A version of this pre-print was accepted for publications in Memory & Cognition (7 December 2021).

A Cognitive Model of Response Omissions in Distraction Paradigms

Karlye A. M. Damaso[1]
Spencer C. Castro[2]
Juanita Todd[1]
David L. Strayer[2]
Alexander Provost[1]
Dora Matzke[3]
Andrew Heathcote[1]

[1] University of Newcastle, Australia
[2] University of Utah, United States of America
[3] University of Amsterdam, Netherlands

Corresponding author:
Karlye Damaso
University of Newcastle, School of Psychology
University Drive
Callaghan NSW 2308
Australia
+61 2 4921 5000
karlye.damaso.psyc@gmail.com

Abstract

The effects of distraction on responses manifest in three ways, prolonged reaction times, and increased error and response omission rates. However, the latter effect is often ignored or assumed to be due to a separate cognitive process. We investigated omissions occurring in two paradigms that manipulated distraction. One required simple stimulus detection of younger participants, the second required choice responses and was completed by both younger and older participants. We fit data from these paradigms with a model that identifies three causes of omissions: two are related to the process of accumulating the evidence on which a response is based: *intrinsic omissions* (due to between-trial variation in accumulation rates making it impossible to ever reach the evidence threshold) and *design omissions* (due to response windows that cause slow responses not to be recorded; a third, *contaminant omissions*, allows for a cause unrelated to the response process. In both data sets systematic differences in omission rates across conditions were accounted for by process omissions. Intrinsic omissions played a lesser role than design omission, even though the presence of design omissions was not evident in descriptive analyses of the data. The model provided an accurate account of all aspects of the detection data and the choice-response data, but slightly under-estimated overall omissions in the choice paradigm, particularly in older participants, suggesting further investigation of contaminant omission effects is needed.

Although the limits of cognitive capacity affect general performance, they are probably most apparent under conditions of distraction – that is, when multiple demands are simultaneously placed upon a person's attention. Distraction and its impact on behavioural performance is a widely studied phenomenon that has been utilised to inform theoretical understandings of cognition and attention (e.g., Lavie 1995, 2005, 2010; Nätäänen, 1990, Nätäänen, Kujala, & Winkler, 2011); functional aberrations in a range of psychiatric conditions (Cortiñas et al., 2008; Gumenyuk et al., 2005); fatigue (Lim & Dinges, 2010); multi-tasking (Sanbonmatsu, Strayer, Medeiros-Ward, & Watson, 2013); and neuroanatomical function (Rinnie et al., 2007; Sabri et al., 2006).

Empirically speaking, distraction manifests in data in three ways: prolonged response times (RTs) and increased error and response-omission rates. Although prolonged RTs and increased error rates have been the subject of extensive investigation (e.g., Liete & Ratcliff, 2010; Luce 1986; Ratcliff, 1978), increased omissions (also known as 'misses' or 'non-responses') have not. In many cases omission rates are not reported, and sometimes they are even pooled with commission error responses and generically classed as errors (e.g., Mager et al., 2006; Miller, Price, Okun, Montijo, Bowers, 2009).

We know of no reason, apart from rarity, that omissions do not have the same status as a measure of distraction as accuracy and RT. On a practical level, step omissions, that is leaving out a necessary step in a task sequence, are regarded as the most common human error, with a proliferation of research aimed at developing strategies to avoid such omissions (for reviews, see Reason, 2002, 1998). On a theoretical level, ignoring omissions represents a commitment to their arising from an entirely independent "contaminant" mechanism that can be safely ignored when interest focuses on the process generating responses. On an empirical level, there are often practical reasons that researchers must stipulate response windows outside of which responses are not recorded, resulting in the creation of omissions by design.

This can sometimes be justified if such “outlying” (i.e., unusually fast or slow) responses are classed as being in part or whole generated by a contaminant process. However, to the degree this is not true, there is the risk of potentially misleading impacts on summary statistics for RTs and error rates particularly when RT distributions have been cut off (Ulrich & Miller, 1994). Although there are ways that these impacts can be avoided (see, Heathcote, 1996; Kendall & Stuart,1967; Ulrich & Miller, 1994), omissions that occur as a consequence of design can be hard to detect, particularly when they occur at relatively low levels, as they are not always apparent in visual inspections of the RT distribution (Ulrich & Miller, 1994).

In the current paper we explore the degree to which omissions arise from a contaminant process versus from the same cognitive mechanisms producing responses, with the latter being described as “process” omissions. We examine two data sets – one from an experiment reported by Castro et al. (2019) where participants made a detection response and a new experiment requiring choice responses. In both experiments, factors causing distraction were manipulated. We propose a cognitive model based on Brown and Heathcote’s (2008) Linear Ballistic Accumulator (LBA) that provides a unified account of the probability of omissions and responses, including the distribution of RTs for both detection and choice responses. The model also allows us to differentiate between two types of process omissions, “*design*” omissions occurring due to slow RT cut-offs imposed in the experiments we examined, and “*intrinsic*” omission caused by failures of the cognitive process to produce a response. We differentiate “*contaminant”* omissions from the process omissions on the basis that the latter are affected by the same factors that influence responses, whereas the former are not.

We aim to establish a methodology to determine the prevalence of the three types (*contaminant*, *design,* and *intrinsic*) of omissions. In the next section, we provide a brief overview of the relatively small body of research that has acknowledged the occurrence of omissions. We then propose our cognitive model of omissions and apply it to data.

Omissions in Past Distraction Research

Past research using omissions as a dependent variable to measure attention and/or distraction has used both detection and choice tasks. Examples of detection tasks include the psychomotor vigilance task (PVT; e.g., Dinges & Powell, 1985; Lim & Dinges, 2008), the detection response task (DRT; e.g., Castro, Strayer, Matzke, & Heathcote, 2019; Howard, Evans, Innes, Brown, & Eidels, 2020), and the continuous performance task (CPT; e.g., Rosvold et al., 1956). Examples of choice tasks include the distraction paradigm (e.g., Schröger & Wolff, 1998) n-back task (Evans, Steyvers, & Brown, 2018) and dual-task perceptual choice (Howard et al., 2020).

In the PVT and DRT participants must maintain a vigil for a set period of time and are required to respond quickly to a target stimulus that occurs at randomly spaced intervals. The tasks differ from each other in the modality, frequency, and presentation duration of the target, as well as the trial durations and conditions under which the tasks are performed. The PVT is used to study fatigue due to sleep loss. Sleep deprived participants are required to monitor a blank computer screen for the sudden onset of a visual millisecond counter. The counter starts from zero and continuously increments until the participant responds or 30 sec has passed. Omissions in the PVT can also be referred to as 'sleep attacks' or 'lapses' and are typically the dependent variable most impacted by fatigue (Lim & Dinges, 2010). The DRT is an International Standards Organization (ISO, 2016) method of quantifying cognitive workload in driving (e.g., Strayer et al., 2015). In the DRT participants are required to respond to a target stimulus occurring randomly every three to five sec while concurrently performing a primary task or tasks (e.g., driving and talking on a cell phone). The target can be presented in the visual or tactile modalities. Increased load in the primary task is found to increase the number of omissions (Castro, Cooper, & Strayer, 2016; Castro et al., 2019; Cooper, Castro, & Strayer, 2016).

The CPT is a detection task used to study sustained attention. However, it differs from the aforementioned tasks in that it routinely features distractor items. Participants are required to respond to a typically rare target symbol (e.g., Beale, Matthew, Oliver, & Corballis, 1978) that randomly appears in a continuous stream of non-target symbols, presented one at a time at a fixed rate. For example, a participant might be presented with a string of letters and instructed to only respond every time the target letter “X” is presented. This describes the classic CPT-X version, however there are many other variants that manipulate the complexity of the target stimuli (e.g., Rosvold et al., 1956; Garfinkel & Klee, 1983; Cornblatt, Lenzenweger, & Erlenmeyer-Limling, 1989), change the task modality to auditory (Earle-Boyer, Serper, Davidson, & Harvey, 1991), or alternate between auditory and visual modalities within the same task (Sanford & Turner, 1995). Increased omission rates in the CPT are observed due to a number of factors: longer task durations; low signal to noise ratio between targets and non-targets; and in participants with clinical diagnoses such as attention deficit hyperactivity disorder, traumatic brain injury, and a history of cerebrovascular accidents (for reviews, see Ballard, 2001; Riccio, Reynolds, Lowe, & Moore, 2001).

The distraction paradigm (e.g., Schröger & Wolff, 1998) is used to investigate how selective attention systems are able to filter extraneous changes in stimulus features. In classic versions a task-irrelevant auditory odd-ball sequence accompanies a primary choice task (e.g., numerical parity; Andrés Parmentier, & Escera, 2006). Participants are instructed to focus exclusively on the primary task. The occurrence of task-irrelevant oddballs has been found to increase omission rates in the primary task (e.g., Domínguez-Borràs et al., 2009; Gumenyuk et al., 2005; Parmentier, Mayberry, & Elsley, 2010; Schröger, 1996; cf. Escera, Corral, & Yago, 2002; Muller-Gass, Stelmack, & Campbell, 2006). The n-back task is used to study sustained attention and working memory (Kirchner, 1958). Participants are presented with a sequence of stimuli, for example letters or pictures, and on each trial they must decide whether the current stimulus matches the stimulus n trials back. In n-back tasks, omissions are found to be more common than commission error responses (Meule, 2017), increase with increasing task difficulty (Meule, 2017), and have a strong genetic basis (Evans et al., 2018).

