



Evidence Accumulation Models: Current Limitations and Future Directions

Nathan J. Evans^a and Eric-Jan Wagenmakers^a

^aDepartment of Psychology, University of Amsterdam, The Netherlands

Abstract ■ Evidence accumulation models (EAMs) have been the dominant models of speeded decision-making for several decades. These models propose that evidence accumulates for decision alternatives at some rate, until the evidence for one alternative reaches some threshold that triggers a decision. As a theory, EAMs have provided an accurate account of the choice response time distributions in a range of decision-making tasks, and as a measurement tool, EAMs have provided direct insight into how cognitive processes differ between groups and experimental conditions, resulting in EAMs becoming the *standard paradigm* of speeded decision-making. However, we argue that there are several limitations to how EAMs are currently tested and applied, which have begun to limit their value as a standard paradigm. Specifically, we believe that a theoretical plateau has been reached for the level of explanation that EAMs can provide about the decision-making process, and that applications of EAMs have started to become restrictive and of limited value. We provide several recommendations for how researchers can help to overcome these limitations. As a theory, we believe that EAMs can provide further value through being constrained by sources of data beyond the standard choice response time distributions, being extended to the entire decision-making process from encoding to responding, and having the random sources of variability replaced by systematic sources of variability. As a measurement tool, we believe that EAMs can provide further value through being a default method of inference for cognitive psychology in place of mean response time and choice, and being applied to a broader range of empirical questions that better capture individual differences in cognitive processes.

Keywords ■ Put keywords here, in a comma separated list.

nathan.j.evans@uon.edu.au

[10.20982/tqmp.16.2.p073](https://doi.org/10.20982/tqmp.16.2.p073)

Acting Editor ■
Cheng-Ta Yang
(Department of Psychology, National Cheng Kung University)

Reviewers
■ **James T. Townsend**
(Indiana University)

■ and an anonymous reviewer

Introduction

Evidence accumulation models (EAMs) have been the dominant models of decision-making for several decades (see Ratcliff, Smith, Brown, & McKoon, 2016, for a modern review) for a modern review). Generally, EAMs propose that the evidence in favour of each decision alternative is accumulated at some rate (known as the “drift rate”), until the evidence for one alternative reaches some level of evidence that triggers a decision (known as the “decision threshold”). The general EAM framework has spawned several variants of models, with these proposals either suggesting differing theoretical accounts of how the decision-making process operates (Ratcliff, 1978; Bussemeyer & Townsend, 1993; Ratcliff & Rouder, 1998; Usher & McClelland, 2001; Brown, Marley, Donkin, & Heathcote, 2008; Ver-

donck & Tuerlinckx, 2014; Tillman & Logan, 2017), or attempting to simplify the process to create measurement tools that are easy to apply to related theoretical questions (Brown & Heathcote, 2005; Wagenmakers, van der Maas, & Grasman, 2007; Brown & Heathcote, 2008; Wagenmakers, van der Maas, Dolan, & Grasman, 2008; Grasman, Wagenmakers, & van der Maas, 2009). EAMs have also differed in whether information is accumulated at discrete points in time (e.g., LaBerge, 1962; Audley & Pike, 1965; Link & Heath, 1975; Link, 1975; Townsend & Ashby, 1983; Smith & Vickers, 1988; Smith & Van Zandt, 2000) or continuously (e.g., Stone, 1960; Ratcliff, 1978; Usher & McClelland, 2001; Brown & Heathcote, 2008; Tillman & Logan, 2017); however, our article primarily focuses on continuous EAMs, such as Ratcliff and colleague’s extensions to the Wiener diffusion model (Ratcliff, 1978; Ratcliff & Rouder, 1998; Rat-



cliff & Tuerlinckx, 2002), as these variants are the most commonly applied EAMs within modern decision-making research.

