *predictability* plays an important role in the processing of distractors. For example, the disrupting effect of salient distractors on search efficiency only emerges when salient trials are sufficiently rare (i.e., 20% of all trials). If salient distractors appear more often (i.e., 80% of all trials), participants can use their prediction of the next trial to prohibit attentional capture by salient distractors (Geyer et al., 2008). A study from the related flanker literature found that perfect predictability reduces the flanker effect substantially (Frings et al., 2019). Specifically, participants were presented with a target letter that could change randomly from trial to trial while irrelevant flanking letters

were also present. Target and flanker letters were either presented in the same or different colors. For one group of participants, target and flanker color changed randomly, while for a second group flanker color was blocked. That is, for each trial of one block, participants saw the target and flanker either always in the same color or always in different colors. Intriguingly, the flanker compatibility effect was reduced in the blocked condition.

Findings from the visual search literature suggest that predictability might modulate the extent of attention a stimulus receives. Authors often operationalize predictability as *statistical regularities* in their paradigms (see Theeuwes & Failing, 2020, chapter 4.3 for an overview). For example, Wang and Theeuwes (2018b) presented participants with a search display that could (irregularly) also contain a salient distractor item that usually leads to large interference effects (e.g., Theeuwes, 1992). However, if the distractor was repeatedly presented at the same location (above chance), interference effects markedly declined. The authors concluded that participants managed to reduce the processing priority of this specific location so that high-priority stimuli (i.e., a salient stimulus) in that location could interfere less. This finding and the conclusion are now supported by several publications (e.g., Failing et al., 2019; Failing & Theeuwes, 2020; Goschy et al., 2014; Wang & Theeuwes, 2018a, 2018b, 2018c). Interestingly, this bias is implemented *proactively* (Huang et al., 2021; van Moorselaar & Theeuwes, 2021; Wang et al., 2019). That is, processing of the predictable location is systematically reduced even before the actual search display is presented (or when a different display is presented).

This is in line with ecological approaches to behavior. Considerations about *curiosity* by Berlyne (1949, 1960) suggest that uncertainty (i.e., low predictability) induces a cognitive state of curiosity that drives participants to explore. Thus, unpredictable stimulus dimensions receive more attention while high predictability dimensions do not. Taken together, mechanistic and ecological approaches suggest that attending to stimuli crucially depends on the predictability of stimuli.

We hypothesized that introducing an element of predictability might alter the S-R binding effect because highly predictable stimuli might receive less attention and less attended stimuli have been demonstrated to show smaller S-R binding effects (Moeller & Frings, 2014; Singh et al., 2018). In fact, several findings from the S-R binding effect literature suggest that the strength of the S-R binding effect depends on attention (Hommel et al., 2014; Moeller & Frings, 2014; Singh et al., 2018; but see Hommel, 2005). In these studies, increased attention allocation towards a stimulus via a spatial cue led to increased S-R binding effects while reduced attention leads to reduced S-R binding effects. Given the findings from the visual attention literature, this would suggest that participants in S-R binding paradigms might proactively withdraw attentional resources from predictable distractors due to the bias they implemented based on their recent experience. This, in turn, would then lead to reduced S-R binding effects because crucial attentional resources are lacking for distractor processing.

The Present Study

S-R binding effects are typically investigated in sequential priming paradigms. Participants execute two consecutive responses to given stimuli in two separate displays. A (first) prime display and a (second) probe display, make up an experimental prime-to-probe episode (i.e., a trial). In the DRB paradigm that we used in this study, participants respond to a central target letter flanked by two identical distractors in the prime and probe display. As mentioned already, probe performance depends on the relationship between prime and probe display configuration, which is systematically varied in DRB studies: Prime and probe can both require the same or different responses (response relation), and prime and probe distractor identity can be the same or different (distractor relation). This leads to four possible combinations: *complete repetitions* with response repetition (RR) + distractor repetition (DR) trials (RRDR), *partial repetitions* with response repetition + distractor change (DC), or response change (RC) + distractor repetition trials (RRDC and RCDR, respectively), and

complete changes with response change + distractor change trials (RCDC). The standard finding is that performance in partial repetition trials is worse than in complete repetition or change trials – the S-R binding effect (Hommel, 1998). Note that since the retrieving stimulus is a distractor, we refer to this effect (also) as distractor-response binding effect (DRB, Frings et al., 2007)

We manipulated predictability by blocking the predictability of the distractors in prime and probe displays (Experiment 1) or only in the prime or probe display (Experiment 2a and 2b). For Experiment 1, participants worked through two blocks of distractor repetition trials (one for each possible distractor identity) and two blocks of distractor change trials (one for each possible distractor identity in the prime and the other in the probe). Block order was counterbalanced over all participants. This ensured that distractors were maximally predictable throughout a block. Target identity varied unpredictably from trial to trial and from prime to probe. In Experiment 2a, 2b, and 3, the identity of the distractors was blocked only in the prime (Experiment 2a) or the probe (Experiment 2b and 3), while the distractor in the not-blocked display varied unpredictably. The target varied unpredictably in prime and probe. In all experiments, half of all trials in a block were RC trials, the other RR trials – in the *high predictability* condition. We compared the performance of this high predictability group to a second *low predictability*² group that worked through the regular DRB task (intermixed presentation of the different trial types).

We expected DRB effects to be lower in the high predictability group compared to the low predictability group. In Experiment 1, we applied the predictability manipulation simultaneously to prime and probe displays, so that we could investigate the general effect of distractor predictability on the DRB effect. In Experiments 2a, 2b, and 3, we applied the

² We chose this label instead of the “no predictability” label because although participants do not know which of two distractor letters will be shown, they at least knew that the upcoming distractor will be one of the two possible distractors.

predictability manipulation exclusively to the prime (Experiment 2a) or the probe (Experiment 2b and 3). To foreshadow the results, distractor predictability diminished DRB in Experiment 1 and we could pinpoint this effect to the retrieval mechanism in Experiment 2b and 3.

Experiment 1

Method

Participants

64 students of Trier University (53 female; 55 right-handed) with a median age of 21 years (range = 18 to 38 years) participated. They received partial course credits for their 0.5h of service. This study followed the ethical standards defined by Trier University. The sample size was calculated according to previous studies investigating effects of DRB effect, which typically led to medium-sized effects ($d_z = 0.50$). Thus, we planned to run at least $N=27$ participants, leading to a power of $1 - \beta = 0.80$ (assuming an alpha = 0.05; GPower 3.1.9.2; Faul et al., 2007), in each group to definitively observe the basic S-R effect without pinpointing the size of the assumed interaction. To account for possible dropout, we recruited five additional participants spread out over the high and low predictability groups.

Design

For this experiment two within-participant factors and one between-participant factor were varied: response relation (within: response repetition vs. response change), distractor relation (within: distractor repetition vs. distractor change), and predictability (between: high vs. low).

Apparatus & Stimuli

The experiment was programmed using Psychopy (Peirce et al., 2019; Version 03.01.2020) and run online via Pavlovia (Peirce & MacAskill, 2018). Two distinct displays were presented (see **Figure 1**): a prime and probe display. In both displays, a string of three letters was presented at the screen center (Font: Arial; Font size 25 pixel). The central letter was

presented in green (RGB₂₅₅: 0, 128, 0) and was thus marked as the target. Target letters were either a ‘J’ or an ‘F’ (both presented equally often as prime and probe target). The two letters flanking the target letter were presented in blue (RGB₂₅₅: 0, 0, 128) and thus marked as distractor letters. Distractor letters were either a ‘D’ or a ‘K’ (both presented equally often as prime and probe distractors, but see Procedure). All stimuli were presented on a black background (RGB₂₅₅: 0, 0, 0).

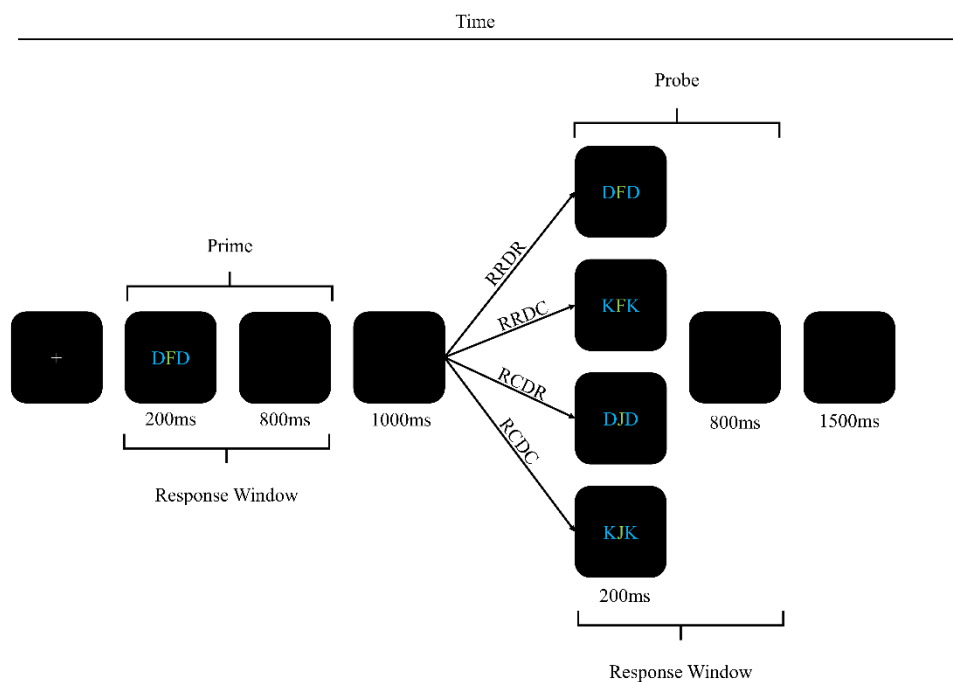


Figure 1. Procedure of the DRB task. Target letters were F or J; Distractor letters were D or K. Participants identified the central, green target letter. Probe display configuration was dependent on trial type and prime configuration. In response repetition trials (RR), the prime target was repeated as the probe target. In response change trials (RC), the probe target was different from the prime target. In distractor repetition trials (DR), the prime distractor was repeated as the probe distractor. In distractor change trials (DC), the probe distractor was different from the prime distractor. Note that in the low predictability conditions of Experiment 1, and 2a and 2b, trial type was determined randomly. In the high predictability condition, trial type was blocked (Experiment 1) or prime / probe distractor identity was blocked (Experiment 2a and 2b). Stimuli are not drawn to scale.