Interestingly, for both detection and choice tasks, there is evidence of changes in omission rates with particular manipulations. Howard et al. (2020) studied both a DRT with a shorter 2.5 sec response window and a modified version that required a choice rather than detection response, again with a 2.5 sec response window. They manipulated the difficulty of a primary object-motion tracking task, but only reported omission results for the detection task, which increased in the highest load condition. Castro et al. (2019) also examined a choice modification of the DRT and reported substantially reduced omissions relative to the standard DRT.

In all of the aforementioned tasks, increased omission rates are interpreted, at least in part, as being indicative of inattention – an inability to sustain attention and inhibit distraction. Interestingly though, even in the areas of research that occasionally feature omissions as a dependent variable, very little effort has gone into their further exploration. Any additional information omissions might contain about decision making processes, particularly when distracted, has been underutilised. Although omissions are often collapsed with commission errors (e.g., Mager et al., 2006; Miller et al., 2009), there is evidence that, at least in some instances, patterns of omission and commission errors can differ (Meule, 2016; Meule et al., 2012), suggesting that it is better to consider omissions separately and that there might be extra psychological insights to be gained by understanding their cause or causes.

Castro et al. (2019) incorporated an account of omissions into an evidence-accumulation model of the standard and choice versions of the DRT using a diffusion process with a single positive barrier (Heathcote, 2004), but they did not differentiate among different possible causes of omissions. Evans et al. (2018) and Howard et al. (2020) took account of design omissions associated with slow RT cut-offs in the data they examined by applying the approach for fitting censored[1] distributions proposed by Kendall and Stuart (1967; also Heathcote, 1996, and Ulrich & Miller, 1994) to modify the likelihood equations of a version of Brown and Heathcote's (2008) LBA) model proposed by Heathcote and Love (2012) that allows only positive accumulation rates. Ratcliff and Van Dongen (2011, see also Ratcliff & Strayer, 2014) fit the single-barrier diffusion model to PVT data and accounted for "lapses" (responses slower than 0.5 sec) through both censoring and intrinsic omissions. Intrinsic omissions are caused by negative evidence-accumulation rates due to Gaussian trial-to-trial variability, which can result in a failure to ever reach the positive barrier.

In the next section we combine these approaches to provide a more comprehensive account of omissions. We use Brown and Heathcote's (2008) original LBA, which allows negative accumulation rates due to Gaussian trial-to-trial variability, to account for intrinsic omissions. We use Kendall and Stuart's (1967) approach to account for design omissions through censoring. Finally, we use a version of the approach of Castro et al. (2019) to model contaminant omissions (see also Matzke, Curley, Gong & Heathcote, 2019).

A Comprehensive Cognitive Model of Omissions

Evidence-accumulation models (EAMs) have long offered a comprehensive solution for analysis of RT and accuracy data. EAMs describe the process of choice, assuming

---

[1] Censoring implies that the number of omissions is known, and this number is taken into account in Kendall and Stuart likelihood modification (Equation 32.36, p. 523). Censoring is to be distinguished from truncation, where the number of omissions is unknown, but the cut-off or cut-offs are known. Truncation requires a different modification of the likelihood, normalisation by the probability of non-omission. Because the number of omissions is not taken into account, the quality of estimates is reduced relative to censoring.

evidence favouring different choice options accumulates over time until a threshold is reached, and an associated response is triggered. EAMs are able to provide a comprehensive account of responses, including the probabilities of correct and error responses and the shapes of the corresponding RT distributions but, in most cases, omissions have been ignored. An additional benefit of EAMs is that they also allow quantification of latent psychological processes involved in decision making and response execution (e.g., Brown & Heathcote, 2008; Donkin & Brown, 2018; Leite & Ratcliff, 2010; Luce 1986). Exploring how different experimental manipulations modulate parameter estimates, or how parameter estimates differ between different populations is ultimately what enables insight into latent psychological processes and informs broader theoretical understandings. The use of EAMs with omissions data similarly elucidates the latent psychological processes that underpin the occurrence of omissions. They also utilise all the available information in the data (i.e., omission rate) – information that has previously been considered noise.

EAM parameters corresponding to different psychological processes include: the *rate* evidence is accumulated at, which can be broken down to the average rate over accumulators for each response (quantifying the urgency with which responses are made) and the difference between the rates for the accumulators that match and mismatch the stimulus (quantifying the quality of the evidence on which the decision is based); the *threshold* (b) amount of evidence required to trigger a response, which can be broken down into the average over accumulators (which quantifies response caution) and the difference between accumulator threshold (which quantifies response bias); and the *non-decision time* (t_{er}) required for processes such as stimulus encoding and response execution. For detection tasks a single threshold and accumulation process is assumed so only the rate, threshold, and non-decision parameters are relevant. Before we outline how these parameters are expressed and their underlying mathematical properties, it is useful to provide a review of the three types of omissions we have defined. The first two – intrinsic omissions (due to between-trial variation in accumulation rates making it impossible to ever reach the evidence threshold) and design omissions (due to a priori response windows that cause slow responses not to be recorded) – are determined by the same process that produces responses; and a third, contaminant omissions, allows for a cause unrelated to the response process.

In some EAMs, one or more of the aforementioned parameters is assumed to randomly vary between trials, with extra parameters corresponding to the level of variation. For the LBA model that we use here, the starting points (or equivalently thresholds) of evidence accumulation have a uniform distribution (between 0 and A), rates have a normal distribution (with mean v and standard deviation s_v), with both assumed independent over accumulators. As the rate distribution is unbounded the sampled rate for a trial can be negative with probability $\Phi\left(\frac{v}{s_v}\right)$, where $\Phi(x)$ is the integral from $-\infty$ to zero of a normal distribution with mean x and a unit standard deviation. In the case of the single-accumulator model for detection tasks (e.g., the DRT) this means that the probability of an intrinsic omission is $p_I = \Phi\left(\frac{v}{s_v}\right)$. In the case of the two-accumulator model for choice tasks (e.g., the distraction paradigm) the probability of an intrinsic omission is given by the probability that both matching and mismatching accumulator have negative rates at the same time, $p_I = \Phi\left(\frac{v_T}{s_{vT}}\right) \times \Phi\left(\frac{v_F}{s_{vF}}\right)$ where the subscript T indicates the matching accumulator and F the mismatching accumulator. Brown and Heathcote (2008) noted that in their fits to data in which omissions were rare or unreported p_I was negligible and so they did not take account of this possibility in the likelihood equations they used to obtain fits.

The probability of design omissions is determined by the response window and the LBA accumulator probability density function, $f(t \mid A, b, v, s_v)$, and cumulative distribution function, $F(t \mid A, b, v, s_v)$ (see Brown & Heathcote, 2008, for the corresponding equations), where t is the decision time (i.e., RT – t_{er}). In the cases we address here, the response window excluded only slow RTs greater than an upper cut-off, U. For the detection case the probability of a design omission is $p_D = 1 - F(U)$. For the choice case we first require the probability densities for correct, $f_c(t) = f\ (t \mid A_T\ , b_T, v_T, s_{vT})\big(1 - F(t \mid A_F, b_F, v_F, s_{vF})\big)$ and error, $f_E(t) = f(t\ |A_F, b_F, v_F, s_{vF})\ \big(1 - F\ (t \mid A_T, b_T, v_T, s_{vT})\big)$, responses. The probability of a design omission is then the sum of the probabilities that correct and error responses exceed $U, p_D = \int_U^\infty f_E(x)\ dx + \int_U^\infty f_c(x)\ dx$, where each probability is obtained by numerical integration.

Both the probability of design and intrinsic omissions, and hence the overall probability of a process omission, p_P, have different value for each cell of the experimental design that differ on one or more accumulator parameters (i.e., A, b, v, or s_v), so for design cell $i,\ p_{Pi} = p_{Ii} + (1 - p_{Ii})\ p_{Di}$. Note that no extra parameters need to be estimated to produce predictions about process contamination. The ability of the model to discriminate process from contaminant comes from the former being constrained by the observed responses (i.e., having to fit response probabilities and RT distributions). The contaminant component adds extra freedom to explain omissions that cannot easily be accommodated by the response process. We make the assumption that only the evidence accumulation process is affected by the factors that constitute the experimental design, and hence that the probability of omission due to contamination, p_C, applies to all design cells. Hence, the overall probability of omission for design cell i is $p_{Oi} = p_C + (1 - p_C)p_{Pi}$. Therefore, differences in omission rates between design cells are entirely accounted for by the evidence accumulation process, whereas some or all of the overall level of omissions is accounted for by contamination at the cost of requiring one extra parameter estimate. This set of assumptions helps to make it possible to identify the parameters of the model. The supplemental materials (available at https://osf.io/hb4dw/) provide the results of a series of parameter-recovery simulations. We evaluated the recovery of the three types of omission parameters on two dimensions: accuracy and the calibration of uncertainty. The detection response task performed well on both dimensions, whereas estimation in the distraction paradigm was accurate, but the uncertainty of the parameter estimates was slightly underestimated.

In the following two sections we assess the evidence for each of the model's three types of omissions, first in the simpler case of Castro et al.'s (2019) DRT data and then in more complex case of our new distraction-paradigm data. In the general discussion we consider alternative assumptions about how omissions are generated.

Omissions in the Detection Response Task

Castro et al. (2019) manipulated cognitive load by having 20 participants count backwards by 3 sec in one condition and not in another. These conditions were crossed with four other conditions, three of which required participants to perform versions of the DRT – either detecting a lower or higher intensity light in two and discriminating the lower versus higher intensity lights in a third choice-DRT condition. No extra task was performed in the fourth condition. Omissions were most frequent in the detection conditions, at an average of around 5%, so we analyse only those conditions, which form a 2 (stimulus: low vs. high intensity) by 2 (load: none vs. 3s) within-subjects design.