Our article discusses the future of EAMs within psychological research, both in terms of potential future directions within the field of decision-making, and their more general role within cognitive psychology as a whole. Specifically, we begin by providing a brief discussion of the recent history of EAMs, and their success both as theories (i.e., providing accurate accounts of empirical data) and as useful measurement tools (i.e., estimating latent variables from data). After this, we discuss the role of EAMs in the broad field of cognitive psychology, and whether EAMs provide a good *standard model* for the field that researchers can use as a benchmark and building-block for more complex models that aim to explain cognition at a general level. Lastly, we discuss some important potential future directions of EAMs – developments which we believe would either allow EAMs to further increase our understanding of the decision-making process, or make EAMs a better measurement tool for making inferences on latent variables within empirical data.

Past Success

As a theory, EAMs provide a detailed account of how humans make decisions. Specifically, EAMs provide predictions for both the response choice and time of each decision in an experiment, meaning that they provide a process-level explanation of the two variables of interest in most rapid decision-making experiments. Although several different variants of EAMs exist, most are able to provide a close quantitative fit to these choice response time distributions observed in a range of different paradigms, and have been shown to qualitatively capture a range of different choice and response time benchmarks (Ratcliff, 1978; Ratcliff & Rouder, 1998; Usher & McClelland, 2001; Ratcliff & Tuerlinckx, 2002; Brown & Heathcote, 2008; Verdonck & Tuerlinckx, 2014). These models have also served as the basis for extensions to explain the choice response time distributions in more complex decisions, such as categorization (Nosofsky & Palmeri, 1997; Nosofsky, Little, Donkin, & Fific, 2011), multi-attribute choice (Roe, Busemeyer, & Townsend, 2001; Usher & McClelland, 2004; Tsetos, Usher, & Chater, 2010; Trueblood, Brown, & Heathcote, 2014), absolute identification (Brown et al., 2008), choice confidence (Van Zandt & Maldonado-Molina, 2004; Ratcliff & Starns, 2009; Pleskac & Busemeyer, 2010), and stop signal paradigms (Matzke, Love, & Heathcote, 2017; Matzke, Hughes, Badcock, Michie, & Heathcote, 2017). Overall, EAMs are one of the most successful frameworks in the history of cognitive psychology, providing an accurate account of data from a range of rapid decision-making tasks,

and serving as a basis for extensions to more complex decisions.

As a measurement tool, EAMs have helped to answer a range of theoretically relevant questions within cognitive psychology. Specifically, EAMs are able to decompose the choice response time distributions into latent variables of the decision-making process, such as the drift rate and the decision thresholds (see Dutilh et al., 2018) for the consistency between different models and methods). The drift rate reflects both people's task ability and the general ease of the task, whereas the decision threshold reflects how cautious people are in their decision strategy. Many cognitive theories can be assessed through the EAM framework based on their predictions for how these latent variables should change over experimental conditions or groups. For example, a robust finding within the ageing literature is that older adults are slower at many cognitive tasks than younger adults (e.g., Brinley, 1965). The primary theoretical explanation given for older adults having slower performance than younger adults was that people undergo a *cognitive slowdown* as they age (Salt-house, 1996), which results in a decrease in mental processing speed. Ratcliff, Thapar, and McKoon (2001) tested this theory using the EAM framework, where the cognitive slowdown account would predict that older adults have a lower drift rate than younger adults. Interestingly, Ratcliff et al. (2001) found that in several tasks younger and older adults had very similar drift rates, and the slower performance of older adults was actually the result of greater caution in responding (i.e., higher decision threshold) and slower perceptual/motor processes (i.e., higher non-decision time), showing evidence against the cognitive slowdown theory through the EAM framework. EAMs have been able to answer similarly posed questions in a range of different paradigms, such as letter identification (Ratcliff & Rouder, 2000), lexical decision-making (Wagenmakers, Ratcliff, Gomez, & McKoon, 2008), sentence comprehension (Lerche, Christmann, & Voss, 2019), genetic heritability (Evans, Steyvers, & Brown, 2018), intelligence testing (Ratcliff, Thapar, & McKoon, 2010), recognition memory (Ratcliff, 1978), personality (Evans, Rae, Bushmakin, Rubin, & Brown, 2017), early life adversity (Knowles, Evans, & Burke, 2019), and performance optimality (Starns & Ratcliff, 2012; Evans & Brown, 2017; Evans, Bennett, & Brown, 2018).

