Cognitive Psychology 95 (2017) 1–16

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Breaking the rules in perceptual information integration

Maxim A. Bushmakin [a,b,*], Ami Eidels [c], Andrew Heathcote [c,d]

[a] Department of Psychological and Brain Sciences, Indiana University, USA
[b] Volen National Center for Complex Systems, Brandeis University, USA
[c] School of Psychology, The University of Newcastle, Australia
[d] School of Medicine, The University of Tasmania, Australia

ARTICLE INFO

Article history:
Accepted 25 March 2017

ABSTRACT

We develop a broad theoretical framework for modelling difficult perceptual information integration tasks under different decision rules. The framework allows us to compare coactive architectures, which combine information before it enters the decision process, with parallel architectures, where logical rules combine independent decisions made about each perceptual source. For both architectures we test the novel hypothesis that participants break the decision rules on some trials, making a response based on only one stimulus even though task instructions require them to consider both. Our models take account of not only the decisions made but also the distribution of the time that it takes to make them, providing an account of speed-accuracy tradeoffs and response biases occurring when one response is required more often than another. We also test a second novel hypothesis, that the nature of the decision rule changes the evidence on which choices are based. We apply the models to data from a perceptual integration task with near threshold stimuli under two different decision rules. The coactive architecture was clearly rejected in favor of logical-rules. The logical-rule models were shown to provide an accurate account of all aspects of the data, but only when they allow for response bias and the possibility for subjects to break those rules. We discuss how our framework can be applied more broadly, and its relationship to Townsend and Nozawa's (1995) Systems-Factorial Technology.

1. Introduction

Human choices often depend on combining noisy signals from multiple sources. When approaching an intersection on a dark and rainy night, for example, a driver must determine whether traffic lights are red or green and whether pedestrians are crossing. If the light is red, or if pedestrians are crossing, braking is required, following an "OR" decision rule for the presence of either one signal *or* the other. Once stopped, and before continuing a trip, the driver must confirm that the light is green and that no pedestrians are in their path, following an "AND" decision rule requiring both one signal *and* the other. The OR rule allows processing to be terminated after only one signal is detected, whereas the AND rule requires processing both signals, leading Townsend (1974) to describe them as stopping rules requiring, respectively, first-terminating and exhaustive processing of stimuli. Stopping rules have been studied in many areas of human cognition, from categorization (Fific, Little, & Nosofsky, 2010), to consumer choices (Fific & Buckmann, 2013), and memory- and visual-search tasks (Fific, Townsend, & Eidels, 2008; Sternberg, 1966).

* Corresponding author at: Volen National Center for Complex Systems, Brandeis University, USA
E-mail address: mbushmak@brandeis.edu (M.A. Bushmakin).

http://dx.doi.org/10.1016/j.cogpsych.2017.03.001
0010-0285/

Early investigations of stopping rules manipulated the number of items in either memory or a visual display and focused on the slope of the response time (RT) function as the number of items increases (Sternberg, 1966, 1969; see Algom, Eidels, Hawkins, Jefferson, and Townsend (2015), for a review). However, later work by Townsend and colleagues (e.g., Townsend & Ashby, 1983; Townsend & Colonius, 1997) showed that this approach was flawed, because different rules could predict the same pattern of slopes. Subsequently, Townsend and Nozawa (1995) showed that AND and OR stopping rules do make unique predictions in a design – often called the "double-factorial paradigm" – that manipulates the relative salience of two or more signals.

It has often been found using the double-factorial paradigm that human observers do appear to apply the stopping rule appropriate to the task at hand (e.g., Eidels, Townsend, & Algom, 2010b; Fific, Nosofsky, & Townsend, 2008; Fific et al., 2008; Fific et al., 2010; Little, Nosofsky, & Denton, 2011; Moneer, Wang, & Little, 2016). However, double-factorial designs are not always easy to implement, and Townsend and Nozawa's (1995) analysis does not take account of the effects of mixtures of stopping rules (i.e., applying different rules on different trials). Cousineau and Shiffrin (2004) found evidence for such inconsistent rule application in a difficult visual-search task. On a proportion of trials it appeared that only one of the two items in a display was processed fully, with participants guessing about the other item.

The use of an inappropriate stopping rule can have detrimental effects on behavioural performance. For example, exhaustively processing under an OR requirement requires more effort and slows response with no benefit to accuracy. However, inappropriate stopping may also have benefits. For example, failing to process exhaustively under an AND requirement, and either ignoring the remaining items, or making guesses about them as in Cousineau and Shiffrin's (2004) study, makes responding faster and easier, although it will also cause errors. Depending on the value a participant places on speed over accuracy—perhaps as a function of trying to complete an experiment with minimal effort, or in order to have time to attempt more decisions in a fixed time period—rule breaking may be an attractive and even optimal strategy (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006).

In the present paper we focus on a perceptual information integration task. Such tasks are of interest because humans are often in a position where they need to integrate information within a sensory modality (Bushmakin & James, 2014; Eidels, Donkin, Brown, & Heathcote, 2010a) as well as across modalities (Alais & Burr, 2004; McGurk & MacDonald, 1976; Stevenson et al., 2014a). Moreover, the irregularities in perceptual integration have been linked to certain disorders like autism and schizophrenia (Stevenson et al., 2014b; Williams, Light, Braff, & Ramachandran, 2010).

Despite both empirical evidence for rule breaking in other tasks, and in at least some cases there being a rational basis for participants to purse a rule-breaking strategy, previous analyses of information integration have not, to our knowledge, considered whether decision makers might sometimes not abide by the rules that experimenters use to score the accuracy of their performance. The present paper develops a framework that takes into account the possibility of rule breaking for tasks requiring both AND and OR rules, and the possibility that a participant can use a mixture of rule-following and rule-breaking strategies. We apply models derived from this framework to the performance of each participant separately to account for the possibility that there will be individual differences in the factors that cause rule breaking, such as the value placed on effort or the speed vs. accuracy.

The nature of the two stopping rules we investigate makes it important to take account of tradeoffs between the speed and accuracy of different responses. These OR and AND designs have built-in biases towards one response or another. Consider the driver OR example presented above; attempting to drive through the intersection, the driver needs to stop if she detects a pedestrian approaching, a red light, or both. In contrast there is only one case in which she can go: if there is a green light and no pedestrians. These contingencies are known to create biases in responding (e.g., Mordkoff & Yantis, 1991). All other things being equal, it is likely that a bias towards responding "YES" will develop under OR instructions and a bias to respond "NO" under AND instructions, because a YES response is required more often in the former case and a NO response more often in the latter case. Attempts to remove those biases by unbalanced presentation of stimuli (e.g., more no-target trials in AND task or more double-target trials in an OR task) create other contingencies and hence other potential biases. In light of this, our experiments used balanced stimulus presentation, and we take account of the potential biases in our models.

In summary, two questions remain unanswered: (1) how well do people abide by AND and OR decision rules, and the related questions of how they manage the associated trade-off between speed and accuracy; and (2) how they select an appropriate level of response bias under each rule. To address these questions, we also investigated two other fundamental questions about the effect of decision rules on the inputs to, and architecture of, the decision process.

First, do response-rule instructions affect only the decision process itself, or do they affect the inputs to the decision process, that is, the evidence on which decisions are based? In particular, is the evidence accumulation rate for each signal the same across decision rules? This question is novel, perhaps because most past analyses have not tried to simultaneously account for responding under two different decision rules by the same participants, as we do here. In contrast, the second question has been the subject of intense scrutiny: is evidence combined before it enters a single decision process (sometimes called a coactive architecture; e.g., Little, Nosofsky, Donkin, & Denton, 2013; Miller, 1978, 1982), or are separate decisions made about each signal and later combined by logical rules (e.g., Eidels et al., 2010a).

To address all of these questions, we develop a unified theoretical framework that expands on Brown and Heathcote's (2008) Linear Ballistic Accumulator (LBA) model of choice RT and on its extension to a logical-rule model of the OR task developed by Eidels, Donkin et al. (2010) and Eidels, Townsend et al. (2010). We generalized Eidels et al.'s model of the OR task to the AND task, and, for both AND and OR models, we also allowed for the possibility that participants sometimes broke the rules and processed only one signal. We also extended Brown and Heathcote's standard LBA to provide coactive AND-task and OR-task models. The coactive models assume that evidence from a number of sources (signals) is combined before entering a single LBA decision process with the possibility of rule breaking. We compared these models to test the effects of response-instructions on architecture (coactive vs. parallel) and rule breaking, and used parameter estimates associated with the models to test for their effects on inputs and response bias. Together the set of models provides a powerful and comprehensive framework for answering questions about noisy perceptual information integration.

We apply our framework to data reported by Eidels, Townsend, Hughes, and Perry (2015), where the same participants made difficult AND and OR decisions about the presence or absence of one or two near-threshold dots of light, one positioned above and one below fixation. Eidels et al. (2015) examined this data with an accuracy-based measure, the "no-response probability contrast" (NRPC; Mulligan & Shaw, 1980). They also analyzed RT data for correct responses collected from the same participants in another experiment with supra-threshold stimuli using Townsend and Nozawa's (1995) nonparametric Systems Factorial Technology (SFT). After describing our parametric model-based analyses and applying it to Eidels et al.'s (2015) experiment with near-threshold stimuli, we discuss the relationship of our results to their NRPC and SFT based results.