Procedure

Participants registered for the study and received the study link from the recruitment platform used at Trier University (Sona Systems; sona-systems.com). Participants were instructed to place their left index finger on the 'F' key and their right index finger on the 'J' key. It was emphasized that responses were to be made as fast as possible while maintaining high accuracy. A training consisting of 20 trials was completed before the main experiment – here, participants received performance feedback after both prime and probe training displays (For a correct response: “Correct!”; For a wrong response: “Wrong!”; both in German language). After the training finished, participants only received feedback when they made an erroneous response after the display the error occurred in (“Wrong! Please respond as fast and correct as possible”; translated from German).

The task consisted of two consecutive responses. In both, prime and probe displays, participants had to classify the identity of the centrally presented, green target letter. If the letter was an F, they responded with the 'F'-key; if the letter was a 'J', they responded with a 'J'-key press.

The experimental block consisted of 400 trials with a break after each 25-trials block (1x 400 trials in the low predictability group; 4x 100 trials in the high predictability group). A single trial consisted of the following chain of events: A trial began with a fixation mark (+) presented at the screen center for 1000 ms. The fixation display was followed by the prime display. The prime display was presented for 200 ms with a response window of 1000 ms, beginning with display onset. If a response was registered, the prime ended before 1000ms had elapsed. After the prime ended, a blank screen was presented for 1000 ms. Then, the probe display was presented. The probe display was, again, presented for 200 ms with a response window of up to 1000 ms. If a response was registered, the probe ended before 1000 ms had elapsed. Each trial was separated from the next by a blank screen for 1500 ms.

The two factors response relation and distractor relation were varied orthogonally. In response repetition trials, the same response required in the prime was also required in the probe. Vice versa, in response change trials a different response was required in prime and probe. In distractor repetition trials, the prime distractor was again presented in the probe. In distractor change trials, the prime distractor was different from the probe distractor.

The order in which the trials were presented differed between the high predictability and the low predictability condition. In the low predictability condition, all possible trial configurations derived from response relation, distractor relation, target, and distractor identity were presented randomly intermixed. In the high predictability condition, four separate blocks were presented, one for each possible trial configuration of distractor relation and distractor identity. Two blocks consisted of only distractor change trials and two blocks consisted only of distractor repetition trials. There were two blocks for each factor level of distractor relation and both possible distractor identities; leading to these four block types: Distractor repetition with only 'D' as distractor; Distractor repetition with only 'K' as distractor; Distractor change with 'D' as a prime distractor and 'K' as probe distractor; Distractor change with 'K' as a prime distractor and 'D' as probe distractor. Each block consisted of 50 response repetition trials and 50 response change trials. The order of blocks was further varied (see the **Supplementary Material** for an analysis of block order; in all three experiments block order did not affect any relevant effect)³. Two groups started with the two distractor repetition blocks and two groups started with the distractor change blocks. Each of the two groups started either with 'D' as the first distractor letter or with 'K' as the first distractor letter.

³ *Block Order A*: 1) Distractor repetition D, 2) Distractor repetition K, 3) Distractor change D, 4) Distractor change K; *Block Order B*: 1) Distractor repetition K, 2) Distractor repetition D, 3) Distractor change K, 4) Distractor change D; *Block Order C*: 1) Distractor change K, 2) Distractor change D, 3) Distractor repetition D, 4) Distractor repetition K; *Block Order D*: 1) Distractor change D, 2) Distractor change K, 3) Distractor repetition K, 4) Distractor repetition D

Results

Data processing and analysis were done with R (R Core Team, 2019; R version 3.6.1). The package ‘dplyr’ (Wickham et al., 2019) was used for data processing and aggregation. The DRB effect was computed as the distractor repetition benefit in response repetition trials minus the distractor repetition cost in response change trials ([RRDC-RRDR]-[RCDC-RCDR]). Binding effects in the two different conditions were compared using a non-parametric Wilcoxon-Signed-Rank Test (Wilcoxon, 1992) supplemented by Bayesian *t*-tests (Rouder et al., 2009), whose Bayes factors (BF_{01}) quantify the evidence in favor of the null hypothesis relative to the evidence in favor of the alternative hypothesis. Values between 1 and 3 indicate weak / anecdotal evidence in favor of the null hypothesis, and values >3 indicate positive / strong evidence for the null hypothesis. Vice versa, values from 1 to .33 indicated weak / anecdotal evidence for the alternative hypothesis, and values $< .33$ indicate positive / strong evidence for the alternative hypothesis (Jarosz & Wiley, 2014). Bayes factors were computed using the package ‘BayesFactor’ (Morey & Rouder, 2018).

Data Processing

For the analysis of probe performance, only probe trials with correct answers in the prime were considered for error and reaction time (RT) analysis. Additionally, only trials with RTs longer than 200 ms and shorter than 1.5 interquartile ranges over the third quartile of each person’s RT distribution were analyzed (see Tukey, 1977). That is, 7 % of all trials were excluded due to these constraints. For the analysis of probe RT, all trials with wrong probe responses were also excluded, that is an additional 6 % of all trials (13 % in total). See the **Supplementary Material** for the associated ANOVAs and **Appendix A** for a plot of mean RTs and error rates.

Reaction Times

A Wilcoxon-Signed Rank Test suggest that the DRB effects in the high predictability condition ($M = 4$ ms, $SD = 16$) were significantly lower than the DRB effects in the low predictability condition ($M = 14$ ms, $SD = 22$), $W = 698$, $p = .006$, $r = 0.31$, $BF_{01} = 0.36$ (one-tailed; see **Figure 2**). A post-hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (one-tailed; $V = 451$, $p < .001$, $r = 0.46$, $BF_{01} = 0.02$) but not for the high predictability condition (one-tailed; $V = 335$, $p = .095$, $r = 0.16$, $BF_{01} = 1.09$).

Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 6% of all trials were relevant for this analysis). A Wilcoxon-Signed Rank Test suggest that the DRB effects in the high predictability condition ($M = 2.52$ %, $SD = 4.84$) were not significantly lower than the DRB effects in the low predictability condition ($M = 5.10$ %, $SD = 8.56$), $W = 584$, $p = .170$, $BF_{01} = 0.85$ (one-tailed). The post-hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (one-tailed, $V = 6798$, $p < .001$, $r = 0.40$, $BF_{01} = 0.03$), and for the high predictability condition (one-tailed, $V = 6448$, $p < .001$, $r = 0.35$, $BF_{01} = 0.07$).

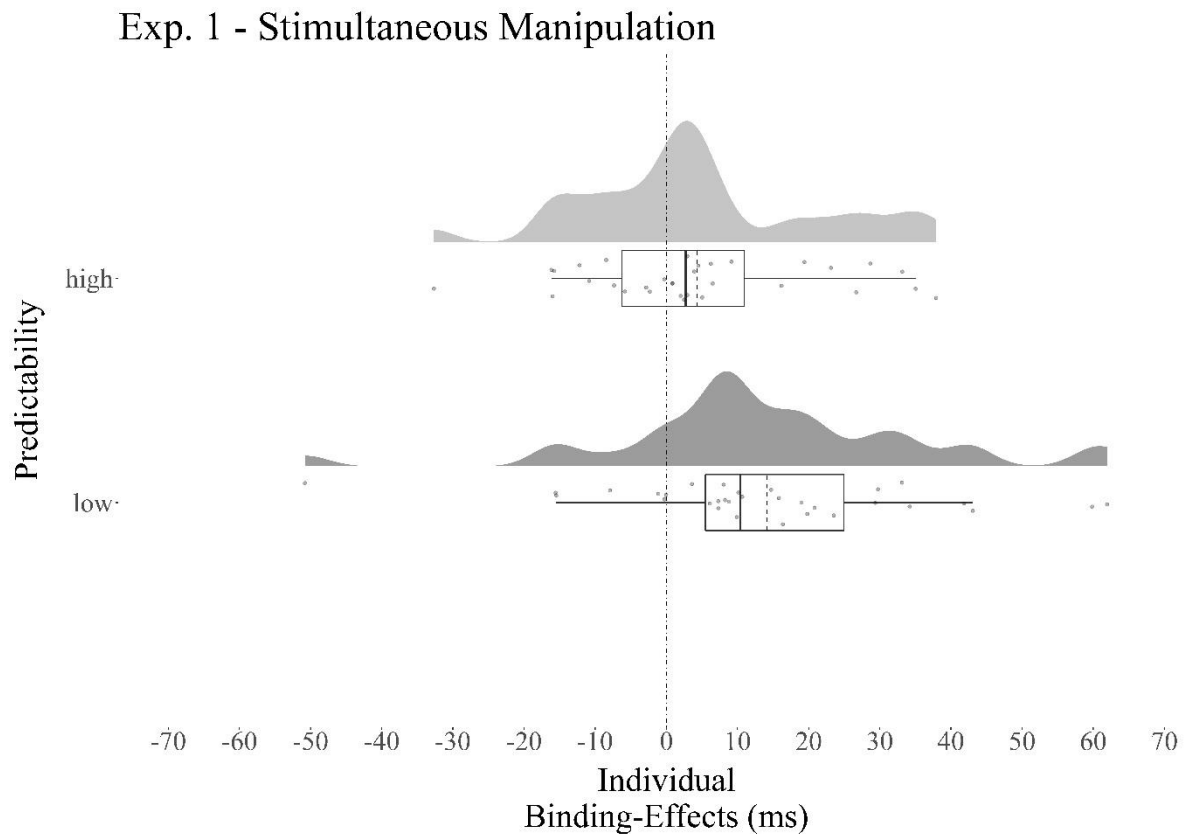


Figure 2. Individual Binding Effects for Experiment 1. Raincloud plots (Allen et al., 2019) for individual DRB effects in the high and low predictability condition for Experiment 1. The solid vertical line in each boxplot represents the median of the distribution; the dashed line represents the mean of the distribution. Upper and lower whiskers extend to the largest / smallest value above / below the respective hinge but at most 1.5 times the interquartile range above and below the third and first quartiles (McGill et al., 1978). Note that participants show lower binding effects when distractor stimuli are predictable.