Omissions were more common with the secondary-task load (6%) than without (4.3%), $\chi^2(1) = 19.8$, $p < .001$. RT was also much slower with the load, by 0.16s, $\chi^2(1) = 1080$, $p < .001$, and slower for low than high intensity stimuli, by 0.02s, $\chi^2(1) = 37.9$, $p < .001$. These results suggest that RT and omission rates are related (i.e., both RT and omissions increase with load) but not entirely redundant (stimulus intensity affects RT but not omissions). Note that throughout this paper tests on manifest measures were carried out using linear mixed models (Bates, Maechler, Bolker, & Walker, 2015), assuming Gaussian error for the log of RT and a binomial model with a probit link function for binary measures, with Type III Wald test statistics in both cases (Fox & Weisberg, 2011).

We first fit seven different versions of the LBA-Omission (LBAO) model that differed in the parameters that were affected by the load factor, constituting combinations of one or more of the mean rate, rate standard deviation, and non-decision time parameters. All models assumed a common *A* parameter across all conditions, and we assumed different mean rate and rate standard deviation parameters for high and low stimuli. In order for the standard LBA model to be identifiable either a rate mean or standard deviation or threshold parameter must be fixed, with the choice of which being arbitrary as a simple re-parametrization makes the effect of fixing one exactly equivalent to fixing another (see Donkin, Brown, & Heathcote, 2009), and a common choice being a rate standard deviation parameter. However, in the LBAO model both rate parameters can affect intrinsic omissions, whereas the threshold does not, and so the equivalence does not hold. Hence, we choose to fix the threshold for identifiability. As low and high stimulus conditions had almost entirely separate parameters, we fixed the threshold at 1 in both. We fit the models using the software and Bayesian methods described in Heathcote et al. (2019) with relatively non-informative priors (see supplementary materials for details, and files at https://osf.io/hb4dw/ for all of the data and R code for all of the model fits described in this paper).

Detection Response Task Modeling Results

Table 1 reports model-selection results based on the Deviance Information Criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002), where smaller values indicate a better trade-off between goodness-of-fit and model complexity. The full model is selected with the largest load effect being on mean rates and the least on non-decision time. This pattern is consistent with Castro et al. (2019) DIC model-selection results, except that they fixed the moment-to-moment variability parameter of the single-barrier diffusion model (which does not include a trial-to-trial variability parameter) to make it identifiable, and we found that load affected the threshold parameter.

Table 1.

DIC model selection results (better models have smaller values) for 7 models constituting all possible combinations of load effects on LBA rate means (v) and standard deviations (s_v) and non-decision time (t_{er}) parameters with the DIC for the best model, min(DIC) = - 1515, subtracted.

Model	v, sv, t_{er}	v, sv	v, t_{er}	sv, t_{er}	v	sv	t_{er}
DIC - min(DIC)	0	50	241	1150	373	2243	1389

Figure 1 plots average omissions and mean RT along with fits of the selected model. In the Bayesian framework uncertainty (indicated by 95% credible intervals) is indicated for the model rather than the data. The model provides a good account of the positively skewed RT distributions and the increase in RT with load as well as the percentage of omissions, although there is slight underestimation for low load stimuli. Figure 1 also plots individual omissions, showing that three participants had particularly high omission rates, and that these were almost entirely accounted for by the model's contaminant omission parameter. However, the figure makes it clear that for most other participants the process omissions are necessary. This is also clear from the fact that omissions vary as a function of the experimental manipulations, which cannot be explained by the contaminant parameter as it is assumed fixed across these conditions.

Figure 2 shows how the two types of process omission were estimated to vary over experimental conditions by the selected model. In every case, omissions appear to increase with load, consistent with the effect of load on the observed omission rates. To provide a test we calculated process omission rates for every posterior parameter sample and computed medians and Bayesian 95% credible intervals (Klauer, 2010; Matzke, Dolan, Batchelder, & Wagenmakers, 2015) on the posterior predictive distributions of differences over manipulations. Credible intervals (provided in square brackets below) were estimated by the range between the 2.5th and 97.5th percentiles of the resulting distribution. There was a main effect of load with both stimuli for intrinsic (low: 0.74% [0.42, 1.04]; high: 0.71% [0.48, 0.99]) and design (low: 1.18% [0.9, 1.41]; high: 1.24% [1.02, 1.24]) omissions. No main effects or interactions with the stimulus factor were supported except a slightly higher rate of design omissions for low than high stimuli with no load (0.21%, [0.04, 0.42]). Intrinsic omissions were less common than design omissions both without load (low: 0.26% [0.16, 0.34]; high: 0.18% [0.08, 0.25]) and more so with load (low: 0.69% [0.53, 0.84]; high: 0.7% [0.56, 0.83]).

We also examined posterior difference distributions to test effects on model parameters. As expected from the model selection results, and consistent with the idea that a secondary task drains cognitive capacity as reflected in accumulation rates (Castro et al., 2019), increased load decreased mean rates from 5.96 to 4.06 (a difference of 1.89 [1.78, 2.02]). Also consistent with their being allowed to vary with stimulus in the models and the standard interpretation of rates reflecting stimulus strength, mean rates were greater for high than low intensity stimuli (0.26 [0.18, 0.34]). Although there was not support for a main effect of stimulus on rate standard deviations there was a bigger load effect for low than high stimuli (0.18 [0.05, 0.32]). The main effect of load was caused by a decrease in rate standard deviations with load from 2.1 to 1.58 (0.52 [0.43, 0.61]), but the decrease was largely proportional to the mean, so the ratio of mean to standard deviation decreased only slightly with load (0.13 [-0.02, 0.27]).

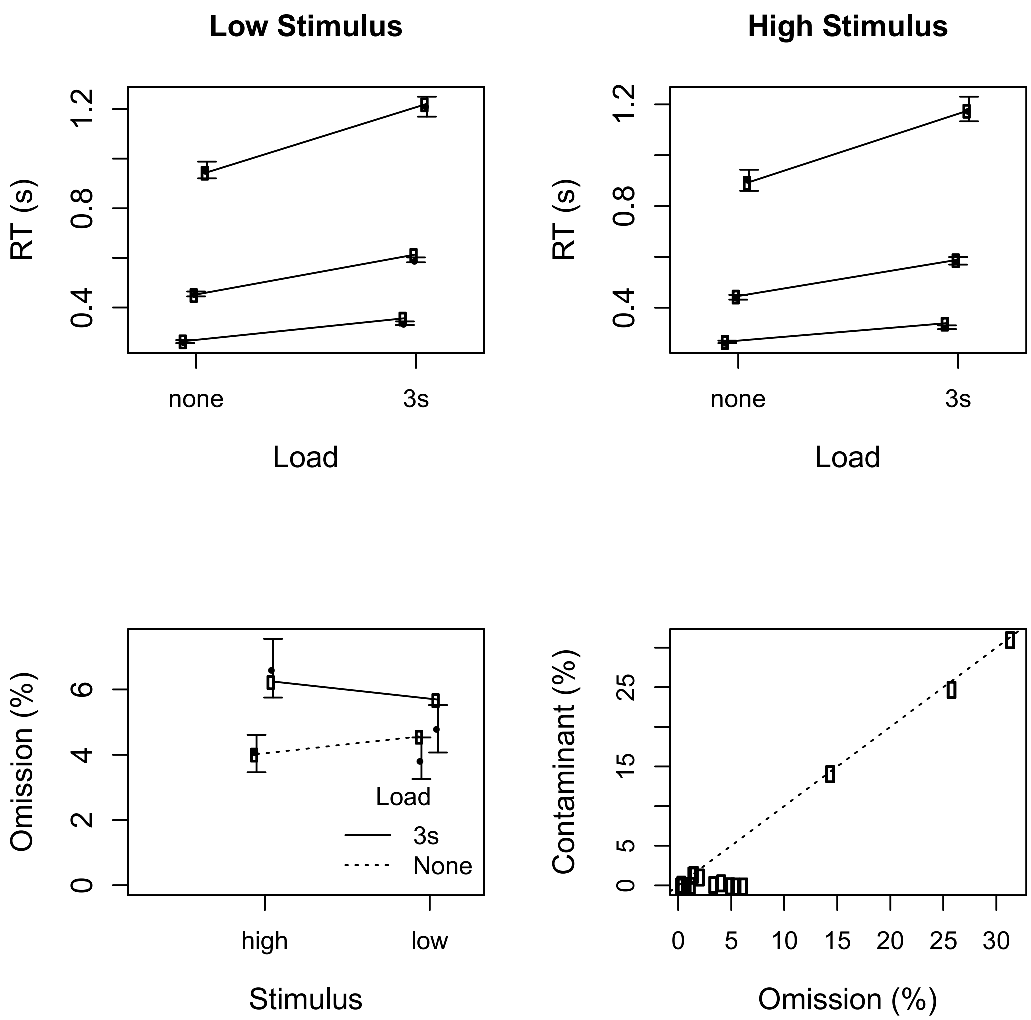

Figure 1. Top row: RT distributions (lower, middle, and upper lines and points are 10th, 50th, and 90th percentiles, respectively). Bottom left panel: Omission percentages data (open symbols), with model fits of the selected model in Table 1 as posterior predictive medians (solid points) and 95% credible intervals. Bottom right panel: Individual participant omission percentages against the posterior median of the contaminant omission parameters estimated by the model. Load conditions: none and 3s (i.e., count backwards by 3s).