2. A Modelling framework for perceptual integration

Our implementation of coactive models uses a standard two-accumulator LBA (see Fig. 1), where each accumulator corresponds to a potential response (i.e., YES or NO) and pools inputs from each position (i.e., potential dot locations: above and below fixation). On each trial accumulators begin with evidence totals drawn independently from a uniform distribution with range (0, A) and a rate of evidence accumulation that is drawn independently from a Gaussian distribution with mean v and standard deviation sv. Evidence accrues linearly, and the response selected corresponds to the accumulator that first reaches its threshold, b. Differences in thresholds between accumulators mediate response bias. RT is the sum of the response-selection time and the time for perceptual encoding of the stimulus and response production, t_{er} ($t_{er} = t_e + t_r$, Fig. 1) which is assumed to be the same for both accumulators and greater than zero.

All models—both coactive and logical rule—were subject to some common parameter constraints. They assumed t_{er} was the same across stimulus conditions defined by the number of targets. The starting point distribution parameter A was allowed to change between conditions yielding two parameters in A_{AND} and A_{OR}. Threshold (b) was re-parameterized in terms of the distance from the top of the start-point distribution to the threshold, $B = b–A$, so that the restriction $B > 0$ enforced the condition that evidence never starts above the threshold (Brown & Heathcote, 2008). We also allowed B to vary between accumulator (i.e., Y = YES or N = NO) and condition to account for the possibility of bias, yielding four parameters: $B_{Y,AND}$, $B_{N,AND}$, $B_{Y,OR}$, and $B_{N,OR}$. Finally, we fixed $sv = 1$ for the accumulator that mismatched the stimulus (i.e., the YES accumulator if no target was present and the NO accumulator if a target was present) in order to make the model identifiable (see Donkin, Brown, & Heathcote, 2009). The sv parameter for the other accumulator was estimated from the data, consistent with previous LBA applications (e.g., Heathcote & Love, 2012; Rae, Heathcote, Donkin, Averell, & Brown, 2014).

The stimulus encoded for each position provides evidence for the presence and absence of a target. We examined both coactive and logical rule models in which rates could differ for upper and lower positions; however, according to our model-selection methods, these models did not provide sufficient improvement in fit over models that assumed same rates to justify the associated doubling in the number of estimated rate parameters.[1] Hence we assume in all of the results reported here the same rates for upper and lower positions.

Rates can differ depending on the response accumulator and whether a target stimulus is actually absent (A) or present (P) at a given position. Hence, there are four types of rates estimated: two when a target is present, for the YES accumulator ($v_{Y,P}$) and for the NO accumulator ($v_{N,P}$), and two when a target is absent, again one for the YES accumulator ($v_{Y,A}$) and one for the NO accumulator ($v_{N,A}$). Hence, $v_{Y,P}$ and $v_{N,A}$ are rates corresponding to the correct response, whereas $v_{Y,A}$ and $v_{N,P}$ indicate rates for the accumulators corresponding to an incorrect response. When a target is present, the ordering $v_{Y,P} > v_{N,P}$ results in above chance responding, as the YES accumulator tends to reach its threshold more quickly than the NO accumulator. Similarly, when the target is absent, $v_{N,A} > v_{Y,A}$ implies above chance responding. These inequalities would be expected to hold under a wide range of reasonable assumptions. This would be the case, for example, if the YES accumulator rate is an increasing function of how far luminance of the stimulus (l) is above a threshold (χ), $rate = g(l - \chi)$, and NO accumulator rate is an increasing function of how far luminance is below that threshold, $rate = h(\chi - l)$, where $g()$ and $h()$ are monotonic increasing functions.

2.1. Coactive models

For the coactive models, pooling of inputs from each position was achieved by simply adding rates associated with upper and lower target positions. Assuming inputs of equal magnitude for each position, the inputs to YES and NO accumulators

[1] The coactive model with separate rates for top and bottom stimuli, had AIC = −1614 and BIC = −324 for the joint parameterization (16 parameters) and AIC = −8087 and BIC = −6324 for the separate parameterization (26 parameters). Logical rule models with separate rates for top and bottom stimuli, had AIC = −9528 and BIC = −8441 for the joint parameterization (16 parameter) and AIC = −10,014 and BIC = −8252 for the separate parameterization (26 parameters). These models were also worse according to the BIC and AIC methods of model selection relative to the rule breaking models with the same accumulation rate for both locations.

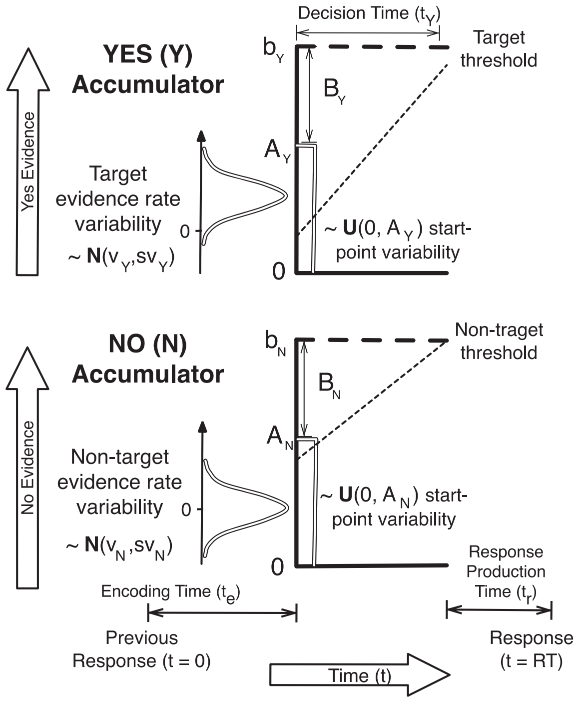

Fig. 1. The standard LBA model, with separate accumulators for target (Y = YES) and non-target (N = NO) responses.

will be $2v_{Y,P}$ and $2v_{N,P}$ respectively when both targets are present. When only one target is present, regardless of where it is positioned, the inputs to YES and NO accumulators are, respectively, $v_{Y,P} + v_{Y,A}$ and $v_{N,P} + v_{N,A}$. When both targets are absent, the inputs to YES and NO accumulators are $2v_{Y,A}$ and $2v_{N,A}$, respectively. In all cases the standard deviation of the sum is $2sv^2$.

For example, suppose $v_{Y,P} = v_{N,A} = 1$ and $v_{Y,A} = v_{N,P} = 0$ (where $v_{Y,P} > v_{N,P}$ and $v_{N,A} > v_{Y,A}$, as required for above chance accuracy). When there are two targets, the input to the YES accumulator is larger than the input to the NO accumulator (2 vs. 0), whereas when there are no targets present the opposite is true (0 vs. 2). When only one target is present both accumulators have the same input, but an appropriate level of response bias can still enable accurate responding for AND and OR decision rules, because these rules are linearly separable. For an AND decision rule, for example, the required response is NO for a single target, so the response threshold must be lower for the NO accumulator. For an OR decision rule, in contrast, the threshold must be lower for the YES accumulator.

All models (both coactive models and the logical rule models detailed below) were fit to each participant's data separately by maximizing the sum of the logarithm of the likelihoods (L) for each observed response-RT pair. Given parameters $\boldsymbol{\theta} = (b, A, v, sv)$, and $t = \mathrm{RT} - t_{er}$, the density function $f(t|\boldsymbol{\theta})$ gives the instantaneous probability that an accumulator reaches its threshold at time t, and the survivor function $S(t|\boldsymbol{\theta})$ gives the probability that an accumulator has not reached threshold at time t. Denoting the coactive model by the subscript CO, we have:

$$\mathrm{L}_{\mathrm{YES|CO}}(\boldsymbol{\theta}_{\mathrm{YES}}, \boldsymbol{\theta}_{\mathrm{NO}}|t) = f_{\mathrm{YES}}(t|\boldsymbol{\theta}_{\mathrm{YES}}) \times S_{\mathrm{NO}}(t|\boldsymbol{\theta}_{\mathrm{NO}}) \tag{1}$$

$$\mathrm{L}_{\mathrm{NO|CO}}(\boldsymbol{\theta}_{\mathrm{YES}}, \boldsymbol{\theta}_{\mathrm{NO}}|t) = f_{\mathrm{NO}}(t|\boldsymbol{\theta}_{\mathrm{NO}}) \times S_{\mathrm{YES}}(t|\boldsymbol{\theta}_{\mathrm{YES}}) \tag{2}$$

In words, these equations say that the likelihood (which is just the probability density function with the conditioning of t and $\boldsymbol{\theta}$ reversed) of a given response at time t equals the density (f) of the corresponding accumulator reaching its threshold at time t times the probability (given by the survivor function, S) that the other accumulator has *not* reached its threshold. Fitting a separate model to either the AND or OR condition required estimation of nine parameters: four *vs* two *B*s (for YES and NO accumulators to allow for response bias) and one A, sv, and t_{er}.

We fit two types of coactive model. The first assumed only the A parameters and the two B parameters (one each for YES and NO accumulators) differed between the AND and OR instruction conditions (i.e., a joint model, with 12 parameters: $2 \times A$, $4 \times B$, $4 \times v$, t_{er}, and sv). This joint coactive model instantiates the assumption that decision-rule instructions only change the decision process, not the inputs to the decision process. We also fit a more flexible model in which all parameters could differ between AND and OR conditions (i.e., two separate fits, with 18 parameters in total), instantiating the assumption that the instructions could also change the nature of the evidence on which decisions are based.