Discussion

In Experiment 1, we found a modulating effect of predictability on the DRB effect: DRB effects were stronger in the low predictability group compared to the high predictability group. Since the manipulation in Experiment 1 targeted both prime and probe (i.e., integration *and* retrieval), it is not possible to pinpoint whether our observation reflects an impact on integration, retrieval, or both. The BRAC framework emphasizes that a truly mechanistic understanding of the role of feature binding in action control requires the separation of effects on integration processes proper on the one hand and effects of the retrieval of event-files on the

other. To meet this requirement, we thus moved on to disentangle integration effects from retrieval effects of predictability. In Experiment 2a and 2b, we thus attempted to target integration (Experiment 2a) and retrieval (Experiment 2b) separately. We achieved this by presenting two blocks with fixed distractor identities in one display but unpredictable distractor identities in the other. In Experiment 2a, the prime distractor letters never changed throughout a block while the probe distractor letters changed unpredictably. In Experiment 2b, the probe distractor letters never changed throughout a block while the prime distractor letters changed unpredictably. We presented two blocks, one for each distractor identity used. Block order was counterbalanced over all participants.

Since we found a modulating influence of predictability on the DRB effect, we further expected to find a modulating influence of predictability on DRB effects in at least one of the two experiments. Specifically, DRB effects should be smaller in the high predictability condition compared to the low predictability condition.

Experiment 2a and 2b

Method

Participants

64 students participated in each experiment. No participant was excluded from Experiment 2a (47 female; 64 right-handed), their median age was 22 years (range: 19 to 37). Two participants were excluded from Experiment 2b because they did not comply with the instructions (they gave no response in any trial) bringing the final sample down to 62 participants (42 female; 57 right-handed); their median age was 22 years (range: 19 to 36). They received partial course credits for their 0.5h of service. This study followed the ethical standards defined by Trier University. The sample size for both experiments was calculated as in Experiment 1.

Design

For both experiments, two within-participant variables were varied: Response relation (repetition vs. change) and distractor relation (repetition vs. change). Additionally, one between-participant variable was varied: predictability of prime or probe distractors (high vs. low).

Apparatus & Stimuli

Apparatus and stimuli were identical to Experiment 1.

Procedure

The procedure was identical to Experiment 1 except for the following changes: In the high predictability condition of Experiment 2a, two blocks were presented (one for each distractor identity, each 200 trials; in the low predictability condition one block of 400 trials was presented). In each block, the same prime distractor letter was shown throughout the block while the probe distractor letter changed unpredictably. In the high predictability condition of Experiment 2b, two blocks were presented (one for each distractor identity). In each block, the same probe distractor letter was shown throughout the block while the prime distractor letter changed unpredictably. Half of the participants started with a block where the letter D was repeated, the other half started with the letter K. The second block was always the other, so far unused letter (see the **Supplementary Material** for an analysis of block order)⁴.

Results

Experiment 2a

Data Processing

For the analysis of probe performance, only probe trials with correct answers in the prime were considered for error and RT analysis. Additionally, only trials with RTs longer than 200 ms and shorter than 1.5 interquartile ranges over the third quartile of each person's RT

⁴ Block Order A: 1) D 2) K; Block Order B: 1) K, 2) D

distribution were analyzed (see Tukey, 1977). That is, 11 % of all trials were excluded due to these constraints. For the analysis of probe RT, all trials with wrong probe responses were also excluded, that is an additional 7 % of all trials (18 % in total). See the **Supplementary Material** for the associated ANOVAs and **Appendix B** for a plot of mean RTs and error rates.

Probe Reaction Times

A Wilcoxon-Signed Rank Test suggested that the DRB effects in the high predictability condition ($M = 17$ ms, $SD = 22$) were not significantly lower than the DRB effects in the low predictability condition ($M = 9$ ms, $SD = 23$), $W = 430$, $p = .138$, $r = 0.14$, $BF_{01} = 8.93$ (one-tailed; see **Figure 3**, upper panel). The post-hoc analysis showed that the DRB effects were significantly different from zero for the low predictability condition (one-tailed, $V = 372$, $p = .022$, $r = 0.25$, $BF_{01} = 0.32$), and for the high predictability condition (one-tailed, $V = 467$, $p < .001$, $r = 0.51$, $BF_{01} < 0.01$).

Probe Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 7 % of all trials were relevant for this analysis).

A Wilcoxon-Signed Rank Test suggested that the DRB effects in the high predictability condition ($M = 6$ %, $SD = 8$) were not significantly lower than the DRB effects in the low predictability condition ($M = 3$ %, $SD = 8$), $W = 446$, $p = .191$, $BF_{01} = 7.99$ (one-tailed). The post-hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (one-tailed, $V = 369$, $p = .007$, $r = 0.31$, $BF_{01} = 0.25$), and for the high predictability condition (one-tailed, $V = 451$, $p < .001$, $r = 0.46$, $BF_{01} = 0.01$).

Results

Experiment 2b

Data Processing

For the analysis of probe performance, only probe trials with correct answers in the prime were considered for error and RT analysis. Additionally, only trials with RTs longer than 200 ms and shorter than 1.5 interquartile ranges over the third quartile of each person's RT distribution were analyzed (see Tukey, 1977). That is, 12 % of all trials were excluded due to these constraints. For the analysis of probe RT, all trials with wrong probe responses were also excluded, that is an additional 7 % of all trials (19 % in total). See the **Supplementary Material** for the associated ANOVAs and **Appendix B** for a plot of mean RTs and error rates.

Probe Reaction Times

A Wilcoxon-Signed Rank Test suggested that the DRB effects in the high predictability condition ($M = 2$ ms, $SD = 19$) were significantly lower than the DRB effects in the low predictability condition ($M = 9$ ms, $SD = 20$), $W = 600$, $p = .047$, $r = 0.21$, $BF_{01} = 0.96$ (one-tailed; see **Figure 3**, lower panel). Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (one-tailed, $V = 357$, $p = .032$, $r = 0.27$, $BF_{01} = 0.17$), but not for the high predictability condition (one-tailed, $V = 285$, $p = .480$, $r = 0.09$, $BF_{01} = 2.90$).

Probe Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 7 % of all trials were relevant for this analysis).

A Wilcoxon-Signed Rank Test suggested that DRB effects in the high predictability condition ($M = 3$ %, $SD = 7$) were not significantly lower than DRB effects in the low

predictability condition ($M = 5\%$, $SD = 8$), $W = 539$, $p = .209$, $r = 0.10$, $BF_{01} = 1.63$ (one-tailed). The post-hoc analysis evidenced that the DRB effect were significantly different from zero for the low predictability condition (one-tailed, $V = 422$, $p < .001$, $r = 0.45$, $BF_{01} = 0.05$), and for the high predictability condition (one-tailed, $V = 346$, $p = .028$, $r = 0.24$, $BF_{01} = 0.40$).

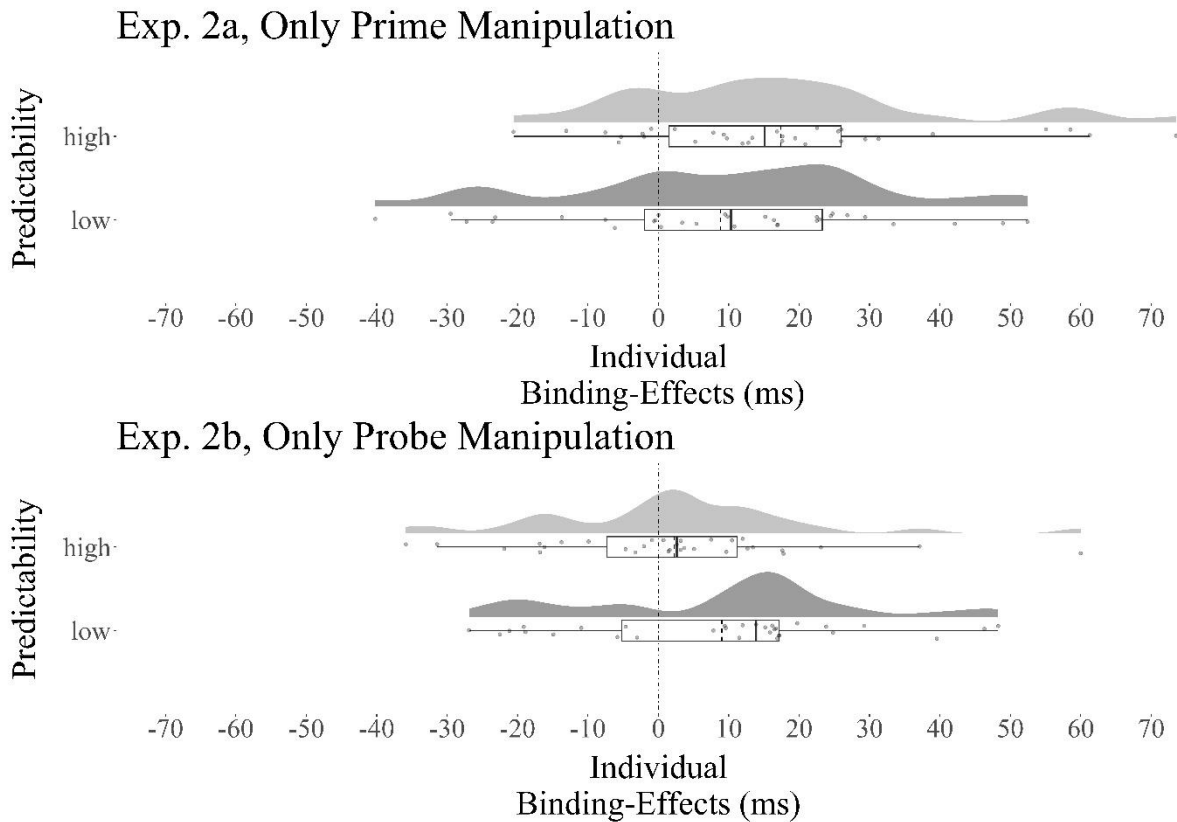


Figure 3. Individual Binding Effects for Experiment 2a and 2b. Raincloud plots (Allen et al., 2019) for individual binding effects in the high and low predictability condition for Experiment 2a (upper panel) and Experiment 2b (lower panel). The solid vertical line in each boxplot represents the median of the distribution; the dashed line represents the mean of the distribution. Upper and lower whiskers extend to the largest / smallest value above / below the respective hinge but at most 1.5 times the interquartile range above and below the third and first quartile (McGill et al., 1978). Note that in Experiment 2b, participants did not show DRB effects when distractors were highly predictable.