Current Role in Cognitive Psychology

In the book *The Structure of Scientific Revolutions*, Thomas Kuhn suggested the scientific progress proceeds in two separate types of phases: *normal science* and *scientific revolutions* (Kuhn, 1962). The first type of phase, *normal science*, occurs when there is a dominant theory within the



field that forms the *standard paradigm*. This standard paradigm is the basis for how researchers explain and conceptualize the phenomenon of interest (e.g., the textbook explanation of a process), and research questions within the field are often answered through the assumptions contained within the standard paradigm. In our opinion, the field of decision-making currently appears to closely resemble Kuhn's concept of normal science, with EAMs being the standard paradigm. Specifically, the noisy integration process of EAMs has had a fundamental impact on how we conceptualize information processing – the cornerstone of cognitive psychology – as a whole. Decision-making is generally thought about within the EAM framework, where EAMs form the theoretical basis for a range of extensions, and researchers often make inferences on the latent variables estimated through EAMs. The overwhelming amount of scientific progress in decision-making – and in some cases, cognitive psychology – made through the assumptions of the EAM framework reflects the process of *normal science*, where EAMs have provided a useful basis for many interesting scientific discoveries (as discussed in the previous section on “Past Success”).

However, the second type of phase, *scientific revolutions*, occur when a standard paradigm becomes decreasingly useful, and a new paradigm rises to replace it. This is often due to increasing evidence against the accuracy of the paradigm, and increasing evidence in favour of another paradigm, resulting in the new paradigm becoming a superior explanation of the phenomenon than the old paradigm. However, a paradigm could also become a candidate for a revolt for other reasons, such as no longer generating testable predictions, becoming increasingly impractical and difficult to work with, or making overly general predictions that allow for the possibility of phenomena that do not occur. Considering the concept of scientific revolutions leads to several important questions surrounding EAMs and their usage within decision-making and cognitive psychology. Do they still provide an adequate account of decision-making phenomena? Can they continue to advance our understanding of human decision-making? Are they the basis that we should be using for testing and extending cognitive theories? Most generally, should EAMs remain the standard paradigm for decision-making, and to a broader extent, cognitive psychology, or should the future of cognitive psychology look elsewhere, in a scientific revolution to a new model and paradigm? Here, we will discuss the reasons in favour of keeping EAMs as the standard paradigm, the reasons against keeping EAMs as the standard paradigm, and what the outcome would be of abandoning EAMs as the standard paradigm.

Benefits of EAMs as the standard paradigm

EAMs provide several benefits for the field of speeded decision-making, and more broadly cognitive psychology, both in terms of their explanatory power and their usefulness in understanding related cognitive processes. These benefits have likely led to their status as the standard paradigm for the field of speeded decision-making. Below we detail some of the advantages of EAMs, which make a case for why they should remain the standard paradigm of speeded decision-making, and the basis for how we understand information processing.

An impressive explanatory benchmark

EAMs have provided one of the best explanations of the decision-making process to date (Ratcliff et al., 2016), which can provide intuitively and neurally plausible explanations for how information processing may operate (Usher & McClelland, 2001). Although many models exist that can explain the average tendencies in response choice (i.e., response proportion) and response time (i.e., mean response time of correct responses) – summary statistics that are relatively easy for models to capture (Luce, 1986) – only EAMs have been able to provide an accurate account of the correct and error response time distributions and their relationship across a range of cognitive tasks (Ratcliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999; Usher & McClelland, 2001). In addition to providing a close quantitative account of the choice response time distributions, EAMs are also able to explain a range of benchmark phenomena in choice response time. At the most basic level, EAMs are able to explain the speed-accuracy tradeoff – where some people perform the task more quickly at the expense of accuracy, and others are slower while remaining more accurate – through differences in the amount of response caution that participants have. At a more detailed level, EAMs can explain (1) the ubiquitous positive skew found in human response time distributions, (2) the relation between the mean and variance in response time, (3) the mean response time for error responses being slower than for correct responses in most situations, and (4) the mean response time for correct responses being slower than for error responses in some situations, such as when quick responding is emphasised. Overall, EAMs appear to provide an impressive account of decision-making behaviour and a clear benchmark for future alternative decision-making frameworks: serious contenders for explaining the decision-making process should be able to explain the entire choice response time distributions with at least the level of accuracy as seen from EAMs. This suggests that EAMs continue to function as an adequate standard paradigm for decision-making, as their underlying architecture continues to provide a accurate explanation of the



key trends within empirical data.