2.2. Logical-rule models

The logical rule models have two pairs of accumulators, one pair taking input from the lower position (LT, lower target) and one from the upper (UT, upper target). Unlike the coactive models, evidence from the two positions is not combined prior to the decision. Rather, a separate decision is made for each position concerning the presence or absence of a target at that location, and the outcome of each sub-decision is then combined logically. For both AND and OR response rules, the likelihood of a YES or NO response at time t can be worked out using multiplication (to get the probability that one event *and* another occurs) and addition (to get the probability that one event *or* another occurs). For a YES response in the AND case, both NO accumulators must be below threshold (i.e., they are both "survivors"), which occurs with probability $S_{NO,LT}(t) \times S_{NO,UT}(t)$. There are two possible cases for the YES accumulators (i.e., one case *or* another). For the first case, the lower YES accumulator has previously reached threshold, with probability $F_{YES,LT}(t) = 1 - S_{YES,LT}(t)$, and the upper YES accumulator reaches threshold at time t, with instantaneous probability $f_{YES,UT}(t)$. The second case swaps the roles of upper and lower YES accumulators. The products of these terms are then summed as shown in the following equation, where for brevity t is omitted:

$$L_{YES|AND} = S_{NO,LT} \times S_{NO,UT} \times (f_{YES,LT} \times F_{YES,UT} + f_{YES,LT} \times F_{YES,UT}) \quad (3)$$

For a NO response both YES accumulators cannot be above threshold, which is most easily calculated by subtracting the probability that both are above threshold (i.e., $F_{YES,LT}$ x $F_{YES,UT}$) from one. Again there are two cases, where the lower NO accumulator finishes before the upper NO accumulator, or vice versa. Hence the likelihood of a NO response is:

$$L_{NO|AND} = (1 - F_{YES,LT} \times F_{YES,UT}) \times (f_{NO,LT} \times S_{NO,UT} + f_{NO,UT} \times S_{NO,LT}) \quad (4)$$

The OR rule, described in detail in Eidels, Donkin et al. (2010) and Eidels, Townsend et al. (2010), yields equations of a very similar form, but swapping the structure of YES and NO equations and the roles of YES and NO accumulators within each:

$$L_{YES|OR} = (1 - F_{NO,LT} \times F_{NO,UT}) \times (f_{YES,LT} \times S_{YES,UT} + f_{YES,UT} \times S_{YES,LT}) \quad (5)$$

$$L_{NO|OR} = S_{YES,LT} \times S_{YES,UT} \times (f_{NO,LT} \times F_{NO,UT} + f_{NO,UT} \times F_{NO,LT}) \quad (6)$$

In an analogous manner to the coactive models (and with the same number of parameters), we fit the logical rule models either separately to each AND and OR conditions, or jointly, assuming the same values for input-related parameters across the two conditions (i.e., the joint parameterization assuming no effect of decision rule on inputs).

2.3. Rule-breaking models

We extended both coactive and logical-rule models by assuming that participants only follow the rules on some trials, which occurs with probability p. When they break the rules, with probability $(1 - p)$, they either process only the lower position, with probability q, or the upper position, with probability $(1 - q)$. In the separate fits, this adds two parameters to each of the AND and OR conditions. In the joint fits, we allowed separate values of p for each decision-rule condition, but a common value of q.[2]

The likelihoods of YES and NO responses in the rule-breaking (RB) case are:

$$L_{YES|RB} = q \times (f_{YES|LT} \times S_{NO|LT}) + (1 - q) \times (f_{YES|UT} \times S_{NO|UT})$$

$$L_{NO|RB} = q \times (f_{NO|LT} \times S_{YES|LT}) + (1 - q) \times (f_{NO|UT} \times S_{YES|UT})$$

For the rule-breaking coactive (RB-CO) model:

$$L_{YES|RB-CO} = p \times L_{YES} + (1 - p) \times L_{YES|RB} \quad (7)$$

$$L_{NO|RB-CO} = p \times L_{NO} + (1 - p) \times L_{NO|RB} \quad (8)$$

As in the rule-following model, AND and OR have identical likelihood equations and are differentiated by parameter estimates.

For the rule-breaking logical-rule models:

$$L_{YES|RB-RULE,AND} = p \times L_{YES|AND} + (1 - p) \times L_{YES|RB} \quad (9)$$

$$L_{NO|RB-RULE,AND} = p \times L_{NO|AND} + (1 - p) \times L_{NO|RB} \quad (10)$$

$$L_{YES|RB-RULE,OR} = p \times L_{YES|OR} + (1 - p) \times L_{YES|RB} \quad (11)$$

[2] Allowing separate values of q for the two rule conditions produced unstable estimates due to multiple local minima and was rejected by model selection. Additionally, all of the subjects demonstrated same preference for one location over the other, so for simplicity we report only results where a common value of q was assumed.

$$L_{NO|RB-RULE,OR} = p \times L_{NO|OR} + (1-p) \times L_{NO|RB} \quad (12)$$

Each rule-breaking model requires estimation of two more parameters than the corresponding rule-following model.

In the following analysis of Eidels et al.'s (2015) data, we report the results for eight types of models: a factorial combination of coactive vs. logical-rule, rule-following vs. rule-breaking, and separate vs. joint parameters. Fits were obtained by maximizing the likelihoods in (1)-(12).

3. Example application

Eidels et al.'s (2015) experiments are similar to those used by Townsend and Nozawa (1995) to develop SFT, with the dots appearing 1° above or below fixation 0.5 s after the offset of a fixation point. SFT provides a non-parametric method of identifying not only stopping rules, but also cognitive architecture (e.g., serial vs. parallel processing) and the capacity constraints (e.g., limited vs. unlimited capacity).

For their first study, Eidels et al. (2015) focused on an SFT analysis in a paradigm using both dimmer or brighter super-threshold dots. SFT's mean- and survivor-interaction contrast analyses, which are diagnostic of architecture and stopping rule, clearly rejected serial processing and favoured a parallel-processing architecture in both conditions, with each following the appropriate rule, first-terminating for OR instructions and exhaustive for AND (see Houpt and Townsend (2010), for a converging analysis).

For our purposes here, Eidels et al.'s (2015) first study introduces a challenge, since the use of super-threshold stimuli caused accuracy to be near ceiling. This makes it difficult to disentangle differences between YES and NO responses caused by bias as opposed to effects caused by a difference in the quality of evidence for each alternative. Response bias seems likely as all four stimulus conditions—no target (NT), lower target only (LT), upper target only (UT), and a double target (DT)—were equi-probable. Hence, a NO response was required on 75% of trials under AND instructions, making it likely that some participants developed a bias to respond NO, whereas a YES response was required on 75% of trials under OR instructions, making a YES bias likely.

Fortunately, when accuracy is lower, bias can be identified because it has a negatively correlated effect on speed and accuracy (i.e., it causes a speed-accuracy trade off), whereas changes in evidence quality have a positively correlated effect. For the same reason, it is difficult to fit evidence-accumulation models such as the LBA to high accuracy data, because such models estimate separate parameters related to bias (i.e., the threshold amount of evidence required to make each response) and evidence (i.e., the rate at which evidence enters accumulators associated with each choice). Given these considerations, the analysis reported here used data from Eidels et al.'s (2015) second study, which re-tested the same participants as the first, but calibrated dot luminance for each participant to a near-threshold level in order to achieve 85% - 95% accuracy.

4. Results

4.1. Model selection

We measured model misfit using deviance (D), which is minus two times the likelihood that was maximized to estimate the best-fitting parameters. Deviance differences between nested models (e.g., joint vs. separate parameterizations; rule-following vs. rule-breaking) have a $\chi^2(df)$ distribution, where df (degrees of freedom) equals the difference in the number of parameters estimated for each model. Model selection was accomplished using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) aggregated over the $i = 1 \ldots S$ participants ($S = 9$): $AIC = \sum_{i=1}^{S} D_i + 2(S \times p)$, and $BIC = \sum_{i=1}^{S} D_i + (S \times p) \times ln(\sum_{i=1}^{S} n_i)$, where n_i is the number of data points for the ith participant and p is the number of model parameters. The term after the plus sign in each case is a penalty for model complexity as measured by the number of model parameters (i.e., p parameters for each of the S subjects). The complexity penalty means that AIC and BIC can prefer worse fitting (i.e., higher deviance) models with fewer parameters to better fitting models with more parameters if the extra parameters do not improve fit sufficiently. The penalty is larger for BIC than AIC (as n is large), and so BIC prefers simpler models.

Table 1 shows a very clear rejection of the coactive models, which always fit much worse than the corresponding (i.e., same number of parameters) logical-rule models. Indeed, the advantage is so great that even the simplest logical-rule model fits better than the most flexible coactive model. Further, none of the three measures in Table 1 supported a coactive model in general and for any individual participant (supplementary materials Tables 1S and 2S, including JS and LB, for whom Eidels et al. (2015) found evidence for a co-active architecture). Hence, we focus only on the logical-rule models in further analyses.

AIC selected the most complex model; the model with both a separate parameterisation and rule-breaking, but BIC selected a simpler model with the joint parameterization. However, when the joint parameterization was enforced, the decrease in fit was significant for both the rule-following models, $\chi^2(54) = 605$, $p < 0.001$, and the rule-breaking models, $\chi^2(63) = 404$, $p < 0.001$. The same was true when rule-following is enforced: for the joint parameterization, $\chi^2(27) = 749$, $p < 0.001$, and for the separate parameterization, $\chi^2(36) = 549$, $p < 0.001$.

Fig. 2 compares the fit of the four logical-rule models to accuracy data, and reveals the source of the misfit for rule-following models: an inability to explain lower accuracy in the single target conditions in the OR task, particularly when

Table 1
Model selection results. Number of parameters (p) is specified per participant. The best AIC and BIC values are in bold font.