Discussion

In Experiment 2a and 2b, our goal was to further pinpoint the modulating effect predictability has on the DRB effect. That is, whether predictability affects integration,

retrieval, or both. The fact that our predictability manipulation affected the size of the DRB effect only when applied to distractors in the probe trial, but not when applied to distractors in the prime trial, strongly suggests that the effect of distractor predictability does not reflect integration (Experiment 2a) but only retrieval (Experiment 2b).

Note that the comparison between high and low predictability conditions in Experiment 2b was only significant on the one-tailed level. On the one hand, using one-tailed testing can clearly be justified based on both the theoretical considerations motivating the present study and the empirical findings obtained in Experiment 1. On the other hand, however, one may ask why the obtained effect was not stronger. To some degree, one could have expected that restricting our manipulation to only one of the two displays would weaken our effects due to encoding specificity (Laub & Frings, 2020). That is, increasing dissimilarity between prime and probe displays should indeed reduce DRB effects and in fact any retrieval-based effect. Thus, by increasing encoding specificity we introduced more noise into the data, which is unfortunate but unavoidable for reaching our theoretical goals. **To reduce the noise in our data at least somewhat, we aimed to replicate the findings of Experiment 2b in Experiment 3. Since the added noise of our manipulation worked on top of the already noisy online environment, we ran our replication in an offline setting.**

Experiment 3

Participants

64 students were recruited for this experiment. Two participants were lost due to a technical error, bringing the final sample down to 62 participants (47 female, 15 male; 56 right-handed, 6 left-handed). They received partial course credits for their 0.5h of service. This study followed the ethical standards defined by Trier University. The sample size was calculated as in Experiment 1.

Design

The experiment's design was identical to Experiment 2b.

Apparatus & Stimuli

Apparatus and Stimuli were identical to Experiment 2b with the important difference that Experiment 3 was run in an offline laboratory setting. That is, participants were individually tested in sound-proof chambers and were seated approximately 50cm away from a 22-inch display monitor (60 Hz refresh rate, 1680 x 1050 pixel resolution). To ensure temporal stability of luminance and color, the monitor was warmed up 5 min before the experiment begin (Poth & Horstmann, 2017).

Procedure

The procedure was identical to Experiment 2b (see the **Supplementary Material** for an analysis of block order).

Results

Data Processing

For the analysis of probe performance, only probe trials with correct answers in the prime were considered for error and RT analysis. Additionally, only trials with RTs longer than 200 ms and shorter than 1.5 interquartile ranges over the third quartile of each person's RT distribution were analyzed (see Tukey, 1977). That is, 4 % of all trials were excluded due to these constraints. For the analysis of probe RT, all trials with wrong probe responses were also excluded, that is an additional 4 % of all trials (8 % in total). See the **Supplementary Material** for the associated ANOVAs and **Appendix C** for a plot of mean RTs and error rates. The analysis followed the preregistered analysis plan.

Probe Reaction Times

A Wilcoxon-Signed Rank Test suggested that the DRB effects in the high predictability condition ($M = 16$ ms, $SD = 15$) were significantly lower than the DRB effects in the low

predictability condition ($M = 23$ ms, $SD = 14$), $W = 341$, $p = .001$, $r = 0.30$, $BF_{01} = 0.91$ (one-tailed; see **Figure 4**). The post-hoc analysis showed that the DRB effects were significantly different from zero for the low predictability condition (one-tailed, $V = 493$, $p < .001$, $r = 1.05$, $BF_{01} < 0.01$), and for the high predictability condition (one-tailed, $V = 492$, $p < .001$, $r = 1.04$, $BF_{01} < 0.01$).

Probe Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 4 % of all trials were relevant for this analysis).

A Wilcoxon-Signed Rank Test suggested that the DRB effects in the high predictability condition ($M = 4$ %, $SD = 5$) were not significantly lower than the DRB effects in the low predictability condition ($M = 5$ %, $SD = 5$), $W = 535$, $p = .224$, $r = 0.10$, $BF_{01} = 2.74$ (one-tailed). The post-hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (one-tailed, $V = 479$, $p < .001$, $r = 0.81$, $BF_{01} < 0.01$), and for the high predictability condition (one-tailed, $V = 454$, $p < .001$, $r = 0.79$, $BF_{01} = 0.01$).

Exp. 3, Only Probe Manipulation

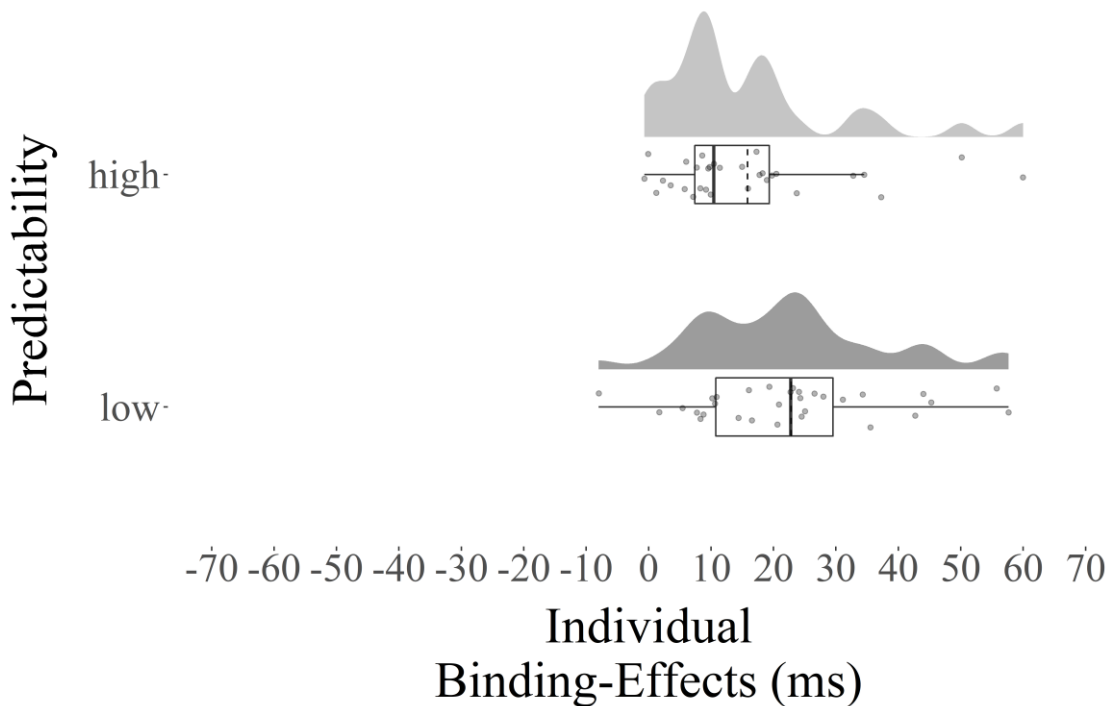


Figure 4. Individual Binding Effects for Experiment 3. Raincloud plots (Allen et al., 2019) for individual DRB effects in the high and low predictability condition for Experiment 3. The solid vertical line in each boxplot represents the median of the distribution; the dashed line represents the mean of the distribution. Upper and lower whiskers extend to the largest / smallest value above / below the respective hinge but at most 1.5 times the interquartile range above and below the third and first quartiles (McGill et al., 1978). Note that participants show lower S-R binding effects when distractor stimuli are predictable.

General Discussion

The present study investigated the influence of predictability on DRB effects. Previous studies found that less attention is focused on predictable stimuli and that S-R binding effects hinge (at least partially) on attention. Against this background, we tested whether S-R binding effects might be reduced when stimulus predictability is high. We further tried to selectively attribute the effect of predictability to S-R feature integration and retrieval (Frings et al., 2020). We did so by first testing whether the DRB effect is sensitive to distractor predictability at all in Experiment 1. As this was the case, we selectively manipulated predictability with respect to

the prime distractors at the probe distractors in Experiments 2a and 2b, respectively. We found that the impact of predictability was restricted to the predictability of probe distractors, suggesting that predictability manipulations affect the retrieval of event-files but not their integration. **In Experiment 3, we could replicate the finding of Experiment 2b in an offline setting – which further underlines the role of predictability for the retrieval process.**

The findings of Experiment 2a confirm that integration seems to be a rather automatic process (Hommel, 2004), in the sense that it does not rely or depend on any particular top-down processes or attentional attitudes. Given the findings from the attention literature, higher predictability should decrease attention (e.g., Failing et al., 2019; Failing & Theeuwes, 2020; Goschy et al., 2014; Wang & Theeuwes, 2018a, 2018b, 2018c, Huang et al., 2021; van Moorselaar & Theeuwes, 2021; Wang et al., 2019), which should have reduced the impact of distractors. As this did not affect the S-R binding effect, it seems fair to assume that integration is not only automatic but also requires little attention to unfold (Hommel, 2005). The fact that we still observed DRB effects in the high predictability condition indicates that the distractors have been processed. Yet, since DRB effects did not differ between conditions, it is reasonable to assume that even superficial stimulus processing is sufficient for integration.

This does not seem to hold for the retrieval process, as predictable distractors apparently reduced the probability that the previous event-file was retrieved (Experiment 2b and 3). **Note also that the replication yielded less noisy data in terms of less overall errors and faster reaction times, which further underlines that replicating our findings offline lead to an overall cleaner estimate of the actual effects. As pointed out in the Discussion of Experiment 2b, we expected effects to be weaker in Experiment 2 compared to Experiment 1 due to the noise-increasing effect of encoding specificity. Nonetheless, given the reliable modulating effect of predictability in Experiment 1, we think it is safe to assume that – even though encoding specificity might have lessened the strength of our modulating effect in Experiment 2b and 3 – the sole source of the modulation must be the retrieval process. Given the ample**

evidence from the visual search literature, we suggest that predictability modulates event-file retrieval via attention. This observation fits also well with other findings that retrieval is particularly sensitive to manipulations of attention (Ihrke et al., 2011; Schmalbrock & Frings, 2022; Schmalbrock et al., 2022). Several mechanisms might come into play here. For instance, as suggested in the introduction, it has been shown that reduced stimulus variance dampens curiosity and thereby the processing of these stimuli (Frings et al., 2019). In addition, continuously ignoring the same distractor might lead to increased inhibition of such a stimulus and as a consequence this stimulus retrieves less (but see Giesen et al., 2012). Finally, perceptual habituation might play a role here, too (see Failing & Theeuwes, 2020 or Narhi-Martinez et al., 2022 for an overview). It goes beyond the current paper to pinpoint which of these explanations affects event-file retrieval here. At an overarching level, it seems clear that predictable distractors are processed in a way that they retrieve less.