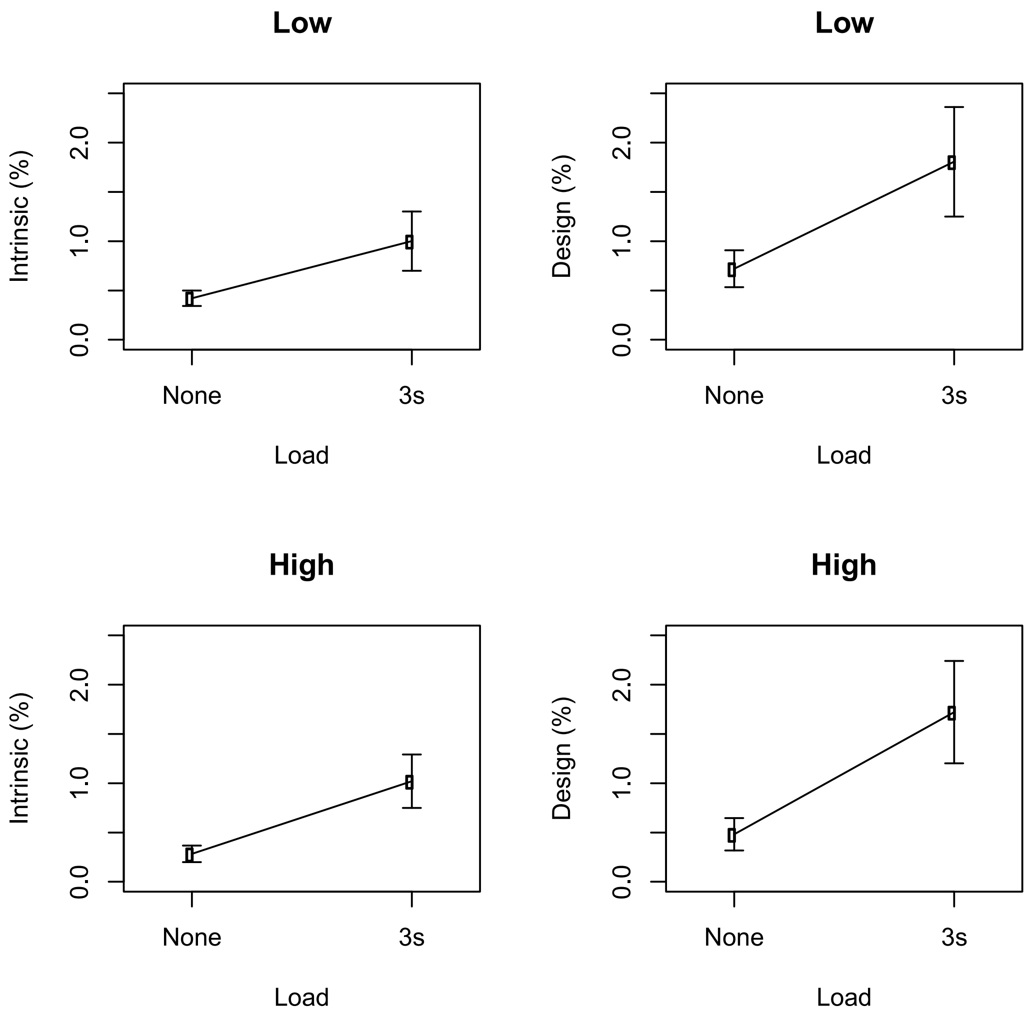

Figure 2. Median and 95% credible intervals for the posterior predictive distribution of process omissions computed for the model selected in Table 1. The left column displays intrinsic omissions and the right column design omissions with the upper row containing results for low stimulus and the lower row for high stimuli.

In light of the results just reported we fit two extra models. Because of the relatively weak intrinsic omission rates, we fit a model in which they are absent as the rate distribution is truncated to be positive (Heathcote & Love, 2012). Consistent with the observed small effects, this model, which was parameterized as for the winning model in Table 1, was preferred by the DIC by a small margin (32). However, as shown in supplementary materials (Figure S1), this model had an exaggerated under-prediction of omissions for low-intensity stimuli. Because three participants had much higher omission rates (15%, 26%, and 32%) that were almost entirely explained by the contaminant parameter (see Figure 1) we fit a model without the contaminant parameter (but otherwise the same as the model selected in Table 1) to see if it was necessary in the remaining participants. The support for the model with the contaminant parameter remained, with the model selected in Table 1 winning on DIC by 55, but the DIC difference was much smaller than for the fit including all participants (369). Further, without contaminant omissions the misfit to observed omission rates was very marked (see Figure S1 in supplementary materials) suggesting that it was impossible to increase the level of process omissions sufficiently to compensate for contaminant omissions.

Detection Response Task Discussion

The LBAO model was able to provide a quite accurate representation of all aspects of performance in the DRT, ranging from RT distribution to omission rates. The model even accurately captured participants with high omission rates who might ordinarily be excluded from analyses due to being deemed non-compliant with experimental instructions. Given that participants had to simultaneously perform the tracking task and counting backwards by 3 sec in one condition, higher omission rates do not seem unreasonable, and may simply represent part of the normal continuum of methods for coping with an attention demanding task. The LBAO model was able to provide a good fit to data from these participants (see Figure S2 in supplementary materials for individual plots), including the RT data for the responses they made. This provides evidence that the LBAO was able to successfully separate out the effects of the contaminant process.

The LBAO model suggested that process omissions were increased by the distraction caused by a secondary-task load (counting backward by 3). This was true of both intrinsic omissions, occurring because the stimulus did not cause any positive evidence accumulation towards the response threshold, and of design omissions, where the positive accumulation was so weak that it did not result in threshold crossing in under the 3 sec time limit dictated by the ISO standard. This result is striking because even the slow 90th percentile of RT was little more than 1 sec on average (see Figure 1), indicating that the LBAO produced distributions with a long thin slow tail that contains non-negligible probability mass at quite long RTs.

The incidence of intrinsic omissions was much less, and indeed there was some evidence that they were not a necessary part of a parsimonious model, although their exclusion led to some noticeable misfit in omission rates.

In the next section we investigate whether the LBAO model can also provide a good description of choice behaviour in the distraction paradigm and how the relative levels of all three types of omissions differ in this case.

Omissions in the Distraction Paradigm

In the distraction paradigm, as well as examining the effect of distraction, we also explored differences in omission behaviour due to aging, comparing the performance of a young and healthy older groups. Participants made a choice response classifying tones as longer or shorter. On a minority of “deviant” trials the task irrelevant tone pitch could either be higher or lower than on the majority of trials, causing an oddball distraction effect. As reviewed in the opening section of this paper, the occurrence of a deviant can increase omission rates, and is known to prolong RT and decrease accuracy rates (Schröger & Wolff, 1998) in the primary task. In light of evidence that distraction also affects performance on the trial after a deviant occurs, we conducted a 2 x 2 x 5 mixed model analysis with the between subject factor of age (young vs. old), and the within subject factors of stimulus (long vs. short) and deviance (S = standard after standard, HD = high deviant after standard, LD = low deviant after standard, SAHD = standard after high deviant, and SALD = standard after low deviant). Before reviewing our EAM results, below we provide an overview of the methodology used and results from traditional analysis techniques.

Method

Participants

Participants were 34 younger (22 females, 18-38 years, *M* = 23 years, *SD* = 4.85 years*)* volunteer community members and undergraduate students from the University of Newcastle, and 23 older (12 females, 59-74 years, *M* = 67 years, *SD* = 4.02 years) volunteer community members.

Apparatus

Assessments of hearing thresholds were conducted using an Earscan 3 ES3S pure tone audiometer in line with the American Speech-Language-Hearing Association Guidelines for Manual Pure-Tone Threshold Audiometry (2005). Experiments were delivered using Presentation© version X by a standard PC running Windows XP, on a 27inch LED monitor (60Hz). Sound stimuli were delivered binaurally via Sennheiser HD 280 professional headphones.

Stimuli

Sounds were equiprobable short (0.1s) and long (0.25s) pure tones with .005s rise/fall times. Tone frequency (i.e., pitch) was manipulated to produce an auditory oddball sequence. Regular tones had a frequency of 700 Hz (p = .75), while rare tones were either of a lower (613 Hz; p = .125) or higher (1560 Hz; p = .125) frequency. All frequencies were presented equiprobably across short and long tones. All tones were delivered at an intensity of 75dB SPL.

Task

Participants completed an auditory duration discrimination task. Specifically, they were asked to indicate whether presented tones were short or long in duration by pressing a left or right button on a custom built two-button response terminal. Response mapping was counterbalanced between participants with the hand-to-button-press requirements (i.e., left index finger to left button and right index finger to right button) explained and demonstrated to participants prior to task onset. Participants completed 800 duration discrimination trials, split equally across four blocks. Blocks were separated by a one-minute rest break. Tones had an inter-stimulus-interval of 1.6 sec. Participants were instructed to respond as quickly and accurately to tones as possible.

Procedure

Participants attended two testing sessions, each approximately two hours in duration. In session one, participants completed assessment of suitability for inclusion based on mental health history, as well as several other psychometric measures for use in another study (Todd et al., in preparation). In session two, participants completed the duration discrimination task. Prior to the beginning of the task participants were fitted with a 64 channel ActiveTwo Biosemi EEG system to record continuous electroencephalogram (EEG) data during completion of the duration discrimination task, also for use in another study (Todd et al., in preparation). Finally, participants were fitted with headphones and asked to stay as still as possible while they followed instructions presented on the screen and completed the duration discrimination task.