Model class	Rules breaking	Parameterization	Parameters per participant (p)	Summed deviance	AIC	BIC
Logical rule	No	Joint	12	−9289	−9073	−8257
		Separate	18	−9892	−9564	−8342
	Yes	Joint	15	−10,042	−9769	−**8751**
		Separate	22	−**10,441**	−**10,040**	−8547
Coactive	No	Joint	12	1015	1181	1997
		Separate	18	−7769	−7608	−6386
	Yes	Joint	15	−6124	−6208	−5189
		Separate	22	−8521	−8158	−6665

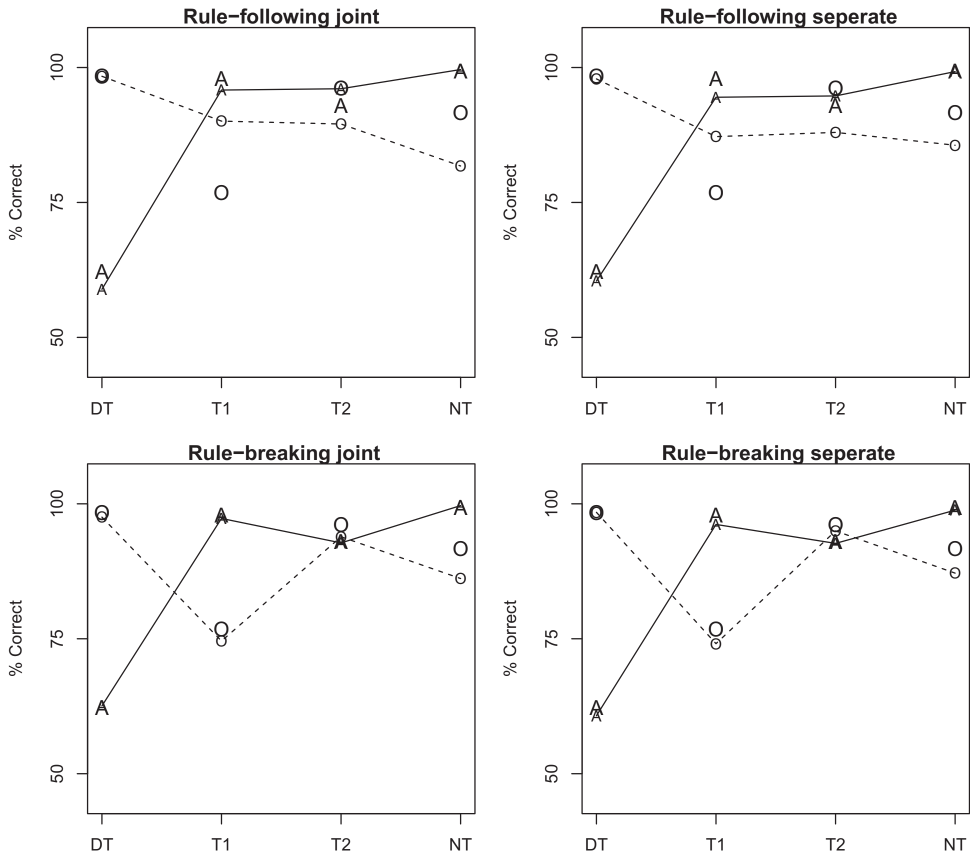

Fig. 2. Logical rule-following and rule-breaking model fits to accuracy data from the four within-subject conditions: NT = non-target, LT = lower target, one dot below, UT = upper target, one dot above, DT = double target, both dots. Larger "O" and "A" symbols indicate, respectively, average observed accuracy in the OR and AND tasks. Solid lines (with smaller "O" symbols) represent predictions for the OR condition and dotted lines (with smaller "A" symbols) predictions for the AND condition.

the target was in the lower location (LT); and to a lesser degree in both decision-rule conditions when the target was in the upper location (UT). The misfit to the AND data is more subtle, being mainly in the UT condition. Fig. 3 shows fits of the joint and separate parameterizations of rule-breaking models to RT distributions, represented by three percentiles indicating the fastest (10th percentile), middle (50th, i.e., median), and slowest (90th) responses. Both joint and separate parameterizations of rule-breaking models provide a good account of the RT data, with the exception of some over-prediction of slow responses in the single target UT condition. The fit of the joint parameterization model is only a little worse, indicating that using the same rate parameters across AND and OR instructions mainly affects the account of accuracy.

Overall these results support logical rule-breaking models. In order to see if there was any inconsistency in the degree of rule breaking between instruction conditions, we examined the evidence for rule breaking in separate fits to each instruction condition for this model. We found that for OR instructions, rule breaking was favoured over the rule following by both AIC (−6730 vs. −6226) and BIC (−5982 vs. −5612). However, under AND instructions, both AIC (−3314 vs. −3342) and BIC

(−2565 vs. −2729) supported rule following. To follow up on this result—and to examine individual differences—we calculated separately for each condition the posterior probability, based on both AIC and BIC (Wagenmakers & Farrell, 2004), for the rule-following and rule-breaking models. As shown in Table 2, the general picture favoured rule breaking in the OR condition and rule-following in the AND condition, but there were also some marked individual differences.

In particular, in the OR condition both AIC- and BIC-based results strongly favoured rule-breaking for all participants except WY, who was classified as clearly rule following by both BIC and AIC. For the AND condition, in contrast, the two measures consistently indicated strong evidence for rule breaking in only participant WY. For six others, the results consistently supported rule following. For the remaining two participants, AIC preferred the more complex rule-breaking model, whereas BIC preferred the simpler rule-following model. However, even for the participant classified as rule breaking in the AND condition, rule breaking was estimated as occurring on only a small number (2%) of trials. Estimates of the probability of rule breaking were much higher for all participants in the OR condition except WY, as shown in Table 3. Table 3 also shows that where rule breaking occurred it was almost always because participants processed only the top position, with the only exception being participant AW, who processed only the bottom position on about one quarter of the 14% of trials where they broke the rules.

Overall, our results suggest that the simpler separate parameterization of the rule-following model provides a good description of all participants in the AND condition, whereas the more complex separate parameterization of the rule-breaking model is required for the OR condition, although the level of rule breaking was quite minimal for one participant. Plots of these models (i.e., rule breaking for OR, rule following for AND, with separately estimated drift rates for each condition) to individual participant accuracy (Fig. 4) and RT distribution (Fig. 5) data confirmed they provide a good fit, with the exception of participant LB in the AND condition. However, the systematic misfit for LB was hardly improved even with the greater flexibility afforded by rule breaking; thus for this participant and condition, the misfit indicates that none of the models considered here was able to provide an entirely satisfactory account.

4.2. Parameter estimates

4.2.1. Evidence accumulation rates

We next performed an analysis of the parameter estimates for the models shown in Figs. 4 and 5. Given the rejection by model selection of the joint-parameterization model (which assumes that AND and OR instructions do not influence the inputs to evidence accumulation), we first examined rate estimates. Supporting the model-selection results, there was a significant main effect of instruction, $F(1,8) = 13.5$, $MSE = 1.77$, $p = 0.006$, and an interaction between the instruction and "match" (i.e., matching vs. mismatching accumulator, where the matching accumulator is the target accumulator for a target stimulus and a non-target accumulator for the a non-target stimulus) factors, $F(1,8) = 25.3$, $MSE = 0.4$, $p = 0.001$.

In order to understand these results, it is useful to focus on the difference between rates for the two levels of the match factor (i.e., the rate for the matching accumulator minus the rate for the mismatching accumulator). This difference indexes the "quality" of the information extracted from the stimulus, with larger values associated with more accurate responding. The two-way interaction indicated quality was greater on average in the OR condition (a difference of 3.5 on average) than the AND condition (a difference of 2). This advantage for the OR condition was very consistent, being present for every participant.[3]

The three-way interaction between the match, instruction and target (presence vs. absence) factors was also significant, $F(1,8) = 5.9$, $MSE = 1.7$, $p = 0.04$. To explore this interaction we conducted two separate two-way ANOVAs on rates for each instruction condition. In the OR condition a significant interaction between the match and target factors, $F(1,8) = 6.4$, $MSE = 0.36$, $p = 0.03$, marked higher quality information on the target absent condition (4.01) than the target present (2.98) condition. This difference in quality was very consistent at the individual level, with a substantial advantage for all but two participants, RM and AW, who had close to equal quality. In the AND condition, in contrast, the difference for target absent (1.96) and target present (2.02) were very similar, and the corresponding interaction did not approach significance, $F(1,8) = 0.03$, $MSE = 0.24$, $p = 0.85$. However, there were still considerable individual differences in the quality of information between the two conditions, ranging from participant MB, whose target-absent difference was 1.7 larger than their target present difference, to participant RM, with the opposite order held by 1.5.

4.2.2. Thresholds

Estimates of the LBA threshold parameter (B) indicated that participants were less cautious (i.e., required less evidence before making a response) in the OR condition ($B = 0.62$) than the AND condition ($B = 0.87$), $F(1,8) = 5.7$, $MSE = 0.1$, $p = 0.04$. This difference was quite consistent, being displayed by all but participant RS. There was also a strong interaction, $F(1,8) = 18.1$, $MSE = 0.02$, $p = 0.003$, because under AND instructions the threshold was lower for the NO accumulator (0.74) than the YES accumulator (1.0), whereas the reverse was true under OR instructions (0.68 vs. 0.56 respectively). This pattern of results of response bias is exactly what would be expected if participants (quite reasonably) set a lower threshold for the more common response in each condition (i.e., 75% NO in the AND condition and 75% YES in the OR condition). It was very

[3] Due to dropouts, session counterbalancing was not achieved, with AW, LB and WY doing the AND session first, whereas the remaining six participants did the OR session first. The OR advantage was on average 2.3 for the former group and 1.1 for the latter, so it is possible that at least some part of the OR advantage was due to, for example, a learning effect.