Following findings from the attention literature, we assumed that predictability might play an important role in S-R binding – an assumption our results support. Further, we showed that both processes contributing to S-R binding effects are not equally affected by our predictability manipulation. Our findings are in line with recent theorizing (Frings et al., 2020) suggesting that experimentally disentangling the two processes that might contribute to S-R binding effects – integration and retrieval – is likely to deepen our understanding of the role of event-files in action control. Our data suggest that integration of features is a rather automatic process that is not modulated by predictability while event-file retrieval is. This fits with recent theorizing on the role of interindividual differences on S-R binding effects (Hommel, 2022).

Data Availability Statement

The data for all experiments are available at PsychArchives under <http://dx.doi.org/10.23668/psycharchives.8396> . Experiment 3 was preregistered, and this preregistration can be found at <https://aspredicted.org/zd94a.pdf> .

Declarations

The research reported in this article was supported by the Deutsche Forschungsgemeinschaft (FR2133/15-1 and FR2133/15-2).

References

- Allen M., Poggiali D., Whitaker, K., Marshall, T. R., & Kievit, R. A. Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Research* 2019, 4(63). <https://doi.org/10.12688/wellcomeopenres.15191.1>
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, 37(3), 379–384. <https://doi.org/10.3758/BF03192707>
- Berlyne, D. E. (1949). Interest as a psychological concept. *British Journal of Psychology. General Section*, 39(Pt. 4), 184–195. <https://doi.org/10.1111/j.2044-8295.1949.tb00219.x>
- Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill Book Company. <https://doi.org/10.1037/11164-000>
- Failing, M., & Theeuwes, J. (2020). More capture, more suppression: Distractor suppression due to statistical regularities is determined by the magnitude of attentional capture. *Psychonomic Bulletin & Review*, 27(1), 86–95. <https://doi.org/10.3758/s13423-019-01672-z>
- Failing, M., Wang, B., & Theeuwes, J. (2019). Spatial suppression due to statistical regularities is driven by distractor suppression not by target activation. *Attention, Perception, & Psychophysics*, 81(5), 1405–1414. <https://doi.org/10.3758/s13414-019-01704-9>
- 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(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Frings, C., & Hommel, B. (2020). The disintegration of event-files over time: Decay or interference? *Psychonomic Bulletin & Review*, 27(4), 751–757. <https://doi.org/10.3758/s13423-020-01738-3>

- Frings, C., Hommel, B., Koch, I., Rothermund, K., Dignath, D., Giesen, C., Kiesel, A., Kunde, W., Mayr, S., Moeller, B., Möller, M., Pfister, R., & Philipp, A. (2020). Binding and Retrieval in Action Control (BRAC). *Trends in Cognitive Sciences*, 24(5), 375–387. <https://doi.org/10.1016/j.tics.2020.02.004>
- Frings, C., Merz, S., & Hommel, B. (2019). The impact of stimulus uncertainty on attentional control. *Cognition*, 183, 208–212. <https://doi.org/10.1016/j.cognition.2018.10.017>
- Frings, C., & Moeller, B. (2012). The horserace between distractors and targets: Retrieval-based probe responding depends on distractor–target asynchrony. *Journal of Cognitive Psychology*, 24(5), 582–590. <https://doi.org/10.1080/20445911.2012.666852>
- Frings, C., Rothermund, K., & Wentura, D. (2007). Distractor repetitions retrieve previous responses to targets. *Quarterly Journal of Experimental Psychology (2006)*, 60(10), 1367–1377. <https://doi.org/10.1080/17470210600955645>
- Geissler, C. F., Frings, C., & Moeller, B. (2021). Illuminating the prefrontal neural correlates of action sequence disassembling in response-response binding. *Scientific Reports*, 11(1), 22856. <https://doi.org/10.1038/s41598-021-02247-6>
- Geyer, T., Müller, H. J., & Krummenacher, J. (2008). Expectancies modulate attentional capture by salient color singletons. *Vision Research*, 48(11), 1315–1326. <https://doi.org/10.1016/j.visres.2008.02.006>
- Giesen, C., Frings, C., & Rothermund, K. (2012). Differences in the strength of distractor inhibition do not affect distractor–response bindings. *Memory & Cognition*, 40(3), 373–387. <https://doi.org/10.3758/s13421-011-0157-1>
- Goschy, H., Bakos, S., Müller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology*, 5, 1195. <https://doi.org/10.3389/fpsyg.2014.01195>

- Henson, R. N., Eckstein, D., Waszak, F., Frings, C., & Horner, A. J. (2014). Stimulus-response bindings in priming. *Trends in Cognitive Sciences*, *18*(7), 376–384.
<https://doi.org/10.1016/j.tics.2014.03.004>
- Hommel, B. (1998). Event-files: Evidence for Automatic Integration of Stimulus-Response Episodes. *Visual Cognition*, *5*(1-2), 183–216. <https://doi.org/10.1080/713756773>
- Hommel, B. (2004). Event-files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*(11), 494–500. <https://doi.org/10.1016/j.tics.2004.08.007>
- Hommel, B. (2005). How much attention does an event-file need? *Journal of Experimental Psychology. Human Perception and Performance*, *31*(5), 1067–1082.
<https://doi.org/10.1037/0096-1523.31.5.1067>
- Hommel, B. (2022). The Control of Event-File Management. *Journal of Cognition*, *5*(1).
- Hommel, B., Memelink, J., Zmigrod, S., & Colzato, L. S. (2014). Attentional control of the creation and retrieval of stimulus-response bindings. *Psychological Research PRPF*, *78*(4), 520–538. <https://doi.org/10.1007/s00426-013-0503-y>
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*(5), 849-78; discussion 878-937.
<https://doi.org/10.1017/s0140525x01000103>
- Huang, C., Vilotijević, A., Theeuwes, J., & Donk, M. (2021). Proactive distractor suppression elicited by statistical regularities in visual search. *Psychonomic Bulletin & Review*, *28*(3), 918–927. <https://doi.org/10.3758/s13423-021-01891-3>
- Ihrke, M., Behrendt, J., Schrobsdorff, H., Michael Herrmann, J., & Hasselhorn, M. (2011). Response-retrieval and negative priming. *Experimental Psychology*, *58*(2), 154–161.
<https://doi.org/10.1027/1618-3169/a000081>

- Jarosz, A. F., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *The Journal of Problem Solving*, 7(1).
<https://doi.org/10.7771/1932-6246.1167>
- Laub, R., & Frings, C. (2020). Distractor-based retrieval in action control: The influence of encoding specificity. *Psychological Research*, 84(3), 765–773.
<https://doi.org/10.1007/s00426-018-1082-8>
- Lawrence, M. A. (2016). *ez: Easy analysis and visualization of factorial experiments*. R Package Version 4.4-0. <https://cran.r-project.org/web/packages/ez/index.html>
- McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. *The American Statistician*, 32(1), 12–16. <https://doi.org/10.1080/00031305.1978.10479236>
- Memelink, J., & Hommel, B. (2013). Intentional weighting: A basic principle in cognitive control. *Psychological Research*, 77(3), 249–259. <https://doi.org/10.1007/s00426-012-0435-y>
- Mocke, V., Weller, L., Frings, C., Rothermund, K., & Kunde, W. (2020). Task relevance determines binding of effect features in action planning. *Attention, Perception, & Psychophysics*, 82(8), 3811–3831. <https://doi.org/10.3758/s13414-020-02123-x>
- Moeller, B., & Frings, C. (2014). Attention meets binding: Only attended distractors are used for the retrieval of event-files. *Attention, Perception, & Psychophysics*, 76(4), 959–978. <https://doi.org/10.3758/s13414-014-0648-9>
- Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.
- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes factors for common designs*. R Package Version 0.9.12-. <https://cran.r-project.org/web/packages/BayesFactor/index.html>

- Narhi-Martinez, W., Dube, B., & Golomb, J. D. (2022). Attention as a multi-level system of weights and balances. *Wiley Interdisciplinary Reviews: Cognitive Science*, e1633. <https://doi.org/10.1002/wcs.1633>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). Psychopy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Peirce, J., & MacAskill, M. (2018). *Building Experiments in PsychoPy*. SAGE.
- Poth, C. H., & Horstmann, G. (2017). Assessing the monitor warm-up time required before a psychological experiment can begin. *The Quantitative Methods for Psychology*, 13(3), 166-173. <http://dx.doi.org/10.20982/tqmp.13.3.p166>
- R Core Team. (2019). *A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2012.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237. <https://doi.org/10.3758/PBR.16.2.225>
- Schmalbrock, P., & Frings, C. (2021). Temporal expectancy modulates Stimulus-Response Integration. *Attention, Perception, & Psychophysics*, 1–10. <https://doi.org/10.3758/s13414-021-02361-7>
- Schmalbrock, P., & Frings, C. (2022). A mighty tool not only in perception: Figure-ground mechanisms control binding and retrieval alike. *Attention, Perception, & Psychophysics*, 1-16. <https://doi.org/10.3758/s13414-022-02511-5>
- Schmalbrock, P., Laub, R., & Frings, C. (2021). Integrating Salience and Action - Increased Integration Strength through Salience. *Visual Cognition*, advanced online publication. <https://doi.org/10.1080/13506285.2020.1871455>