A common underpinning for all tasks

EAMs provide one of the most broadly applicable formalized frameworks in cognitive psychology. EAMs describe a general theoretical process where decisions are made, allowing them to theoretically extend to any cognitive task that requires a decision to be made, such as those in the areas of learning (Evans, Brown, Mewhort, & Heathcote, 2018), memory (Ratcliff, 1978; Osth & Farrell, 2018), language processing (Wagenmakers, Ratcliff, et al., 2008; Lerche et al., 2019), and consumer choice (Trueblood et al., 2014; Evans, Holmes, & Trueblood, 2019), making them an excellent standard paradigm for cognitive psychology. However, EAMs generality also means that they only provide a vague explanation of the process involved in any specific cognitive task – similar to signal detection theory (Green & Swets, 1966) – meaning that they do not provide the in-depth explanations of specific cognitive tasks that less general, task-specific theories can. Researchers have overcome the vagueness of EAMs by creating specific extensions of the framework for different cognitive tasks. Importantly, the EAM framework is simple to extend to more task-specific models, where researchers can define a task-specific function – often based on task-specific models – which takes stimulus information and transforms it into the drift rates that feed into the EAM framework for the different alternatives. These are commonly referred to as *front-end* models, as the task-specific function acts as an initial filter for the stimulus information before it feeds into the *back-end* EAM process. Front-end models provide a more constrained model of the task, as the drift rates are now directly constrained by the stimulus information rather than being free parameters, and can also provide a more specific explanation of the process, as the creation of the front-end function can be guided by task-specific theories. Front-end models have also proved successful in accounting for empirical data in a range of specific cognitive tasks, such as the exemplar-based random walk in categorization (Nosofsky & Palmeri, 1997) and memory (Nosofsky et al., 2011, see also the exemplar-based linear ballistic accumulator; Donkin & Nosofsky, 2012; Osth, Jansson, Dennis, & Heathcote, 2018; Cox & Shiffrin, 2017, and several other front-end extensions in memory research), the multi-attribute linear ballistic accumulator in multi-attribute choice (Trueblood et al., 2014), and RTCON in choice confidence (Ratcliff & Starns, 2009). This suggests that EAMs provide a suitable standard paradigm for cognitive psychology, being both sufficiently general to be applicable to all tasks, and simple enough to easily combine with task-specific theories for more detailed explanations of specific cognitive tasks.

A well-developed tool for application

EAMs have become one of the most developed modelling frameworks within cognitive psychology, which is likely the result of their ability to decompose observed variables into latent parameters, allowing for meaningful psychological conclusions to be drawn from data within cognitive psychology. As discussed previously, several simple EAM variants have been developed that can either be solved in closed form (Wagenmakers et al., 2007), estimated using simple methods of moments (Grasman et al., 2009), or have an analytically tractable likelihood function (Brown & Heathcote, 2008). These simple variants provide quick and easy methods for researchers to apply EAMs to estimate the latent variables of the decision-making process, and use them to answer theoretically relevant questions about human cognition. Even for the more complex models, such as the diffusion model (Stone, 1960) with between-trial variability parameters for drift rate (Ratcliff, 1978), starting point (Ratcliff & Rouder, 1998), and non-decision time (Ratcliff & Tuerlinckx, 2002) – commonly referred to as the *full* diffusion model – several frameworks have been developed that perform the fitting process for researchers, such as *fast-dm* (Voss & Voss, 2007), *DMAT* (Vandekerckhove & Tuerlinckx, 2008), *HDDM* (Wiecki, Sofer, & Frank, 2013), and *DMC* (Heathcote et al., 2018). Importantly, these EAM fitting frameworks allow researchers with basic programming skills to easily implement complex EAMs with state-of-the-art methods, even if they only have a rudimentary understanding of EAMs and these methods, meaning that researchers can focus on learning how to correctly interpret the estimated latent variables (see Dutilh et al., 2018, for a comparison of several methods). These frameworks make EAMs an extremely useful standard paradigm, being accessible to researchers who are not experts at applying these models, allowing for broad applications to answer a range of theoretically relevant questions in a range of different tasks.