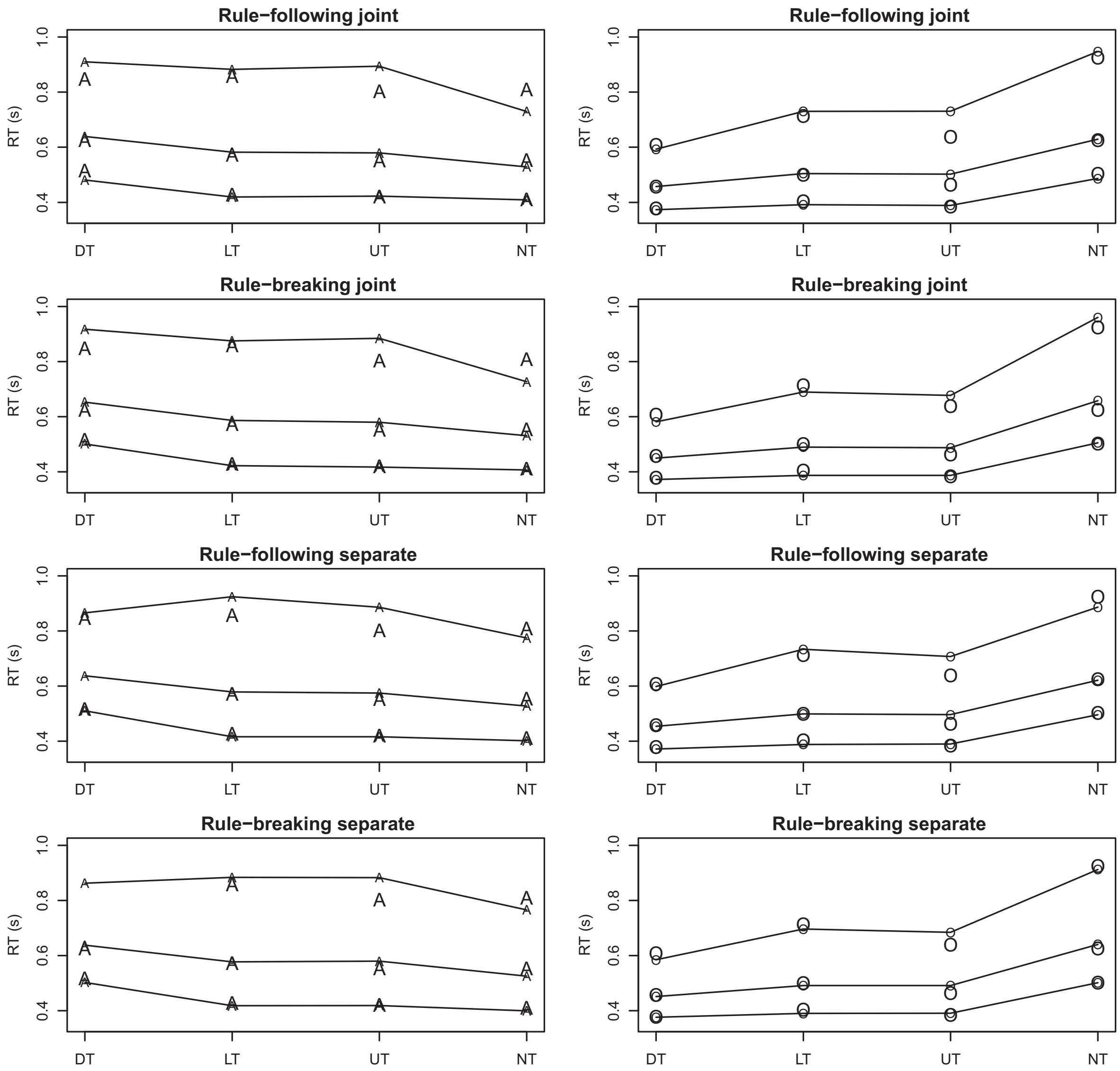

Fig. 3. Logical rule-following and rule-breaking separate and joint parameterization model fits to response RT distribution data (lower line = 10th percentile, middle line = 50th percentile, upper line = 90th percentile) from the four within-subject conditions: NT = non-target, LT = lower target, one dot below, UT = upper target, one dot above, DT = double target, both dots. Larger "O" and "A" symbols indicate, respectively, average observed RT percentiles in the OR and AND tasks. Solid lines with smaller "O" symbols represent predictions for the OR condition and solid lines with smaller "A" symbols represent the predictions for the AND condition.

Table 2
Posterior model probabilities across the set of eight models based on BIC for each participant for the separate parameterization rule breaking and rule following model. Probabilities are rounded to two decimal places.

		Models	BJ	RS	JS	MB	RM	LB	JG	WY	AW
AIC	AND	Rule following	**1**	**0.9**	**0.88**	**1**	0.13	0	.16	0	**0.99**
		Rule breaking	0.01	0.1	.12	0	**0.87**	**1**	**0.84**	**1**	0.01
	OR	Rule following	0	0	0	0	0	0	0	**0.98**	0
		Rule breaking	**1**	**1**	**1**	**1**	**1**	**1**	**1**	0.02	**1**
BIC	AND	Rule following	**1**	**1**	**1**	**1**	**0.94**	0.01	**0.95**	0.02	**1**
		Rule breaking	0	0	0	0	0.06	**0.99**	0.05	**0.98**	0
	OR	Rule following	0	0	0	0	0	0	0	**1**	0.05
		Rule breaking	**1**	**1**	**1**	**1**	**1**	**1**	**1**	0	**0.95**

Table 3
Parameter estimates for each participant of the probability of rule following (p) and the probability of processing the lower position given rule breaking (q) for the OR condition.

	BJ	RS	JS	MB	RM	LB	JG	WY	AW
p	0.79	0.68	0.88	0.53	0.54	0.74	0.61	1	0.84
q	0	0.04	0	0.01	0	0.07	0.05	0.02	0.25

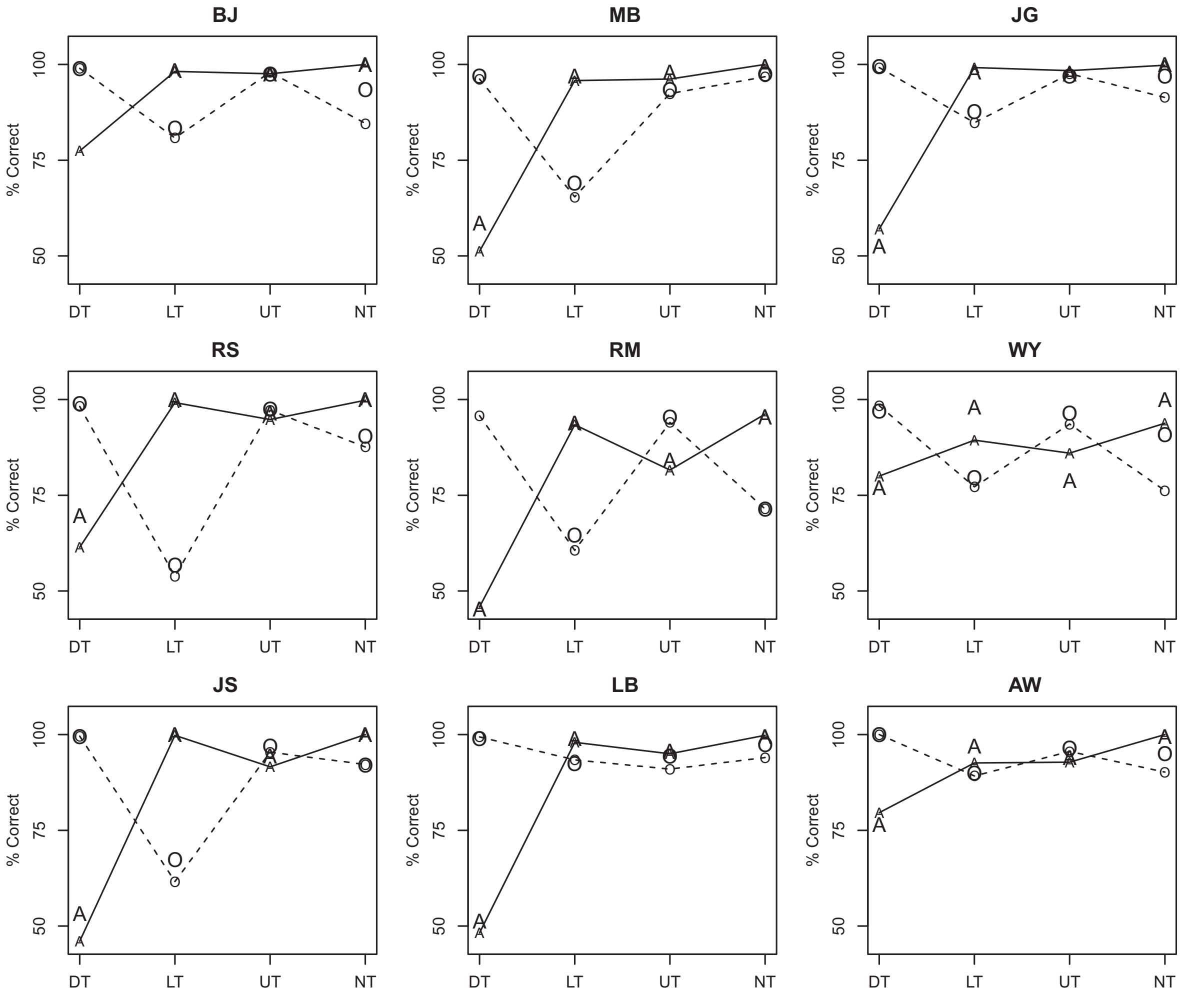

Fig. 4. Rule-breaking separate-parameterization model fits to accuracy data from the four within-subject conditions: NT = non-target, LT = lower target, one dot below, UT = upper target, one dot above, DT = double target, both dots. Larger "O" and "A" symbols indicate, respectively, observed accuracy in the OR and AND tasks. Solid lines (with smaller "O" symbols) represent predictions of the rule-breaking model for the OR condition and dotted lines (with smaller "A" symbols) predictions of the rule-following model for the AND condition.

consistent at the individual level, being displayed by every participant except LB, who had the opposite pattern of bias in the AND condition, the same condition where the model displayed systematic misfit.