- Schmalbrock, P., Kiesel, A., & Frings, C. (2022). What Belongs Together Retrieves Together—The Role of Perceptual Grouping in Stimulus-Response Binding and Retrieval. *Journal of Cognition*, 5(1). <http://doi.org/10.5334/joc.217>
- Singh, T., Moeller, B., Koch, I., & Frings, C. (2018). May I have your attention please: Binding of attended but response-irrelevant features. *Attention, Perception, & Psychophysics*, 80(5), 1143–1156. <https://doi.org/10.3758/s13414-018-1498-7>
- Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, 51(6), 599–606. <https://doi.org/10.3758/BF03211656>
- Theeuwes, J., & Failing, M. (2020). Attentional Selection. *Elements in Perception*. <https://doi.org/10.1017/9781108891288>
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1), 97-136. [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- Tukey, J. W. (1977). *Exploratory data analysis*. Addison-Wesley.
- van Moorselaar, D., & Theeuwes, J. (2022). Spatial suppression due to statistical regularities in a visual detection task. *Attention, Perception, & Psychophysics*, 84(2), 450-458. <https://doi.org/10.3758/s13414-021-02330-0>
- Wang, B., & Theeuwes, J. (2018a). How to inhibit a distractor location? Statistical learning versus active, top-down suppression. *Attention, Perception, & Psychophysics*, 80(4), 860–870. <https://doi.org/10.3758/s13414-018-1493-z>
- Wang, B., & Theeuwes, J. (2018b). Statistical regularities modulate attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, 44(1), 13–17. <https://doi.org/10.1037/xhp0000472>
- Wang, B., & Theeuwes, J. (2018c). Statistical regularities modulate attentional capture independent of search strategy. *Attention, Perception, & Psychophysics*, 80(7), 1763–1774. <https://doi.org/10.3758/s13414-018-1562-3>

- Wang, B., van Driel, J., Ort, E., & Theeuwes, J. (2019). Anticipatory Distractor Suppression Elicited by Statistical Regularities in Visual Search. *Journal of Cognitive Neuroscience*, 31(10), 1535–1548. https://doi.org/10.1162/jocn_a_01433
- Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. R Package Version 0.8.3. <https://cran.r-project.org/package=dplyr>
- Wilcoxon, F. (1992). *Individual comparisons by ranking methods*. *Breakthroughs in statistics*. Springer.

Appendix A

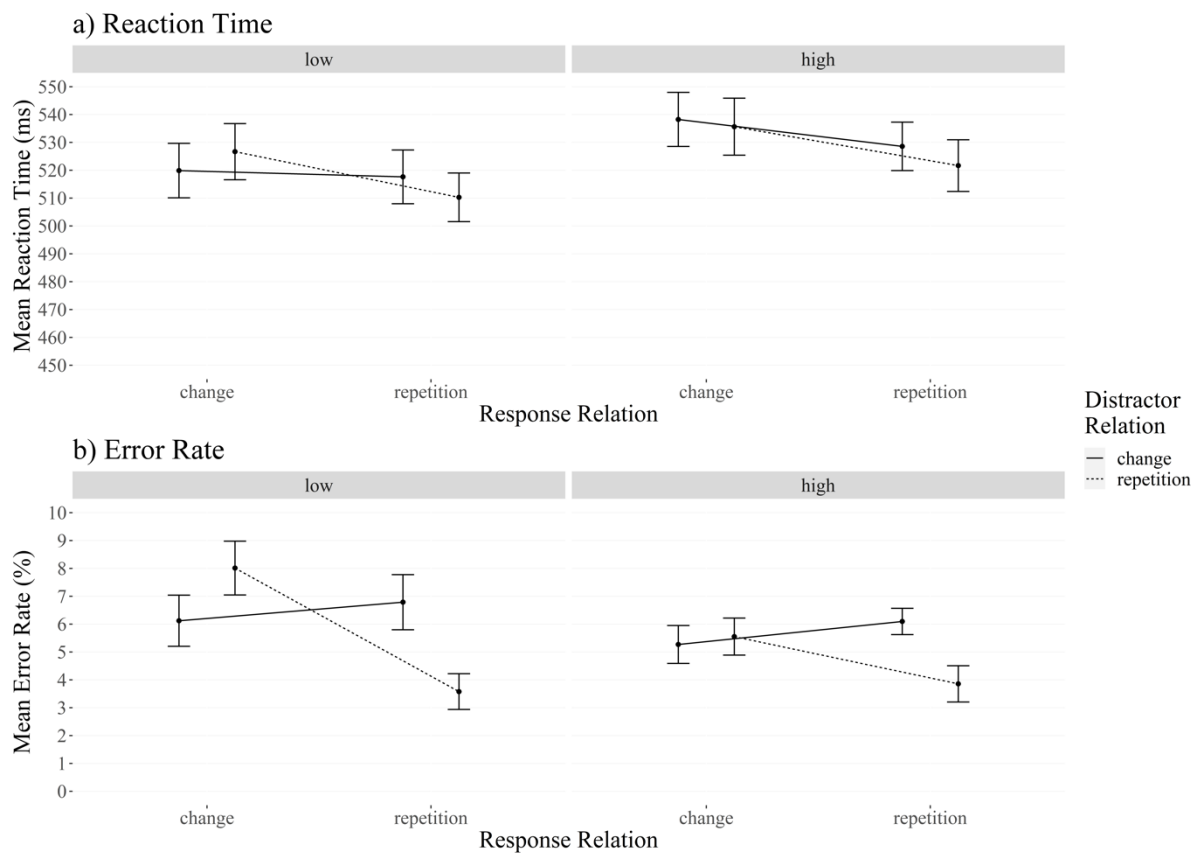


Figure A1. Mean reaction time and Error rates for Experiment 1. Reaction time (a) and error rates (b) as a function of response relation (repetition vs. change), distractor relation (repetition vs. change), and predictability (high vs. low). Error bars indicate between-participants standard error of the mean.

Appendix B

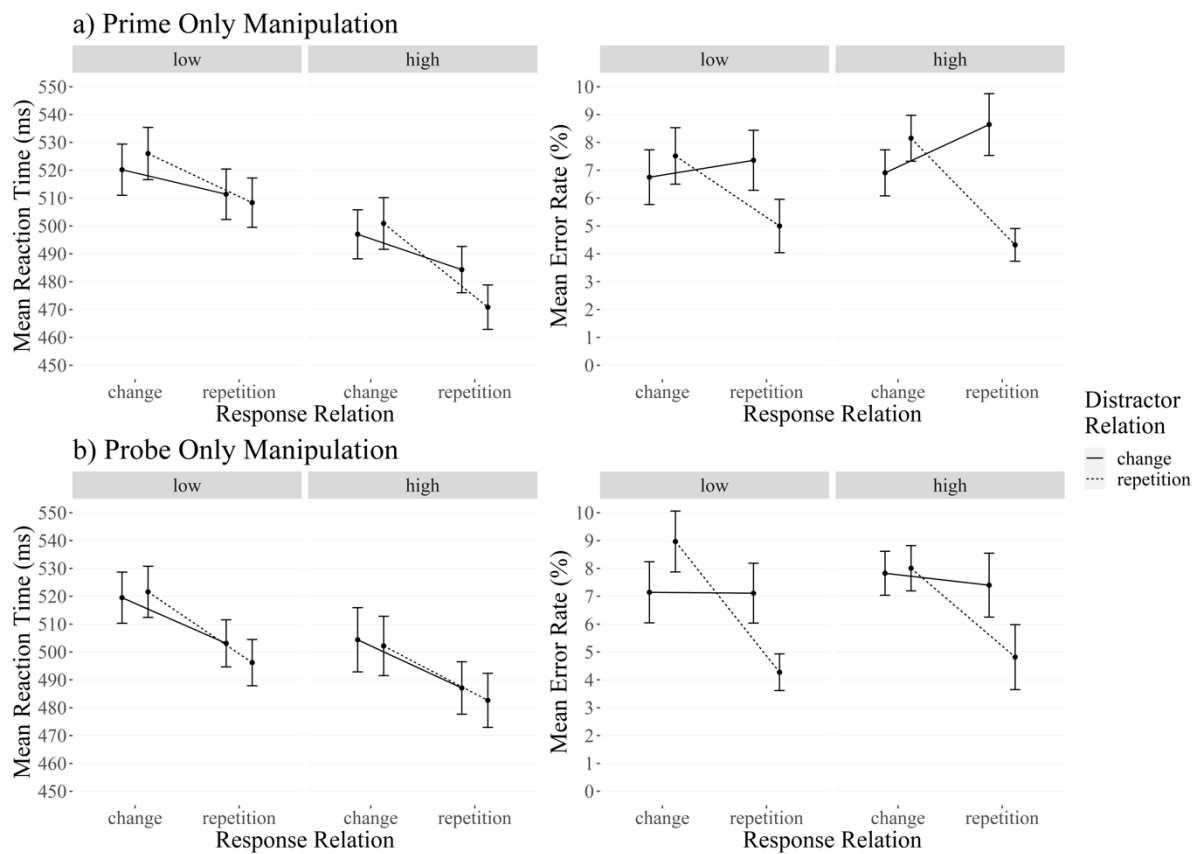


Figure B1. Mean reaction time and Error rates for Experiment 2a and 2b. Reaction time and error rates as a function of response relation (repetition vs. change), distractor relation (repetition vs. change), and predictability (high vs. low) for Experiment 2a (panel a) and 2b (panel b). Error bars indicate between-participants standard error of the mean.