Results

The first 3 trials of every block were removed from analysis as is standard with most distraction paradigm tasks. Responses faster than 0.2 sec were then removed from analysis, which reflected 0.78% of responses by old and 0.075% of responses by young participants. The programmed response window closed at 1.45 sec, however due to computer timing jitter, eight responses slightly greater than 1.45 sec duration were identified and removed from analysis. Of these, five were for old (0.03%) and three for young (0.01%) participants.

Mean RT was faster for short than long stimuli (0.616 sec vs. 0.633 sec), $\chi^2(1) = 96$, $p < .001$, and was clearly shorter for standard trials than other trials (S: 0.601 sec, HD: 0.645 sec, LD: 0.651 sec, SAHD: 0.656 sec, SALD: 0.635 sec), $\chi^2(4) = 925$, $p < .001$. Although the main effect of age just failed conventional significance, the interaction of age and stimulus was significant; older participants responded 0.044 sec slower to short stimuli and 0.035 sec slower to long stimuli, $\chi^2(1) = 13.4$, $p < .001$.

Accuracy was greater for short than long stimuli (90% vs. 88%), $\chi^2(1) = 41$, $p < .001$, highest in the standard condition, and slightly lower in the other deviance conditions except in the standard-after-high-deviant condition where it was markedly lower (S = 91%, HD = 78%, LD = 90%, SAHD = 89%, SALD = 90%), $\chi^2(4) = 937$, $p < .001$. The latter effect was driven by an interaction between tone length and frequency whereby long duration high deviants were perceived as short duration tones (accuracy 61%), $\chi^2(3) = 899$, $p < .001$. Older participants were less accurate on average than younger participants (82% vs. 93%), $\chi^2(1) = 19.7$, $p < .001$. The misperception effect was also stronger, with accuracy for long duration high deviants less than chance for older participants (45%) but better for younger participants (72%), $\chi^2(1) = 49.6$, $p < .001$.

Omissions were more common for older than younger participants (6.4% vs. 2.2%), $\chi^2(1) = 7.3$, $p = .007$, and for short than long stimuli (4.2% vs. 3.6%), $\chi^2(1) = 15.7$, $p < .001$.

They were most common following the misperceived high-deviant in the standard-after-high-deviant condition, $\chi^2(4) = 49.7$, $p < .001$, particularly for young participants, $\chi^2(4) = 23.8$, $p < .001$. Detailed patterns are shown along with model fits below.

Distraction Paradigm Modeling Results

We fit young and old participants separately, and as with the DRT we fixed threshold parameters in order to make the LBAO model identifiable. We reasoned that thresholds could mediate a response bias (i.e., long and short response accumulators could have different thresholds) and that thresholds might differ on trials after deviants relative to other trials (as participants could easily detect the occurrence of a deviant and then had sufficient time before the next trial to adjust their threshold). As the most complex model we considered had different rate parameters for all stimuli, we fixed the threshold at 1 for the short accumulator for all conditions and estimated separate long accumulator thresholds for trials after a deviant and after a standard trial to allow for differential response bias.

In light of the strong differences among the 10 stimulus and deviance conditions we allowed separate mean rate parameters for each, resulting in a total of 20 estimated parameters, 10 for matching accumulators and 10 for mismatching accumulators. Given their critical role in determining both design and intrinsic omissions, we explored three models with progressively more complex rate standard deviations, either only differing between matching and mismatching accumulators (2 estimated parameters), also differing with the deviance factor (10 estimated parameters), or also differing with deviance and stimulus factors (20 estimated parameters). We fully crossed these three cases with three types of omission model: the full omissions model allowing for contaminant and process (design and intrinsic) omissions and models removing one mechanism, a second with only contaminant and design mechanisms to test the necessity of our new proposal of an intrinsic mechanism, and a third removing the contaminant mechanism to test the efficacy of the process mechanisms in isolation. All nine resulting models we assumed the same value of non-decision time and start-point noise for each condition.

Table 2 shows that for every combination of contaminant mechanism, the most complex model of rate standard deviations was preferred. Older participants showed the same pattern of results as for the DRT, with support for design and contaminant omissions but not for intrinsic omissions, whereas younger participants displayed support for design and intrinsic but not contaminant omissions.

Table 2. DIC minus minimum DIC separately for young (minimum DIC = -19156) and old (minimum DIC = 2744) participants for the 9 models crossing the three combinations of omission mechanisms with the three parameterizations of rate standard deviations (M = match vs. mismatch, D = deviance and S = stimulus factors). N = number of estimated parameters.

Omissions	Full Omission Model			Design & Intrinsic			Design & Contaminant		
Rate SD	M	M,D	M,D,S	M	M,D	M,D,S	M	M,D	M,D,S
N	27	35	45	26	34	44	27	35	45
Old	1102	606	44	1320	832	243	1025	505	0
Young	1047	496	76	952	438	0	1065	651	214

Figure 3 shows the fit of the full omission model. Although its account of responses (i.e., accuracy and RT distributions) is good, it under-estimates omissions by a constant amount across all conditions, particularly for older participants and short stimuli. We speculated that the perceptual interaction between tone length and frequency might be the cause, and so fit a model in which the probability of contaminant omissions, which modulates the overall level of omissions across the different deviance conditions, differed for long and short stimuli. This model was preferred by DIC, with the same type of models winning by an extra 192 and 128 units for old and young participants, respectively. Although the fit to omissions was better misfit remained (see Figure S3 in supplementary materials). Given the fact that allowing contaminant omissions to differ as a function of experimental manipulations blurs the boundary between them and decision-process related omissions, we chose to focus further analysis on the model with fits depicted in Figure 3.

The left column of Figure 4 plots the individual participant omission percentages against the model's contaminant estimates. In contrast to the DRT data, while one of the higher omission rate cases was predominantly explained by contamination, two were not. The remaining columns of Figure 4 focus on the model's ability to fit the pattern of omission differences across deviant conditions by subtracting out the standard condition. As the contaminant process adds a constant to all conditions it does not play a role in this explanation. The figure shows that the model's process omission components are generally quite good at accommodating the pattern of differences, with the exception of the underestimation for old participants with long stimuli in the high deviant condition and for young participants for short stimuli in the standard after high deviant condition.

Figure 5 plots a breakdown of process omission estimates for correct and error responses. Intrinsic omissions are notably rare for older participants, and although generally low for younger participants there are noticeable elevations in some deviant conditions. This is consistent with the DIC support for the inclusion of intrinsic omissions in Table 2. However, the large credible intervals indicate that the elevation in intrinsic omissions applies to only some participants. Design omissions are elevated for older participants, consistent with their generally slower responses.

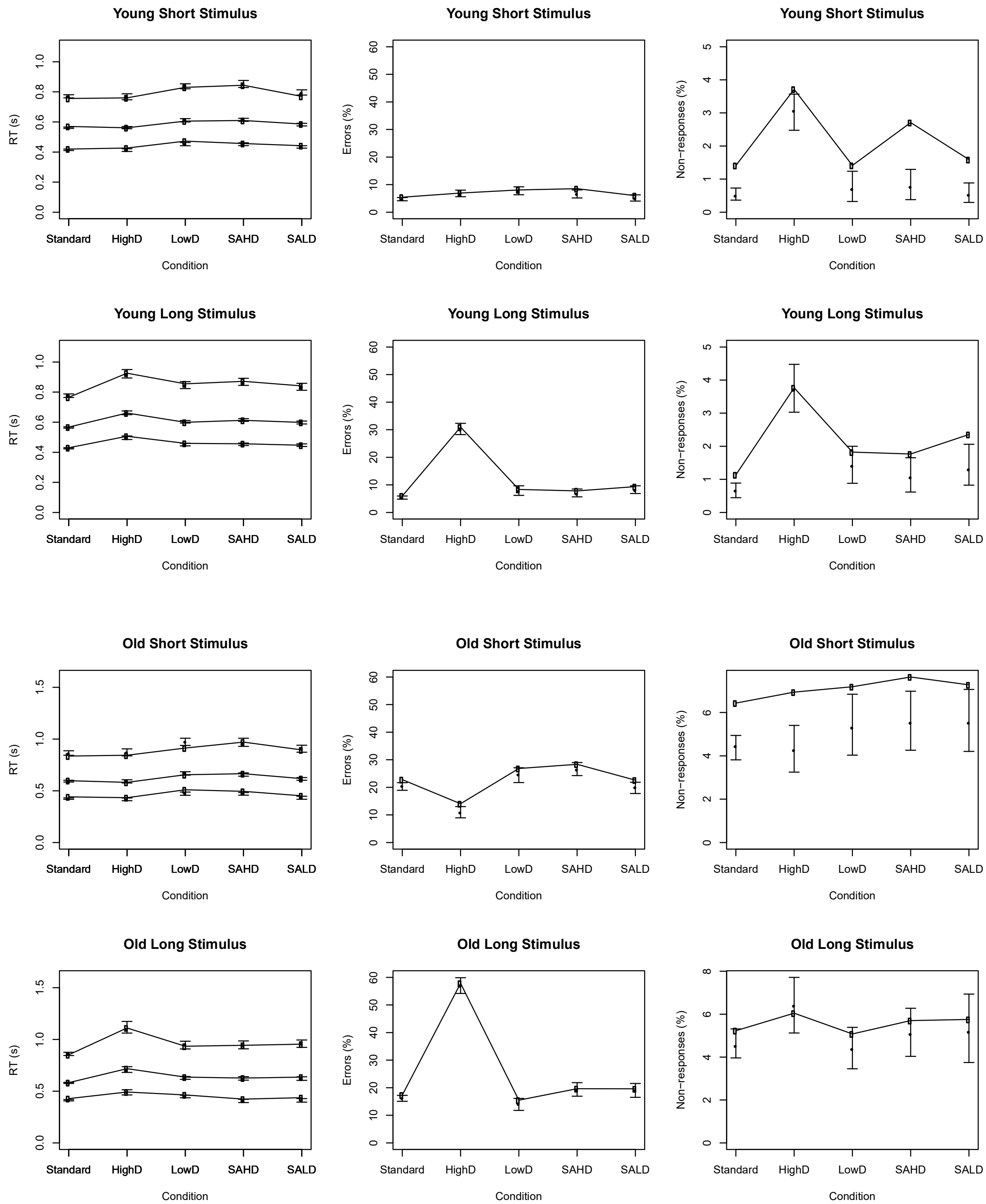

Figure 3. Fit of the most flexible intrinsic and contaminant model in Table 2. The three lines in the left column are the 10^{th}, 50^{th} and 90^{th} percentiles of RT distributions. Data are open circles joined by lines. Fits are shown with 95% credible intervals. HighD = High Deviant, LowD = Low Deviant, SAHD = Standard After High Deviant, SALD = standard after low deviant.