The two remaining parameters, t_{er} (non-decision time) and A (start-point variability), did not differ significantly between instruction conditions: $F(1,8) = 0.3$, $MSE = 0.008$, $p = 0.61$, and $F(1,8) = 3.8$, $MSE = 0.38$, $p = 0.09$, respectively.

5. Discussion

In this paper we have developed a parametric framework for understanding the integration of information from noisy perceptual signals. The general framework uses mixtures of independent racing evidence-accumulation processes. In the detailed implementation of the framework we developed here, we assumed that these processes followed Brown and Heathcote's (2008) LBA model. For the data set we examined in detail here (Eidels et al., 2015), these assumptions

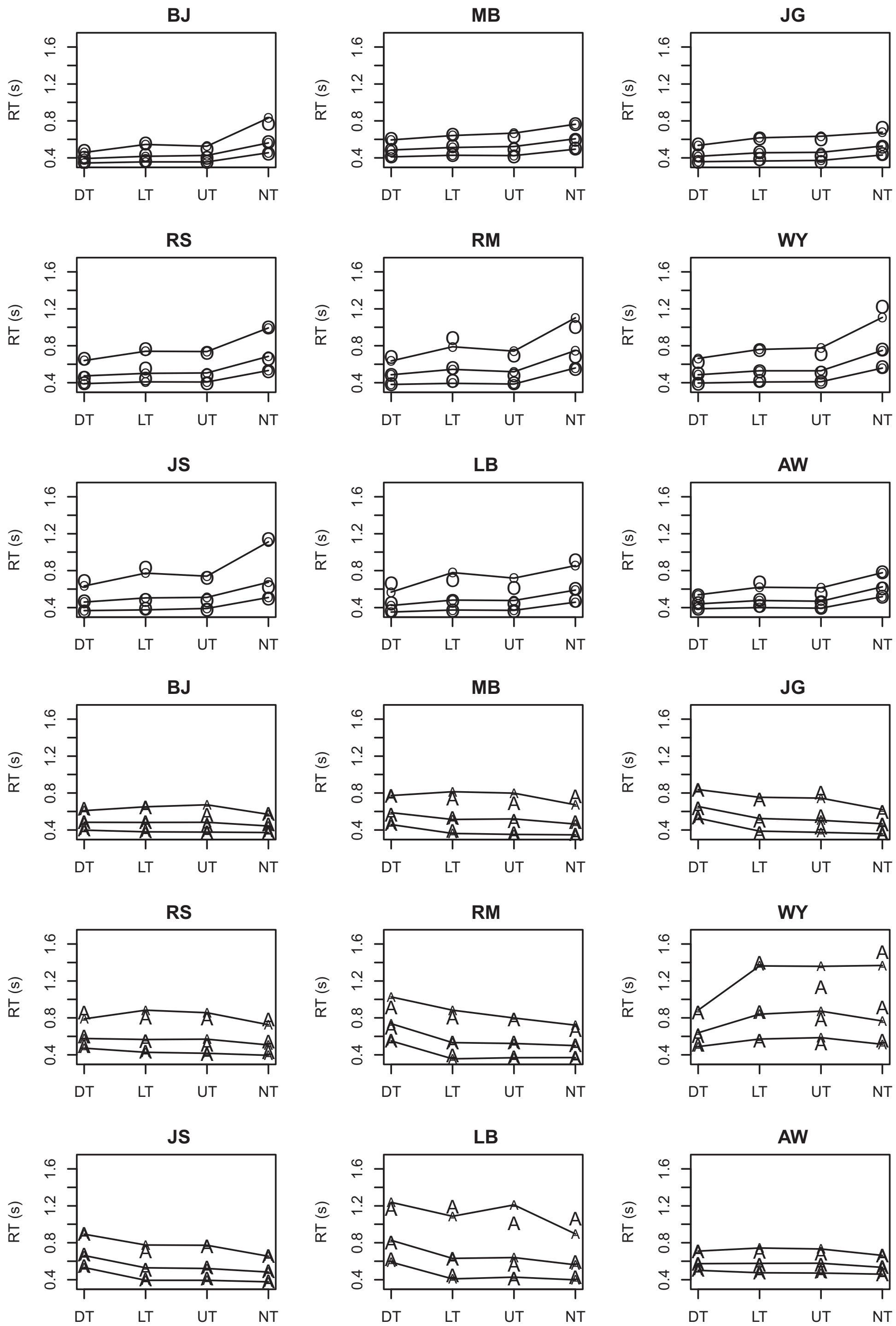

Fig. 5. Rule-breaking separate-parameterization model fits to response RT distribution data (lower line = 10th percentile, middle line = 50th percentile, upper line = 90th percentile) from the four within-subject conditions: NT = non-target, LT = lower target, one dot below, UT = upper target, one dot above, DT = double target, both dots. Larger "O" and "A" symbols indicate, respectively, observed RT percentiles in the OR and AND tasks. Solid lines (with smaller "O" symbols) represent predictions of the rule-breaking model for the OR condition and dotted lines (with smaller "A" symbols) predictions of the rule-following model for the AND condition.

proved sufficient to provide an accurate and detailed description of all aspects of each participant's performance, including the frequency of choices and the distribution of response time (RT), with only one exception, which we discuss further below.

For Eidels et al.'s (2015) data, our analysis provided clear answers to four questions concerning the integration of information from pairs of near-threshold stimuli with response rules of the OR (i.e., respond YES if one signal *or* another was present) or AND (i.e., respond YES if one signal *and* another is present) type. The first question concerns whether information from each stimulus is combined before it enters a single decision process—a "coactive" architecture—or whether separate decisions are made about each signal, with these decisions then combined by logical rules. Our model fits decisively rejected the coactive architecture in favor of logical rules. These findings largely agree with Eidels et al.'s analyses based on a contrast between the probability of responding NO in different conditions, which also rejected coactive processing for all participants in the OR condition and all but two in the AND condition.

The second question is, how well do people abide by AND and OR decision rules? Past attempts to model information integration have assumed participants are compliant, but we found otherwise. Under OR instructions, for about one quarter of trials on average, participants only processed one stimulus. However, there was substantial individual variation, with rule breaking occurring on anything from 4% to 45% of trials for different participants. For the same participants under AND instructions most were best described as never breaking the rules, with at best two of the nine participants failing to process both stimuli on 3% or less of trials. Thus, although there were individual differences, there was also quite a systematic effect of the type of response rule on the occurrence of rule breaking. We discuss the implications of these novel findings below.

Our third question was whether response rules affect only the decision process itself, or whether they also affect the inputs to the decision process, that is, the evidence on which decisions are based. Our results supported the latter answer, with two clear differences related to the quality of the information extracted from the stimuli. First, quality was higher in the OR condition than the AND condition. Second, in the OR condition, information quality for target-absent displays was systematically higher than for target-present displays, whereas in the AND condition there was no systematic difference. The advantage for the models that allowed response rules to affect rates was very consistent, occurring for all participants in all conditions.

The findings with respect to our third question are related to our fourth question—how participants managed the tradeoff between the speed and accuracy of responding—as we found that decision-processes thresholds also differed systematically between AND and OR conditions. In particular, most participants set a lower threshold in the AND than OR condition. This difference largely counteracted the better quality of information in the OR condition, so overall accuracy in the OR condition (91%) was only marginally greater than in the AND condition (88%). We also found that most participants selected an appropriate level of response bias that favoured the more commonly correct response (i.e., YES in the OR condition, and NO in the AND condition). We now discuss the implications of all of our findings for understanding information integration more generally, both with noisy near-threshold stimuli and with super-threshold stimuli.

5.1. Understanding information integration

The question of whether information integration involves coactivation versus an independent parallel architecture—as is assumed by our logical-rule models—has traditionally been addressed by non-parametric analyses focused on RT in paradigms where accuracy is near ceiling. Miller's (1978, 1982) seminal studies found violations of the RT-based race model inequality, consistent with coactive processing of visual and auditory signals, but later studies have failed to support this conclusion (Alais & Burr, 2004; Wuerger, Hofbauer, & Meyer, 2003). Mordkoff and Danek (2011) suggested that coactivation occurs only when redundant targets (i.e., two features with an OR response rule) are part of the same object. Eidels et al. (2015) hypothesized that processing architecture might also depend on the difficulty of the required discriminations. In particular, they suggested that, even though the signals in their second experiment were spatially separated, a coactive architecture might be favoured because pooling near-threshold stimuli might improve discrimination. This possibility could have important implications in real-world information integration tasks where, as suggested by the example with which we began the paper, integration of noisy perceptual information can be required.