Appendix C

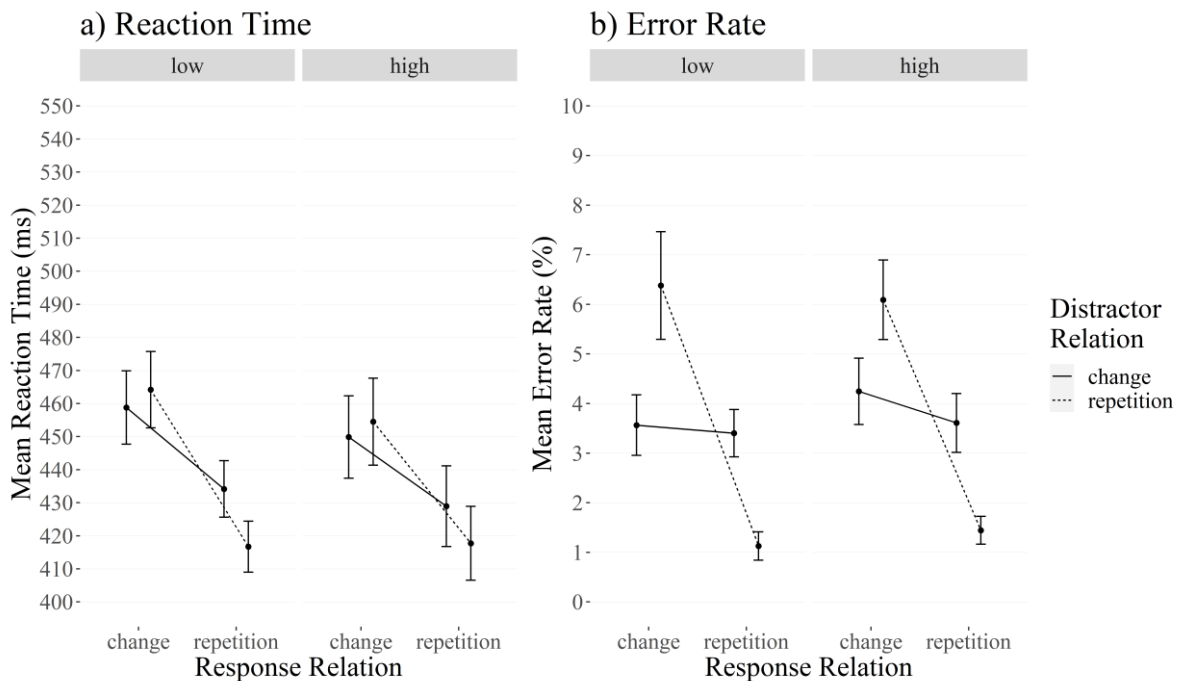


Figure C1. Mean reaction time and Error rates for Experiment 3. Reaction time (a) and error rates (b) as a function of response relation (repetition vs. change), distractor relation (repetition vs. change), and predictability (high vs. low). Error bars indicate between-participants standard error of the mean.

Supplementary Material

Experimental conditions were compared using a repeated-measures analysis of variance (ANOVA) with type-III sums of square, using the ‘ezAnova’- function from the package ‘ez’ (Lawrence, 2016). We report two effect sizes for ANOVAs: η_p^2 , and η_G^2 (Bakeman, 2005). The distractor-response binding effect is computed as the distractor repetition benefit in response repetition trials minus the distractor repetition interference in response change trials ($[RRDC-RRDR]-[RCDC-RCDR]$). Binding effects were compared using post-hoc t-tests complemented by Bayesian t-tests (Rouder et al., 2009) whose Bayes factor (BF_{01}) quantify the evidence in favor of the null hypothesis relative to the evidence in favor of the alternative hypothesis. Bayes factors were computed using the package ‘BayesFactor’ (Morey & Rouder, 2018).

Results ANOVA Experiment 1

Probe Reaction Time

A 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 4 (block order: A vs. B vs. C vs. D) mixed-effects ANOVA on probe RTs yielded no main effect for block order, $F(3, 28) = 0.87$, $p = .470$, $\eta_G^2 = .08$, $\eta_P^2 = .08$, indicating that there was no general performance difference between each group. No further interaction with block order reach significance, all $F_s < 1.3$ and $p_s > .290$. We, therefore, conclude that there was no influence of the order the trials in the high predictability condition were presented in and will exclude it henceforth from any following analysis.

A 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 2 (predictability: high vs. low) mixed-effects ANOVA on probe RTs yielded a significant interaction between response relation and distractor relation, $F(1, 62) = 14.42$, $p < .001$, $\eta_G^2 < .01$, $\eta_P^2 = .19$, indicating significant DRB effects. Intriguingly, this interaction was further modulated by predictability, resulting in a significant three-way interaction, $F(1, 62) = 4.12$, $p = .047$, $\eta_G^2 < .01$, $\eta_P^2 = .06$, suggesting that the DRB effect is modulated by the level of predictability (see **Figure 2**). This effect is further supplemented when Bayes factors are considered: A Welch two-sample t -test underlined that the DRB effects for the high predictability condition ($M = 4$ ms, $SD = 16$) were significantly lower from the DRB effect in the low predictability condition ($M = 14$ ms, $SD = 22$), two-tailed $t(62) = 2.03$, $p = .047$, $d_z = 0.25$, $BF_{01} = 0.70$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(31) = 3.64$, $p = .001$, $d_z = 0.64$, $BF_{01} = 0.03$) but not for the high predictability condition (two-tailed $t(31) = 1.47$, $p = .151$, $d_z = 0.26$, $BF_{01} = 1.99$).

Additionally, a main effect for response relation emerged, $F(1, 62) = 19.37$, $p < .001$, $\eta_G^2 = .01$, $\eta_P^2 = .23$. No further main effect or interaction reached significance, all $F_s < 1.16$ and $p_s > .280$.

Probe Error Rate

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 6% of all trials were relevant error trials).

The same mixed-effects ANOVA including block order yielded no main effect for block order, $F(3, 28) = 0.33, p = .805, \eta_G^2 = .02, \eta_P^2 = .03$, indicating that there was no general performance difference between each group. No other main interaction including block order reached significance, we excluded the factor block order from the following analysis, all $F_s < 2.37$ and $p_s > .09$.

The mixed-effects ANOVA excluding block order, but including predictability yielded a significant interaction between response relation and distractor relation, $F(1, 62) = 19.23, p < .001, \eta_G^2 = .05, \eta_P^2 = .24$, again, indicating significant DRB effect. However, this interaction was not further modulated by predictability, $F(1, 62) = 2.19, p = .144, \eta_G^2 = .01, \eta_P^2 = .03$. A Welch two-sample t -test underlined that the DRB effects for the low predictability condition ($M = 5.10\%$, $SD = 8.56$) were not significantly different from the DRB effects in the high predictability condition ($M = 2.52\%$, $SD = 4.84$), two-tailed $t(62) = 1.48, p = .145, d_z = 0.19, BF_{01} = 1.56$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(31) = 3.37, p = .002, d_z = 0.60, BF_{01} = 0.06$) and also for the high predictability condition (two-tailed $t(29) = 2.69, p = .006, d_z = 0.52, BF_{01} = 0.15$).

Additionally, several main effects emerged. A main effect for response relation was observed, $F(1, 62) = 4.28, p = .043, \eta_G^2 = .02, \eta_P^2 = .06$. Participants made more errors when responses changed from prime to probe ($M = 6.24\%$, $SD = 4.70$), compared to when the response repeated from prime to probe ($M = 5.08\%$, $SD = 4.23$). A main effect for distractor relation emerged, $F(1, 62) = 4.57, p = .037, \eta_G^2 < .01, \eta_P^2 = .07$. Participants made more errors when the distractor changed ($M = 6.07\%$, $SD = 4.46$), compared to when it repeated ($M = 5.25\%$, $SD = 4.52$). No further main effect or interaction reached significance, all $F_s < 2.20$ and $p_s > .142$.

Results ANOVA Experiment 2a

Probe Reaction Times

A 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 4 (block order: A vs. B) mixed-effects ANOVA on probe RTs yielded no main effect for block order, $F(1, 30) = 0.11, p = .747, \eta_G^2 < .01, \eta_P^2 < .01$, indicating that there was no general performance difference between each group. The interaction between block order and response relation reached

significance, $F(1, 30) = 6.60, p = .015, \eta_G^2 < .01, \eta_P^2 = .18$. This is however of no concern to our research question. No further interaction with block order reach significance, all F s < 2.65 and p s $> .11$. We, therefore, conclude that there was no influence of the order the trials in the high predictability condition were presented in and will exclude it henceforth from any following analysis.

For the main analysis, a 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 2 (predictability: high vs. low) mixed-effects ANOVA on probe RTs yielded a significant interaction between response relation and distractor relation, $F(1, 62) = 22.05, p < .001, \eta_G^2 < .01, \eta_P^2 = .26$, indicating significant DRB effects.

Intriguingly, this interaction was not further modulated by predictability, $F(1, 62) = 2.33, p = .132, \eta_G^2 < .01, \eta_P^2 = .03$, suggesting that distractor-response integration is not modulated by the level of predictability (see **Figure 2**). A Welch two-sample t -test also underlined that the DRB effects for the high predictability condition ($M = 17$ ms, $SD = 22$) were not significantly lower compared to the DRB effects in the low predictability condition ($M = 9$ ms, $SD = 23$), two-tailed $t(62) = -1.53, p = .132, d_z = 0.19, BF_{01} = 1.47$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(31) = 2.21, p = .034, d_z = 0.39, BF_{01} = 0.64$) and also for the high predictability condition (two-tailed $t(31) = 4.46, p < .001, d_z = 0.79, BF_{01} < 0.01$).

Additionally several main and interaction effects emerged. A main effect for response relation emerged, $F(1, 62) = 42.04, p < .001, \eta_G^2 = .03, \eta_P^2 = .40$. Participants responded faster when the response repeated ($M = 494$ ms, $SD = 51$) compared to when it changed ($M = 511$ ms, $SD = 53$). Further, a main effect for predictability emerged, $F(1, 62) = 5.43, p = .023, \eta_G^2 = .07, \eta_P^2 = .08$. Participants responded faster when the prime distractor predictability was high ($M = 488$ ms, $SD = 50$) compared to when it was low ($M = 516$ ms, $SD = 51$). Also an interaction between response relation and predictability emerged, $F(1, 62) = 5.47, p = .023, \eta_G^2 < .01, \eta_P^2 = .08$. Responses were slowest when predictability was low and the response changed ($M = 523$ ms, $SD = 52$), somewhat faster when predictability was low and the response repeated ($M = 510$ ms, $SD = 50$), even faster when predictability was high and the response changed ($M = 499$ ms, $SD = 51$), and fastest when predictability was high and the response repeated ($M = 478$ ms, $SD = 46$). No further main effect or interaction reached significance, all F s < 2.34 and p s $> .132$.

Probe Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 7% of all trials were relevant error trials).

The same mixed-effects ANOVA including block order yielded no main effect for block order, $F(1, 30) = 0.31, p = .581, \eta_G^2 < .01, \eta_P^2 = .01$, indicating that there was no general performance difference between each group. No other main or interaction effect including block order reached significance, we excluded the factor block order from the following analysis, all $F_s < 0.94$ and $p_s > .330$.