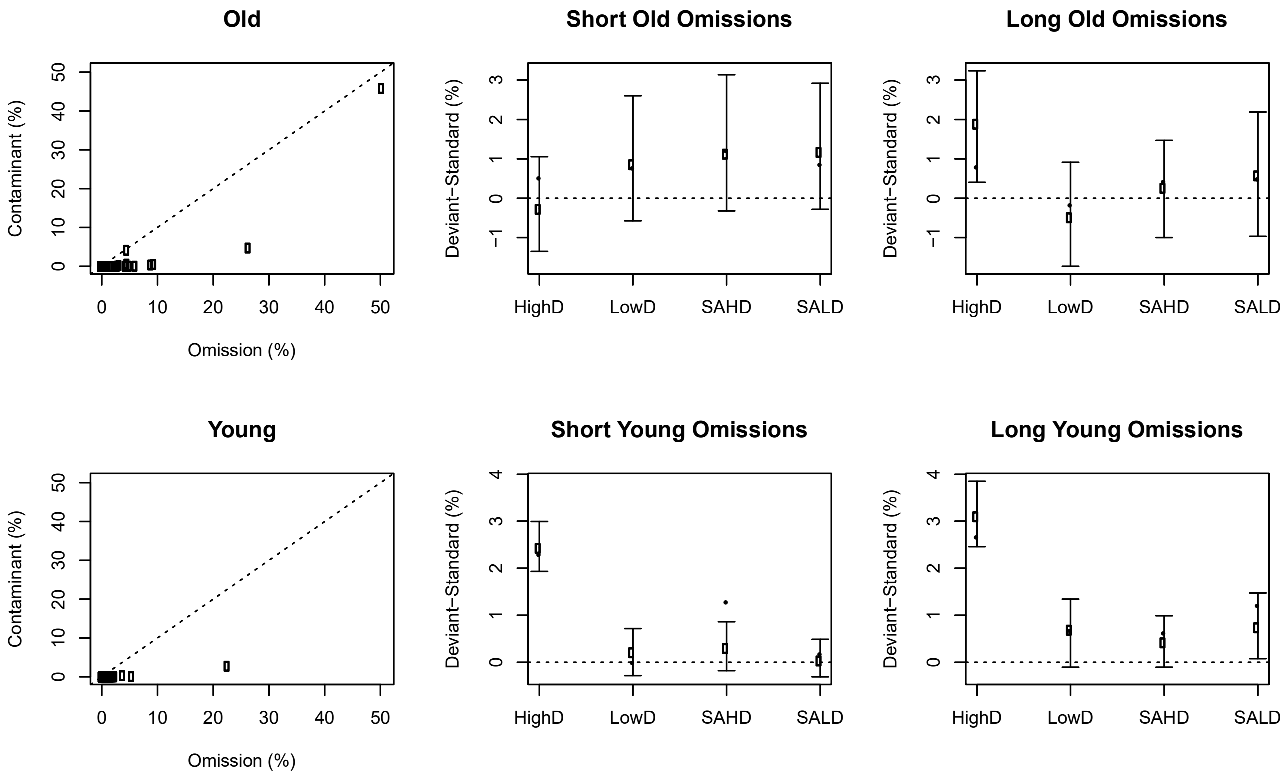

Figure 4. Parameters and fit of the most flexible intrinsic and contaminant model in Table 2. The left column shows individual participant omission percentages against the posterior median of the contaminant omission parameters estimated by the model. The middle and right columns show differences between the standard and deviant conditions omission percentages for the data (solid symbols) and the model (open symbols with 95% credible intervals). HighD = High Deviant, LowD = Low Deviant, SAHD = Standard After High Deviant, SALD = standard after low deviant.

Detailed parameter estimates are found in supplementary materials (see Figures S4, S5, and S6) and we summarize the most relevant aspects here. On average over accumulators mean rates were higher for short than long stimuli but did not vary greatly over deviance conditions. The difference between mean rates for the matching and mismatching accumulators also did not vary much over stimulus and deviance conditions. The only exception to this trend is for a sharp decrease in the high deviant condition for long stimuli, particularly for older participants. This reflects the marked tendency for old participants to classify long high deviant stimuli as short. Rate standard deviations averaged over accumulators did not differ much

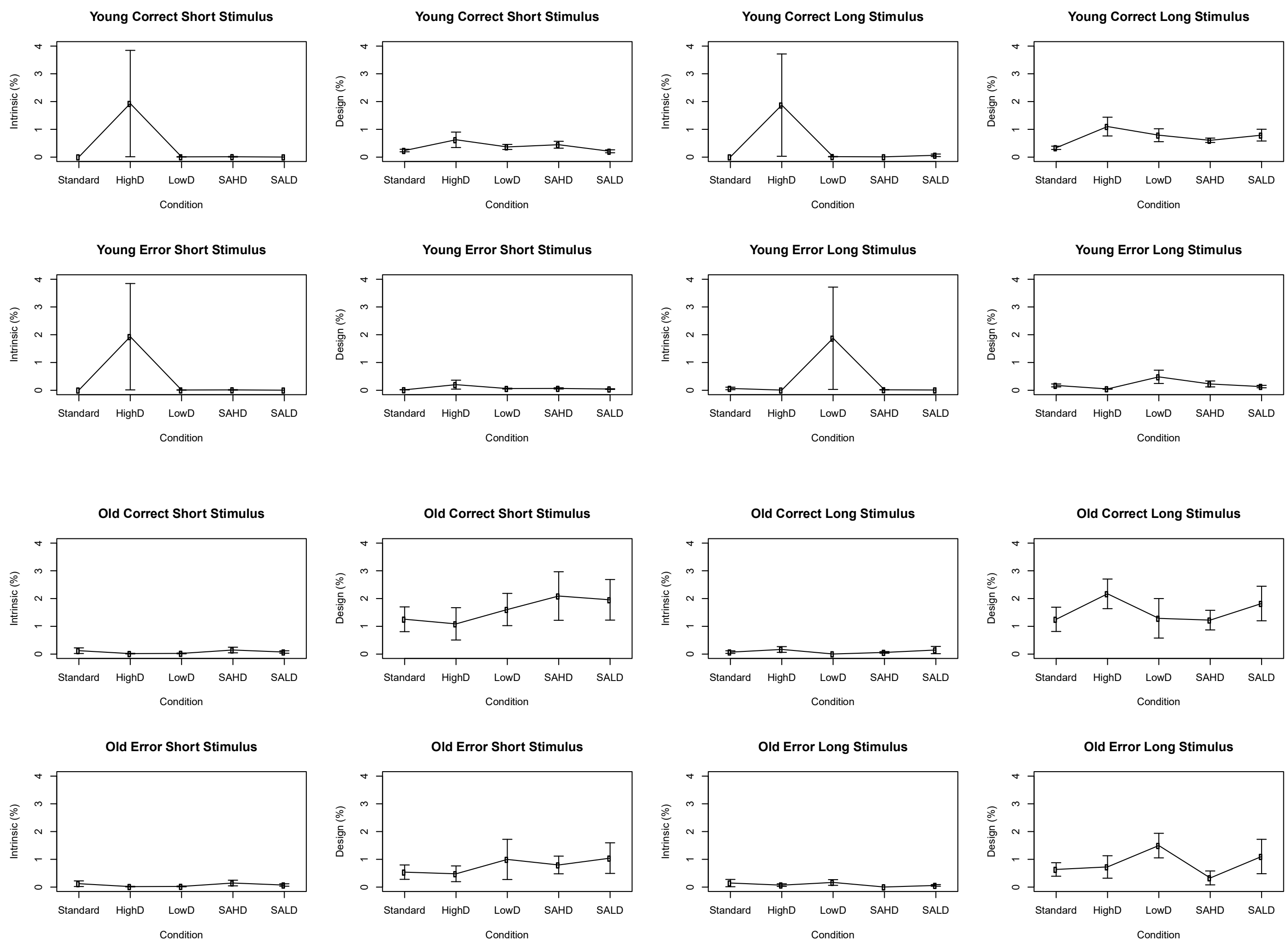
Young Correct Short Stimulus
Young Correct Short Stimulus
Young Correct Long Stimulus
Young Correct Long Stimulus
Young Error Short Stimulus
Young Error Short Stimulus
Young Error Long Stimulus
Young Error Long Stimulus
Old Correct Short Stimulus
Old Correct Short Stimulus
Old Correct Long Stimulus
Old Correct Long Stimulus
Old Error Short Stimulus
Old Error Short Stimulus
Old Error Long Stimulus
Old Error Long Stimulus
Intrinsic (%)
Design (%)
Condition
Standard
HighD
LowD
SAHD
SALD

Figure 5. Process omissions with 95% credible intervals for the most flexible intrinsic and contaminant model in Table 2 for correct and error responses in each condition. HighD = High Deviant, LowD = Low Deviant, SAHD = Standard After High Deviant, SALD = standard after low deviant.

between conditions. As is usually found with the LBA, the mismatch accumulator had a greater standard deviation than the match accumulator, with this difference being greater for short than long stimuli. Again, for the difficult high deviant condition things differed, with no difference for short stimuli and a reversal for long stimuli. As also shown in supplementary materials, participants were generally biased towards short responses. Consistent with all models allowing thresholds for each accumulator to differ on trials following a standard and following a deviant, the bias for young participants was greater after a deviant, whereas for old participants it was greater after a standard.