Unfortunately, the study of near-threshold stimuli is complicated by the fact that differences between conditions in accuracy as well as RT often occur. Accuracy differences can potentially confound tests based solely on RT, such as the non-parametric race model inequality and related non-parametric techniques using the full distribution of RT, as developed in Townsend and Nozawa's (1995) Systems Factorial Technology (SFT). In order to address data where accuracy is well below ceiling, Eidels et al. (2015) extended seminal work by Mulligan and Shaw (1980), and developed a test of coactivation based purely on the probability of responding NO in different conditions. In close agreement with our findings, this test—which is also parametric as it assumes choices arise from a latent Gaussian distribution—rejected coactivation for all but two participants in the AND condition of their second experiment.

Taken together, these results indicate that near-threshold stimuli do not necessarily, or even not very often, cause coactive processing. Future work could examine whether this finding generalizes to other stimuli and paradigms, following up, for example, on evidence for coactivation with integral but not separable dimensions (Fific et al., 2008; Little et al., 2013) and on Mordkoff and Danek's (2011) suggestions that coactivation occurs when redundant targets (i.e., two features with an OR response rule) are part of the same object. For such purposes, the parametric framework we developed here has a number of advantages. Like Eidels et al.'s (2015) test of coactivation, our framework requires strong distributional assumptions (Jones & Dzhafarov, 2014). However, in contrast to their test, these assumptions make testable predictions that can be evaluated by the goodness-of-fit of the models (Heathcote, Brown, & Wagenmakers, 2014). Our framework also has the advantage that it addresses all aspects of the data, including the probability of both NO and YES responses as well as the full distribution of

RT (see Townsend and Altieri (2012) for a similarly comprehensive but non-parametric approach). A further advantage is that it is able to address the possibility that, at least on some occasions, participants break the rules of the information integration task.

Why might Eidels et al.'s (2015) participants have broken the rules and processed only one position on some trials? An enabling condition for this behavior is that this can be done while maintaining above chance accuracy, which is the case for both the AND and OR tasks. If participants processing only one position responded YES based on detecting a single target and NO otherwise, they will be 75% accurate in both the OR task (making an error only when the other position contained a target, which occurs on 25% of trials) and the AND task (making an error only when the other position contains a non-target, which again occurs on 25% of trials). Thus, the penalty for rule breaking is not particularly large; at the average rate observed in Eidels et al.'s OR condition, it adds only around 6% extra errors. Given that processing only one stimulus might speed at least a subset of responses (e.g., a NO response in the OR condition and a YES response in the AND condition, as both otherwise require classification of stimuli at both locations), this type of rule breaking provides a mechanism for achieving a speed-accuracy tradeoff.

One puzzle raised by our results is why rule breaking rarely occurred in the AND condition. Perhaps this was because participants construed the task as being about detecting the presence of dots, although logically it is equally about detecting their absence. Participants may have perceived the AND task as requiring the performance of two such target detections. This possibility suggests that rule breaking may be less prevalent when non-targets are defined by the presence rather than absence of a particular feature. Future research might examine this possibility using a variety of stimuli.

Given the possibility that rule breaking can mediate a speed-accuracy tradeoff, future research might also look at the effect of instructions to respond rapidly. Such instructions are likely to also cause changes in thresholds for decisions about individual features, and even potentially in the rate of evidence accumulation (e.g., Rae et al., 2014). The analysis framework developed here, which can accommodate such changes, will be appropriate for future analyses of such paradigms.

The occurrence of rule breaking, at least in the OR task, raises a potential concern that it has confounded earlier analyses of perceptual integration. This may not be a serious concern if rule breaking is associated with only near-threshold stimuli, since most prior research has used super-threshold stimuli. Super-threshold stimuli are associated with near-ceiling accuracy, which is not consistent with the errors caused by rule breaking. However, the calculations reported above—namely that only around 6% errors are added by rule breaking on 25% of trials—indicate that fairly high accuracy does not exclude rule breaking, at least on a sizeable minority of trials.

A potential experimental solution to the problem of rule breaking is to use an "exclusive or" (XOR) response rule. The XOR task requires a YES response when one or the other target is present, and a NO response when both targets are present or when neither target is present. Processing only one feature under these instructions leads to chance performance. Heathcote et al. (2015) also suggested an XOR response rule could be useful in removing response bias, which they also found in AND and OR tasks. As discussed previously, when the four conditions occur with equal frequency under AND and OR rules, NO and YES responses, respectively, are correct on 75% of trials, which induces response bias. One solution could have been to induce unequal condition frequencies so that correct YES and NO responses are equally likely. However, tampering with the relative frequencies of trial types may induce sequential dependencies that can cause other problems (i.e., anticipation of upcoming trials). Therefore, Heathcote et al. proposed using an XOR rule, which naturally has 50% correct YES and NO responses under equal trial-types frequencies.

The XOR task can also be useful to the non-parametric SFT, since the latter makes use of measures composed of responses that occur with different frequencies that can be distorted by response bias. Townsend and Eidels (2011) suggested as a remedy composing SFT contrasts from a mixture of different AND- and OR-rule responses that are correct with equal frequencies in their respective tasks (e.g., NO in AND and YES in OR). Unfortunately, this strategy may be confounded by other factors, such as any a priori tendency to prefer one response, or overall differences in speed between the sessions in which AND and OR tasks were performed. The XOR rule does not have these potential problems, and can achieve the goal of equating correct response frequency within a single task.

Perhaps our most surprising finding relates to the strong and systematic differences in the inputs to the decision process (i.e., the rates of evidence accumulation) induced by differences in response rules. Non-systematic differences between participants might be explained by fluctuations in the average level of attention occurring because the AND and OR tasks were performed in different sessions. However, this cannot explain why the quality of information about the stimuli was systematically higher in the OR session than in the AND session, and also why the quality of information about non-targets was systematically higher than the information about targets in the OR session, whereas it did not differ systematically in the AND session. To understand these findings it is important to note that, although the stimuli themselves—which did not differ between sessions—are one major determinant of evidence accumulation rates, rates are also influenced by factors related to attention (e.g., working memory capacity, Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007) and response instructions (e.g., speeded responding instructions can reduce quality; Rae et al., 2014; Starns, Ratcliff, & McKoon, 2012). Our results indicate that response-rule instructions can also affect the rate of evidence accumulation. Thus, these findings are at least not inconsistent with previous applications of evidence accumulation models, but further work is clearly required to understand why they occur.

Although the application of the framework we developed was largely successful, two limitations are of note. First, we assumed unlimited processing capacity, in the sense that we assumed the same rate of evidence accumulation for each position regardless of whether the other position contained a target or not (see Eidels, Donkin et al. (2010) and Eidels, Townsend et al. (2010), for a demonstration of the relationship between capacity and evidence accumulation rates). SFT provides a non-parametric RT-based method of testing this assumption, which was applied to Eidels et al. (2015) data from their super-threshold experiment. For the OR task they found all participants had more capacity than if the two signals shared a fixed pool, although it was a little less than unlimited. In the AND condition, in contrast, there were strong individual differences: four participants had severely limited capacity (i.e., less than fixed), one was intermediate, and four had super-capacity (i.e., more resources for two than one target), at least for slower responses. Although these findings might cast some doubt on our unlimited-capacity assumption, the fact that we mostly obtained good fits provides some reassurance that it is reasonable for Eidels et al.'s near-threshold experiment.

A second limitation relates to the misfit observed for participant LB in the AND condition. This misfit underlines the point that evidence accumulation models such as the LBA, are constrained to only accommodate a limited range of patterns of behavior (Heathcote et al., 2014), and that the same is true of our general framework. Although this is a good characteristic of our approach, because it allows critical evaluation via goodness-of-fit, it also raises the question of how the observed misfit might be explained. There are many possibilities: data may have not conformed to the unlimited capacity assumption, participants may have broken the response rules in a different way to that which we assumed (e.g., applied an OR rule in the AND condition), their data may have been contaminated by mind wandering (e.g., Mittner et al., 2014) or the assumptions made by the LBA may be inappropriate.

Fortunately, if misfit were to occur on a sufficiently wide-spread basis in future data, the general framework which we have proposed can be elaborated to accommodate such possibilities, including incorporation of alternatives to the LBA (e.g., Heathcote & Love, 2012; Leite & Ratcliff, 2010; Logan, Van Zandt, Verbruggen, & Wagenmakers, 2014; Terry et al., 2015; Van Zandt, Colonius, & Proctor, 2000). Similarly, extensions of the framework to the integration of information from more than two sources, and to perform hierarchical Bayesian rather than maximum-likelihood estimation (e.g., where data per participant are limited), are relatively straightforward. In order to facilitate applications of our framework for understanding information integration we provide in Supplementary Materials the code required to obtain likelihoods in the open-source R language (R Core Team, 2015), along with examples of how to perform maximum-likelihood fitting to the data we examined in the present paper.