The mixed-effects ANOVA excluding block order, but including predictability yielded a significant interaction between response relation and distractor relation, $F(1, 62) = 19.92, p < .001, \eta_G^2 = .04, \eta_P^2 = .24$, again, indicating significant DRB effects. However, this interaction was also not further modulated by predictability, $F(1, 62) = 1.57, p = .214, \eta_G^2 < .01, \eta_P^2 = .02$. A Welch two-sample *t*-test underlined that the DRB effects for the low predictability condition ($M = 3\%, SD = 8$) were not significantly different from the DRB effects in the high predictability condition ($M = 6\%, SD = 8$), two-tailed $t(62) = -1.26, p = .214, d_z = 0.16, BF_{01} = 2.01$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(31) = 2.35, p = .025, d_z = 0.42, BF_{01} = 0.49$) and also for the high predictability condition (two-tailed $t(31) = 3.91, p < .001, d_z = 0.69, BF_{01} = 0.02$).

Additionally, a main effects for distractor relation was observed, $F(1, 62) = 5.94, p = .018, \eta_G^2 = .01, \eta_P^2 = .09$. Participants made more errors when the distractor changed from prime to probe ($M = 7\%, SD = 6$), compared to when the distractor repeated from prime to probe ($M = 6\%, SD = 5$). No other main effect or interaction reached significance, all $F_s < 2.33$ and $p_s > .130$.

Results ANOVA Experiment 2b*Probe Reaction Times*

A 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 4 (block order: A vs. B) mixed-effects ANOVA on probe RTs yielded a main effect for block order, $F(1, 29) = 5.14, p = .031, \eta_G^2 = .14, \eta_P^2 = .15$. Participants in Group A ($M = 472$ ms, $SD = 43$) were significantly faster than Group B ($M = 515$ ms, $SD = 62$). Since this is a between-participant comparison

and concerns a main effect, it is of no relevance to our main research question. The interaction between block order and response relation reached significance, $F(1, 29) = 4.86, p = .036, \eta_G^2 < .01, \eta_P^2 = .14$. This is however of no concern to our research question. No further interaction with block order reach significance, all $F_s < 0.07$ and $p_s > .930$. We, therefore, conclude that there was no influence of the order the trials in the high predictability condition were presented in and will exclude it henceforth from any following analysis.

For the main analysis, a 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 2 (predictability: high vs. low) mixed-effects ANOVA on probe RTs yielded a significant interaction between response relation and distractor relation, $F(1, 60) = 5.29, p = .025, \eta_G^2 < .01, \eta_P^2 = .08$, indicating significant S-R binding.

This interaction was not further modulated by predictability, $F(1, 60) = 1.92, p = .171, \eta_G^2 < .01, \eta_P^2 = .03$, suggesting that distractor-response retrieval is not modulated by the level of predictability (see **Figure 3**). A Welch two-sample t -test also underlined that the DRB effects for the high predictability condition ($M = 2$ ms, $SD = 19$) were not significantly lower compared to the DRB effects in the low predictability condition ($M = 9$ ms, $SD = 20$), two-tailed $t(60) = 1.39, p = .171, d_z = 0.18, BF_{01} = 1.73$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(30) = 2.55, p = .016, d = 0.45, BF_{01} = 0.33$), but not for the high predictability condition (two-tailed $t(30) = 0.66, p = .661, d = 0.12, BF_{01} = 4.27$).

Additionally several main and interaction effects emerged. A main effect for response relation emerged, $F(1, 62) = 56.51, p < .001, \eta_G^2 = .04, \eta_P^2 = .49$. Participants responded faster when the response repeated ($M = 492$ ms, $SD = 50$) compared to when it changed ($M = 512$ ms, $SD = 57$). Further, a main effect for distractor relation emerged, $F(1, 60) = 5.63, p = .021, \eta_G^2 < .01, \eta_P^2 = .09$. Participants responded faster when the distractor repeated ($M = 501$ ms, $SD = 54$) compared to when it changed ($M = 504$ ms, $SD = 55$). No further main effect or interaction reached significance, all $F_s < 1.48$ and $p_s > .220$.

Probe Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 7% of all trials were relevant error trials).

The same mixed-effects ANOVA including block order yielded a main effect for block order, $F(1, 29) = 6.86, p = .014, \eta_G^2 = .11, \eta_P^2 = .19$. Participants in Group A ($M = 9\%$, $SD = 6$) made more errors than Group B ($M = 5\%$, $SD = 4$). Since this is a between-participant comparison and concerns a main effect, it is of no relevance to our main research question. No other interaction including block order reached significance, we excluded the factor block order from the following analysis, all $F_s < 2.90$ and $p_s > .100$.

The mixed-effects ANOVA excluding block order, but including predictability yielded a significant interaction between response relation and distractor relation, $F(1, 60) = 13.90, p < .001, \eta_G^2 = .03, \eta_P^2 = .19$, again, indicating significant DRB effects. However, this interaction was not further modulated by predictability, $F(1, 60) = 0.91, p = .344, \eta_G^2 < .01, \eta_P^2 = .01$. A Welch two-sample t -test underlined that the DRB effects for the low predictability condition ($M = 5\%$, $SD = 8$) were not significantly different from the DRB effect in the high predictability condition ($M = 3\%$, $SD = 7$), two-tailed $t(60) = 0.95, p = .344, d_z = 0.12, BF_{01} = 2.63$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(30) = 3.13, p < .001, d_z = 0.56, BF_{01} = 0.10$) and also for the high predictability condition (two-tailed $t(30) = 2.09, p = .045, d_z = 0.38, BF_{01} = 0.78$).

Additionally, a main effects for distractor relation was observed, $F(1, 60) = 4.17, p = .045, \eta_G^2 < .01, \eta_P^2 = .06$. Participants made more errors when distractor changed from prime to probe ($M = 7.37\%$, $SD = 6$), compared to when the distractor repeated from prime to probe ($M = 6.51\%$, $SD = 6$). Also a main effects for response relation was observed, $F(1, 60) = 9.71, p = .003, \eta_G^2 = .03, \eta_P^2 = .13$. Participants made more errors when response changed from prime to probe ($M = 8\%$, $SD = 5$), compared to when the response repeated from prime to probe ($M = 6\%$, $SD = 6$). No other main effect or interaction reached significance, all $F_s < 0.70$ and $p_s > .400$.

Results ANOVA Experiment 3

Probe Reaction Times

A 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 4 (block order: A vs. B) mixed-effects ANOVA on probe RTs yielded no main effect for block order, $F(1, 29) = 0.63, p = .435, \eta_G^2 = .02, \eta_P^2 = .02$, indicating that there was no general performance difference between each group. No further interaction with block order reach significance, all $F_s < 0.17$ and $p_s > .680$. We, therefore, conclude that there was no influence of the order the trials in the high predictability condition were presented in and will exclude it henceforth from any following analysis.

For the main analysis, a 2 (response relation: repetition vs. change) x 2 (distractor relation: repetition vs. change) x 2 (predictability: high vs. low) mixed-effects ANOVA on probe RTs yielded a significant interaction between response relation and distractor relation, $F(1, 60) = 106.43, p < .001, \eta_G^2 < .01, \eta_P^2 = .64$, indicating significant DRB effects.

Intriguingly, this interaction was not further modulated by predictability, $F(1, 60) = 3.47, p = .068, \eta_G^2 < .01, \eta_P^2 = .05$, suggesting that S-R integration is not modulated by the level of predictability (see **Figure 2**). A Welch two-sample t -test also underlined that the DRB effects for the high predictability condition ($M = 16$ ms, $SD = 15$) were not significantly lower compared to the DRB effects in the low predictability condition ($M = 23$ ms, $SD = 14$), two-tailed $t(59.69) = 1.86, p = .068, d_z = 0.24, BF_{01} = 0.91$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(30) = 8.32, p < .001, d_z = 1.49, BF_{01} < 0.01$) and also for the high predictability condition (two-tailed $t(30) = 6.21, p < .001, d_z = 1.11, BF_{01} < 0.01$).

Additionally, several main effects emerged. A main effect for response relation emerged, $F(1, 60) = 84.21, p < .001, \eta_G^2 = .07, \eta_P^2 = .58$. Participants responded faster when the response repeated ($M = 424$ ms, $SD = 56$) compared to when it changed ($M = 457$ ms, $SD = 67$). Further, a main effect for the distractor relation emerged, $F(1, 60) = 33.99, p < .001, \eta_G^2 < .01, \eta_P^2 = .36$. Participants responded faster when the prime distractor repeated ($M = 438$ ms, $SD = 65$) compared to when it changed ($M = 443$ ms, $SD = 63$). No further main effect or interaction reached significance, all $F_s < 2.94$ and $p_s > .09$.

Probe Error Rates

For the same analysis on probe error rates, only trials with correct prime responses but incorrect probe responses were considered (i.e., 4% of all trials were relevant error trials).

The same mixed-effects ANOVA including block order yielded no main effect for block order, $F(1, 29) = 1.72, p = .199, \eta_G^2 = .03, \eta_P^2 = .03$, indicating that there was no general performance difference between each group. No other main or interaction effect including block order reached significance, we excluded the factor block order from the following analysis, all $F_s < 1.40$ and $p_s > .240$.

The mixed-effects ANOVA excluding block order, but including predictability yielded a significant interaction between response relation and distractor relation, $F(1, 60) = 57.64, p < .001, \eta_G^2 = .09, \eta_P^2 = .49$, again, indicating significant DRB effects. However, this interaction was also not further modulated by predictability, $F(1, 60) = 0.81, p = .371, \eta_G^2 < .01, \eta_P^2 = .01$. A Welch two-sample t -test underlined that the DRB effects for the low predictability condition ($M = 5\%, SD = 5$) were not significantly different from the DRB effects in the high predictability condition ($M = 4\%, SD = 4$), two-tailed $t(59.42) = 0.90, p = .371, d_z = 0.11, BF_{01} = 2.74$. Post-Hoc analysis evidenced that the DRB effects were significantly different from zero for the low predictability condition (two-tailed $t(30) = 5.73, p < .001, d_z = 1.03, BF_{01} < 0.01$) and also for the high predictability condition (two-tailed $t(30) = 4.98, p < .001, d_z = 0.89, BF_{01} < 0.01$).

Additionally, a main effects for response relation was observed, $F(1, 60) = 31.43, p < .001, \eta_G^2 = .12, \eta_P^2 = .34$. Participants made more errors when the response repeated from prime to probe ($M = 3\%, SD = 2$), compared to when the response changed from prime to probe ($M = 5\%, SD = 5$). No other main effect or interaction reached significance, all $F_s < 0.49$ and $p_s > .490$.