General Discussion

We proposed a variant of Brown and Heathcote's (2008) LBA evidence-accumulation model in which participants fail to make a response due to factors related to the decision process or to a contamination. We defined three types of omission: the first two – intrinsic omissions (due to between-trial variation in accumulation rates making it impossible to ever reach the evidence threshold) and design omissions (due to a priori response windows that cause slow responses not to be recorded) – are determined by the same process that produces responses; the third – contaminant omissions – allows for a cause unrelated to the response process. Because of the latter stipulation, contaminant omissions were assumed to be unrelated to the factors affecting process omissions, and so were fixed to be constant across experimental condition that we examined.

We fit this LBA Omission (LBAO) model to both detection data from a DRT task (Castro et al., 2019) and to new choice response data from a distraction paradigm task (Schröger & Wolff, 1998). Both paradigms manipulated experimental factors that were likely to cause omissions through causing distraction. In Castro et al. this was the load from other tasks that had to be performed in parallel with the detection responses. In the distraction paradigm this was relatively rare and task-irrelevant variations in an otherwise repetitive auditory stream. In both data sets responses longer than a slow cut off (3 sec in the DRT and 1.45 sec in the distraction paradigm) were not recorded. Also in both paradigms, a few participants had unusually high omission rates, including cases in which almost half of the response were omitted. Commonly, the latter participants are classified as non-compliant with experimental instructions and excluded from analysis. However, in the present study we included these participants in our analysis in order to see if the LBAO model was able to accommodate such a wide variation of omission rates.

We found that the LBAO model displayed varying degrees of success. It provided a quite accurate account of the detection data, both in terms of the distribution of RTs and the probability of omissions, with both being elevated under conditions of increased load. While the unusually large omission rates displayed by three participants were predominantly accounted for by the contaminant process, variation in omissions rates for the remaining 17 participants was related to process omissions. Design omissions played a more prominent role than intrinsic omissions, however, both were necessary to provide an accurate account. Both were also elevated under load, consistent with their common cause lying in the rate of evidence accumulation. The effect of design omissions is perhaps surprising – the associated slow cut off was more than twice as long as the 90th percentile of observed RTs (see Figure 1) and histograms of individual participants' RT distributions did not reveal any obvious signs of slow responses being cut off. This result suggests that it can be hard to detect design omissions by a simple visual inspection of the RT distribution (Ulrich & Miller, 1994), particularly when they occur at relatively low levels.

The LBAO model was less successful with the choice data from the distraction paradigm. It did provide a good account of choice accuracy and RT distributions in all conditions. This is the first demonstration that EAMs in general, and an LBA-type model in particular, can be successful and informative with data from the distraction paradigm. The

LBAO model also captured the high error rates and slowing apparent when long stimuli were accompanied by a high deviant. This was apparently due to a tendency for higher frequency deviants with a long duration to be perceived as shorter than they really were. Interactions between tone length and frequency, and a greater susceptibility to this effect with increasing age has been observed in a prior distraction paradigm study (Mager et al., 2005). This type of effect has previously gone unexplained, and so our results point to new avenues for further applications of EAM analysis within distraction paradigm data.

The LBAO model, did, however, have a tendency to underestimate overall omission rates by a small amount that was approximately the same amount for all levels of the deviance factor. This global tendency was most marked for older participants with short choice stimuli. In contrast, the model’s account of differences in omission rates among the different levels of the deviance factor was quite good. The latter account, which relies only on omissions related to the evidence-accumulation process, suggests that intrinsic omissions were essentially absent for older participants, and played a predominant role only for some younger participants that were susceptible to misperception of long high deviant stimuli. Design omissions played a more generally prominent role, being present in all conditions. These were most pronounced in older participants, perhaps less surprisingly, given that the slow cut off was less than twice the 90th RT percentile (see Figure 3) and older participants had comparatively slower RTs.

Overall, our results suggest that omission rates can provide information relevant to the psychological processes related to distraction. Further, we have demonstrated that it is possible to integrate the information provided by omissions with the information provided by accuracy and RT through an EAM approach. It is, however, important to recognise that not all of the potential causes of omission always play a prominent role. Our results clearly support consideration of design omissions through determining the proportion of responses that fall outside of response windows (see also Evans et al., 2018; Howard et al., 2020). Support for intrinsic omissions was weaker, suggesting that further work is required to extend investigation into the factors that determine their occurrence with detection responses (see Ratcliff & Van Dongen, 2011; Ratcliff & Strayer, 2014) versus choice responses.

The role played by contaminant omissions is less clear. Our work suggests that they are certainly necessary to explain the occurrence of high omission rates in some participants. However, the account provided by our specific assumption that they occur entirely independently of the decision process and experimental manipulations of distraction may be problematic, at least in choice paradigms. One possibility is that the passage of time itself plays some role. Recently, Hawkins and Heathcote (in press) proposed the addition of a "timing" accumulator (Simen, Vlasov, & Papadakis, 2016) to an evidence-accumulation race architecture, so that responses can be based on the passage of time as well as on the accumulation of evidence. In their proposal, this enables a time-out criterion with an associated guessed response. A possible extension of Hawkins and Heathcote's proposition to omissions is that a time-out might also sometimes result in a failure to respond.

Finally, the LBAO adds to the body of EAM research that seeks to further constrain models by including omissions data (Evans, Dutilh, Wagenmakers, & van der Maas, 2020; Evans et al., 2018; Howard et al., 2020; Ratcliff & Rouder, 1998). We conclude that omissions provide a theoretically sensible constraint, particularly in paradigms that examine distraction. Future studies might refine the methods we have developed here, and even potentially combine them with other approaches to providing extra constraint, such as taking account of the effects of the passage of time.

Acknowledgments:

This research was supported in part by the National Science Foundation Graduate Research Fellowship Program, the National Health and Medical Research Council Australia, the AAA Foundation for Traffic Safety, the Veni Grant (451-15-010) from the Netherlands Organization for Scientific Research, the Australian Research Council (Grant DP160101891), a Competitive Evaluation Research Agreement (CERA247) from The Defence Science and Technology Group (DST) of the Defence Science Institute of Australia, and an Australian Postgraduate Award Scholarship from the Department of Education and Training.

Open Practices Statement:

All data and code used in this manuscript are available at https://osf.io/hb4dw/. No experiments featured in this manuscript were preregistered.

References

Andrés, P., Parmentier, F. B. R., & Escera, C. (2006). The effect of age on involuntary capture of attention by irrelevant sounds: A test of the frontal hypothesis of aging. *Neuropsychologia, 44*(12), 2564-2568. doi:https://doi.org/10.1016/j.neuropsychologia.2006.05.005

Ballard, J. C. (2001). Assessing attention: Comparison of response-inhibition and traditional continuous performance tests. *Journal of Clinical and Experimental Neuropsychology, 23*(3), 331-350. doi:10.1076/jcen.23.3.331.1188

Beale, I. L., Matthew, P. J., Oliver, S., & Corballis, M. C. (1987). Performance of disabled and normal readers on the continuous performance test. *Journal of Abnormal Child Psychology, 15*(2), 229-238. doi:10.1007/BF00916351

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153-178. doi:https://doi.org/10.1016/j.cogpsych.2007.12.002

Castro, S., Cooper, J., & Strayer, D. (2016). Validating two assessment strategies for visual and cognitive load in a simulated driving task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 60*(1), 1899-1903. doi:10.1177/1541931213601432

Castro, S. C., Strayer, D. L., Matzke, D., & Heathcote, A. (2019). Cognitive workload measurement and modeling under divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 45*(6), 826–839. doi:https://doi.org/10.1037/xhp0000638

Cooper, J. M., Castro, S. C., & Strayer, D. L. (2016). Extending the detection response task to simultaneously measure cognitive and visual task demands. *Proceedings of the*

Human Factors and Ergonomics Society Annual Meeting, 60(1), 1962-1966. doi:10.1177/1541931213601447

Cornblatt, B. A., Lenzenweger, M. F., & Erlenmeyer-Kimling, L. (1989). The continuous performance test, identical pairs version: II. Contrasting attentional profiles in schizophrenic and depressed patients. *Psychiatry Research, 29*(1), 65-85. doi:https://doi.org/10.1016/0165-1781(89)90188-1

Cortiñas, M., Corral, M.-J., Garrido, G., Garolera, M., Pajares, M., & Escera, C. (2008). Reduced novelty-P3 associated with increased behavioral distractibility in schizophrenia. *Biological Psychology, 78*(3), 253-260. doi:https://doi.org/10.1016/j.biopsycho.2008.03.011

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers, 17*(6), 652-655. doi:10.3758/BF03200977

Domínguez-Borràs, J., Trautmann, S.-A., Erhard, P., Fehr, T., Herrmann, M., & Escera, C. (2009). Emotional context enhances auditory novelty processing in superior temporal gyrus. *Cerebral Cortex, 19*(7), 1521-1529. doi:10.1093/cercor/bhn188