6. Concluding remarks

We developed a comprehensive parametric framework for identifying important processing attributes such as architecture (in particular, parallel-independent vs. coactive) and stopping rule (exhaustive vs. minimum time). The approach complements existing non-parametric tools that focused exclusively on either RT (Miller's, 1982; Townsend & Nozawa, 1995, race model inequality) or accuracy (the no response probability contract, Eidels et al., 2015; Mulligan & Shaw, 1980), but which do not combine the two dependent measures (see also Townsend & Altieri, 2012, and Townsend, Houpt, & Silbert, 2012, for other approaches that combine these measures with respect to a single aspect of processing). Our approach uses an evidence accumulation model, LBA (Brown & Heathcote, 2008), as a basic building block, and offers a principled way for combining a number of these building blocks to construct a system capable of complex decisions. In this paper we used combinations that form coactive and independent logical-rule process models, using either OR or AND rules, but future studies could form almost an endless number of combinations to suit many tasks and conditions.

The approach we presented offers a number of advantages over alternatives: (i) simultaneous fitting of OR and AND decisions (given the right empirical data), (ii) consideration of speed-accuracy trade-off and, related, (iii) decomposition of complex decisions to latent variables, such as evidence accumulation rate, bias, and non-decision processes. Finally, (iv) it allowed a "rule-breaking" parameter, to indicate whether participants were conforming to the tasks' demands. The results with respect to the latter were quite marked. Participants in the data we analysed consistently ignored the rule dictated by the OR task. Eidels et al.'s (2015) dot-detection task requires divided attention to two spatial locations, each of which could contain a target. Yet, with some frequency, participants processed only one location and not the other, despite the fact such behavior leads to reduced accuracy.

The latter result has profound theoretical implications for perceptual judgments and models of complex decision-making. The prevalent view in simple decision-making is that errors are intrinsic to the decision process. In evidence accumulation models, such as the LBA or Ratcliff's (1978) diffusion model, errors occur when a noisy accumulation process reaches threshold associated with an incorrect response. The exact characteristics of these errors, and how models could produce errors that are either slower than correct responses on some instances or faster on other instances had been a matter of close scrutiny (e.g., Ratcliff & Rouder, 1998). Our modelling results reveal an interesting phenomenon, another source of errors driven by participants not fully conforming to the task instructions. Such failures in processing have been noted in other complex decision tasks (e.g., "Failure to Engage" in task switching: de Jong, 2000; Poboka, Karayanidis, & Heathcote, 2014). The broader implication is, therefore, that such failures should be more widely considered in extending models of simple decision making to more complicated choices.

Acknowledgments

This research was funded in part by the NSF Brain-Body-Environment Systems IGERT grant to Randy Beer and in part by the Faculty Research Support Program, administered through the Office of the Vice President of Research at Indiana

University, Bloomington, and an Australian Research Council (ARC) Professorial Fellowship for Heathcote and ARC Discovery Project grants for Eidels and Heathcote.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cogpsych.2017.03.001.

References

Alais, D., & Burr, D. (2004). No direction-specific bimodal facilitation for audiovisual motion detection. *Cognitive Brain Research, 19*(2), 185–194.
Algom, D., Eidels, A., Hawkins, R. X., Jefferson, B., & Townsend, J. T. (2015). Features of response times: Identification of cognitive mechanisms through mathematical modeling. *The Oxford Handbook of Computational and Mathematical Psychology*, 63–98.
Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*, 700–765.
Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*, 153–178.
Bushmakin, M. A., & James, T. W. (2014). The influence of feature conjunction on object inversion and conversion effects. *Perception, 43*(1), 31–42.
Cousineau, D., & Shiffrin, R. M. (2004). Termination of a visual search with large display size effects. *Spatial Vision, 17*(4), 327–352.
de Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 357–376). Cambridge, MA: MIT Press.
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.
Eidels, A., Donkin, C., Brown, S. D., & Heathcote, A. (2010a). Converging measures of workload capacity. *Psychonomic Bulletin & Review, 17*, 763–771.
Eidels, A., Townsend, J. T., & Algom, D. (2010b). Comparing perception of Stroop stimuli in focused versus divided attention paradigms: Evidence for dramatic processing differences. *Cognition, 114*(2), 129–150.
Eidels, A., Townsend, J. T., Hughes, H. C., & Perry, L. A. (2015). Evaluating perceptual integration: Uniting response-time-and accuracy-based methodologies. *Attention, Perception, & Psychophysics, 77*, 659–680.
Fific, M., & Buckmann, M. (2013). Stopping Rule Selection (SRS) theory applied to deferred decision making. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Cooperative minds: Social interaction and group dynamics. Proceedings of the 35th annual conference of the cognitive science society* (pp. 2273–2278). Austin, TX: Cognitive Science Society.
Fific, M., Little, D. R., & Nosofsky, R. M. (2010). Logical-rule models of classification response times: A synthesis of mental-architecture, random-walk, and decision-bound approaches. *Psychological Review, 117*(2), 309–348.
Fific, M., Nosofsky, R. M., & Townsend, J. T. (2008). Information-processing architectures in multidimensional classification: A validation test of the systems factorial technology. *Journal of Experimental Psychology: Human Perception and Performance, 34*(2), 356–375.
Fifić, M., Townsend, J. T., & Eidels, A. (2008). Studying visual search using systems factorial methodology with target—Distractor similarity as the factor. *Perception & Psychophysics, 70*, 583–603.
Heathcote, A., Brown, S. D., & Wagenmakers, E.-J. (2014). The falsifiability of actual decision-making models. *Psychological Review, 121*, 676–678.
Heathcote, A., Coleman, J. R., Eidels, A., Watson, J. M., Houpt, J., & Strayer, D. L. (2015). Working memory's workload capacity. *Memory & Cognition, 43*(7), 973–989.
Heathcote, A., & Love, J. (2012). Linear deterministic accumulator models of simple choice. *Frontiers in Psychology, 3*.
Houpt, J. W., & Townsend, J. T. (2010). The statistical properties of the Survivor Interaction Contrast. *Journal of Mathematical Psychology, 54*(5), 446–453.
Jones, M., & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modeling schemes for choice reaction time. *Psychological Review, 121*(1), 1–32.
Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics, 72*(1), 246–273.
Little, D. R., Nosofsky, R. M., & Denton, S. E. (2011). Response-time tests of logical-rule models of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(1), 1.
Little, D. R., Nosofsky, R. M., Donkin, C., & Denton, S. E. (2013). Logical rules and the classification of integral-dimension stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(3), 801–820.
Logan, G. D., Van Zandt, T., Verbruggen, F., & Wagenmakers, E.-J. (2014). On the ability to inhibit thought and action: General and special theories of an act of control. *Psychological Review, 121*(1), 66–95.
McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746–748.
Miller, J. (1978). Multidimensional same–different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 411–422.
Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology, 14*, 247–279.
Mittner, M., Boekel, W., Tucker, A., Turner, b., Heathcote, A., & Forstmann, B. U. (2014). When the brain takes a break: A model-based analysis of mind wandering. *Journal of Neuroscience, 34*(49), 16286–16295.
Moneer, S., Wang, T., & Little, D. R. (2016). The processing architectures of whole-object features: A logical-rules approach. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 1443–1465.
Mordkoff, J. T., & Danek, R. H. (2011). Dividing attention between color and shape revisited: Redundant targets coactivate only when parts of the same perceptual object. *Attention, Perception, & Psychophysics, 73*, 103–112.
Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 520.
Mulligan, R. M., & Shaw, M. L. (1980). Multimodal signal detection: Independent decisions vs. integration. *Perception & Psychophysics, 28*, 471–478.
Poboka, D., Karayanidis, F., & Heathcote, A. (2014). Extending the failure-to-engage theory of task switch costs. *Cognitive Psychology, 72*, 108–141.
R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. (2014). The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1226–1243.
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*, 347–356.
Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136*(3), 414–429.
Starns, J. J., Ratcliff, R., & McKoon, G. (2012). Evaluating the unequal- variability and dual-process explanations of zROC slopes with response time data and the diffusion model. *Cognitive Psychology, 64*, 1–34.
Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*, 652–654.
Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., ... Wallace, M. T. (2014b). Identifying and quantifying multisensory integration: A tutorial review. *Brain Topography, 27*(6), 707–730.
Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., & Wallace, M. T. (2014a). Evidence for diminished multisensory integration in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(12), 3161–3167.
Terry, A., Marley, A. A. J., Barnwal, A., Wagenmakers, E.-J., Heathcote, A., & Brown, S. D. (2015). Generalising the drift rate distribution for linear ballistic accumulators. *Journal of Mathematical Psychology, 68*, 49–58.
Townsend, J. T., & Altieri, N. (2012). An accuracy–response time capacity assessment function that measures performance against standard parallel predictions. *Psychological Review, 119*, 500–516.
Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge University Press.
Townsend, J. T., & Colonius, H. (1997). Parallel processing response times and experimental determination of the stopping rule. *Journal of Mathematical Psychology, 41*(4), 392–397.
Townsend, J. T., & Eidels, A. (2011). Workload capacity spaces: A unified methodology for response time measures of efficiency as workload is varied. *Psychonomic Bulletin & Review, 18*(4), 659–681.
Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology, 56*, 476–494.
Townsend, J. T. (1974). Issues and models concerning the processing of a finite number of inputs. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Potomac, Md: Erlbaum.
Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39*(4), 321–359.
Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review, 7*(2), 208–256.
Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11*(1), 192–196.
Williams, L. E., Light, G. A., Braff, D. L., & Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia, 48*(10), 3128–3136.
Wuerger, S. M., Hofbauer, M., & Meyer, G. F. (2003). The integration of auditory and visual motion signals at threshold. *Perception & Psychophysics, 65*(8), 1188–1196.