Donkin, C., & Brown, S. D. (2018). Response times and decision-making. In E.-J. Wagenmakers & J.T. Wixted (Eds.), *Stevens' handbook of experimental psychology and cognitive neuroscience, Volume five: Methodology (4th ed.)*, pp. 349–377. John Wiley & Sons, Inc. doi: https://doi.org/10.1002/9781119170174.epcn509

Donkin, C., Brown, S. D., & Heathcote, A. (2009). The overconstraint of response time models: Rethinking the scaling problem. *Psychonomic Bulletin & Review, 16*(6), 1129-1135. doi:10.3758/PBR.16.6.1129

Earle-Boyer, E. A., Serper, M. R., Davidson, M., & Harvey, P. D. (1991). Continuous performance tests in schizophrenic patients: Stimulus and medication effects on performance. *Psychiatry Research, 37*(1), 47-56. doi:https://doi.org/10.1016/0165-1781(91)90105-X

Escera, C., Corral, M.-J., & Yago, E. (2002). An electrophysiological and behavioral investigation of involuntary attention towards auditory frequency, duration and intensity changes. *Cognitive Brain Research, 14*(3), 325-332. doi:https://doi.org/10.1016/S0926-6410(02)00135-0

Evans, N. J., Dutilh, G., Wagenmakers, E.-J., & van der Maas, H. L. J. (2020). Double responding: A new constraint for models of speeded decision making. *Cognitive Psychology, 121*, 101292. doi:https://doi.org/10.1016/j.cogpsych.2020.101292

Evans, N. J., Steyvers, M., & Brown, S. D. (2018). Modeling the covariance structure of complex datasets using cognitive models: An application to individual differences and the heritability of cognitive ability. *Cognitive Science, 42*(6), 1925-1944. doi:10.1111/cogs.12627

Fox, J., & Weisberg, S. (2018). *An R companion to applied regression*. Sage publications.

Garfinkel, B. D., & Klee, S. H. (1983). A computerized assessment battery for attention deficits. *Psychiatric Hospital, 14(3)*, 163–166.

Gumenyuk, V., Korzyukov, O., Escera, C., Hämäläinen, M., Huotilainen, M., Häyrinen, T., . . . Alho, K. (2005). Electrophysiological evidence of enhanced distractibility in ADHD children. *Neuroscience Letters, 374*(3), 212-217. doi:https://doi.org/10.1016/j.neulet.2004.10.081

Hawkins, G., & Heathcote, A. (in press). Racing against the clock: Evidence-based vs. time-based decisions. *Psychological Review*.

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Behavior Research Methods, Instruments, & Computers, 28*(3), 427-445. doi:10.3758/BF03200523

Heathcote, A. (2004). Fitting Wald and ex-Wald distributions to response time data: An example using functions for the S-PLUS package. *Behavior Research Methods, Instruments, & Computers, 36*(4), 678-694. doi:10.3758/BF03206550

Heathcote, A., Lin, Y.-S., Reynolds, A., Strickland, L., Gretton, M., & Matzke, D. (2019). Dynamic models of choice. *Behavior Research Methods, 51*(2), 961-985. doi:10.3758/s13428-018-1067-y

Heathcote, A., & Love, J. (2012). Linear deterministic accumulator models of simple choice. *Frontiers in Psychology: Cognitive Science, 3*(292). doi:10.3389/fpsyg.2012.00292

Howard, Z. L., Evans, N. J., Innes, R. J., Brown, S. D., & Eidels, A. (2020). How is multi-tasking different from increased difficulty? *Psychonomic Bulletin & Review*. doi:10.3758/s13423-020-01741-8

International Standards Organisation DIS 17488 (2016). Road Vehicles -Transport information and control systems -Detection Response Task (DRT) for assessing selective attention in driving. International Standard, ISO TC 22/SC39/WG8.

Kendall, M., & Stuart, A. (1967). The Advanced Theory of Statistics, vol. 2, Inference and Relationship, Haffner, New York. *Kendall2The Advanced Theory of Statistics1967*.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*(4), 352–358. doi: https://doi.org/10.1037/h0043688

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451-468. doi:10.1037//0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences, 9*(2), 75-82. doi:https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science, 19*(3), 143-148.

Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics, 72*(1), 246-273. doi:10.3758/APP.72.1.246

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences, 1129*(1), 305-322. doi:10.1196/annals.1417.002

Lim, J., & Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological Bulletin, 136*(3), 375–389. doi:https://doi.org/10.1037/a0018883

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*: Oxford University Press on Demand.

Mager, R., Falkenstein, M., Störmer, R., Brand, S., Müller-Spahn, F., & Bullinger, A. H. (2005). Auditory distraction in young and middle-aged adults: A behavioural and event-related potential study. *Journal of Neural Transmission, 112*(9), 1165-1176. doi:10.1007/s00702-004-0258-0

Matzke, D., Curley, S., Gong, C.Q., & Heathcote, A. (2019). Inhibiting responses to difficult choices. *Journal of Experimental Psychology: General, 148,* 124-142.

Meule, A. (2017). Reporting and interpreting working memory performance in n-back tasks. *Frontiers in Psychology: Emotion Science, 8*(352). doi:10.3389/fpsyg.2017.00352

Miller, K. M., Price, C. C., Okun, M. S., Montijo, H., & Bowers, D. (2009). Is the n-back task a valid neuropsychological measure for assessing working memory? *Archives of Clinical Neuropsychology, 24*(7), 711-717. doi:10.1093/arclin/acp063

Muller-Gass, A., Stelmack, R. M., & Campbell, K. B. (2006). The effect of visual task difficulty and attentional direction on the detection of acoustic change as indexed by the Mismatch Negativity. *Brain Research, 1078*(1), 112-130. doi:https://doi.org/10.1016/j.brainres.2005.12.125

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences, 13*(2), 201-233. doi:10.1017/S0140525X00078407

Näätänen, R., Kujala, T., & Winkler, I. (2011). Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology, 48*(1), 4-22. doi:10.1111/j.1469-8986.2010.01114.x

Parmentier, F. B. R., Maybery, M. T., & Elsley, J. (2010). The involuntary capture of attention by novel feature pairings: A study of voice—location integration in auditory sensory memory. *Attention, Perception, & Psychophysics, 72*(2), 279-284. doi:10.3758/app.72.2.279

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*(5), 347-356. doi:https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., & Strayer, D. (2014). Modeling simple driving tasks with a one-boundary diffusion model. *Psychonomic Bulletin & Review, 21*(3), 577-589. doi:10.3758/s13423-013-0541-x

Ratcliff, R., & Van Dongen, H. P. A. (2011). Diffusion model for one-choice reaction-time tasks and the cognitive effects of sleep deprivation. *Proceedings of the National Academy of Sciences, 108*(27), 11285-11290. doi:10.1073/pnas.1100483108

Reason, J. (1998). How necessary steps in a task get omitted: Revising old ideas to combat a persistent problem. *Cognitive Technologies, 3*, 24-32.

Reason, J. (2002). Combating omission errors through task analysis and good reminders. *Quality and Safety in Health Care, 11*(1), 40-44. doi:10.1136/qhc.11.1.40

Riccio, C. A., Reynolds, C. R., Lowe, P., & Moore, J. J. (2002). The continuous performance test: A window on the neural substrates for attention?☆. *Archives of Clinical Neuropsychology, 17*(3), 235-272. doi:10.1093/arclin/17.3.235

Rinne, T., Kirjavainen, S., Salonen, O., Degerman, A., Kang, X., Woods, D. L., & Alho, K. (2007). Distributed cortical networks for focused auditory attention and distraction. *Neuroscience Letters, 416*(3), 247-251. doi:https://doi.org/10.1016/j.neulet.2007.01.077

Rosvold, H. E., Mirsky, A. F., Sarason, I., Bransome Jr, E. D., & Beck, L. H. (1956). A continuous performance test of brain damage. *Journal of Consulting Psychology, 20*(5), 343-350. doi:https://doi.org/10.1037/h0043220

Sabri, M., Liebenthal, E., Waldron, E., Medler, D. A., & Binder, J. R. (2006). Attentional modulation in the detection of irrelevant deviance: A simultaneous ERP/fMRI study. *Journal of Cognitive Neuroscience, 18*(5), 689-700.

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PloS one*, *8*(1), e54402.

Sanford, J., & Turner, A. (2004). The Integrated Visual and Auditory Continuous Performance Test. Interpretive manual. In: Richmond, VA: BrainTrain.

Schröger, E. (1996). A neural mechanism for involuntary attention shifts to changes in auditory stimulation. *Journal of Cognitive Neuroscience, 8*(6), 527-539.

Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research, 7*(1), 71-87. doi:https://doi.org/10.1016/S0926-6410(98)00013-5

Simen, P., Vlasov, K., & Papadakis, S. (2016). Scale (in) variance in a unified diffusion model of decision making and timing. *Psychological Review, 123*(2), 151-181. doi:https://doi.org/10.1037/rev0000014

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583-639. doi:10.1111/1467-9868.00353

Strayer, D. L., Turrill, J., Cooper, J. M., Coleman, J. R., Medeiros-Ward, N., & Biondi, F. (2015). Assessing cognitive distraction in the automobile. *Human Factors, 57*(8), 1300-1324. doi:10.1177/0018720815575149

Ulrich, R., & Miller, J. (1994). Effects of truncation on reaction time analysis. *Journal of Experimental Psychology: General, 123*(1), 34 –80. doi:https://doi.org/10.1037/0096-3445.123.1.34