Drawbacks of EAMs as the standard paradigm

Although there are many benefits to having EAMs as the standard paradigm of decision-making, there are also drawbacks. Specifically, despite EAMs initially providing rapid progress in our understanding of both decision-making and the related areas of cognitive psychology that EAMs were applied to, this progress has begun to reach a plateau, where EAMs are providing fewer advances to our understanding of cognitive processes. Note that we are not attempting to claim that research involving EAMs is devoid of advances in our understanding of cognitive processes, and we cite many examples within our “Future Directions for EAMs as Theories” and “Future Directions for EAMs as Measurement Tools” sections where we be-



lieve that our understanding of cognitive processes have continued to flourish through EAMs. However, we believe that the overall progress of the field has begun to reach a plateau, and we hope that discussing these issues will galvanize researchers to explore more novel directions, resulting in the rate of progress once again becoming more rapid. Below we detail some of the disadvantages of EAMs, which create some potential reasons for why researchers may consider searching for alternative frameworks to replace EAMs as the standard paradigm of decision-making and information processing.

A theoretical plateau has been reached

EAMs have provided researchers with an explanation of the general decision-making process, with EAMs proposing that speeded decision-making is the result of evidence accumulating for each alternative until one reaches a threshold. However, progress beyond this general explanation of decision-making has been limited, and few insights have been gained into the more specific dynamics of the decision-making process. Several different variants of EAMs exist, which each propose different specific dynamics that occur during the accumulation process. For example, the leaky-competing accumulator (LCA; Usher & McClelland, 2001) proposes that different alternatives *laterally inhibit* one another, where alternatives with more accumulated evidence lessen the accumulation for other alternatives, whereas the diffusion model (Ratcliff, 1978) proposes that alternatives interact through *feed-forward inhibition*, where evidence for one alternative is evidence against the other alternatives. However, these different theoretical accounts display a high level of mimicry, making it difficult to assess which specific dynamics provide the best explanation of decision-making process (Leite & Ratcliff, 2010; Teodorescu & Usher, 2013). Specifically, most EAMs provide an accurate account of the choice response time distributions of most cognitive tasks, making their predictions difficult to distinguish between in choice and response time data. EAMs also make several assumptions that seem implausible based on research in related fields. For example, EAMs assume that the decision-making and motor processes occur sequentially, where motor responding only begins after a response has been chosen, and the decision-making process has terminated (Ratcliff, 1978; Usher & McClelland, 2001; Brown & Heathcote, 2008). However, electromyography (EMG) data – where participants response movements are recorded – has indicated that the decision and motor processes are intertwined, with decisions commonly involving people initially beginning a response for one alternative, before stopping and responding for the other alternative (Coles, Gratton, Bashore, Eriksen, & Donchin, 1985; Schweickert, 1989; Townsend & Fikes,

1995; Burle, Spieser, Servant, & Hasbroucq, 2014; Servant, White, Montagnini, & Burle, 2015; Servant, White, Montagnini, & Burle, 2016). These EMG findings suggest that EAMs currently provide an inadequate explanation of the complete decision-making process, though little effort has been directed towards developing variants that do not assume decision and motor processes occur sequentially, as this assumption does not prevent EAMs from providing an accurate account of the data that they are usually fit to: the choice response time distributions (though see Servant et al., 2016, for an example of how EMG data can be integrated into the EAM framework). This suggests that little progress is currently being made in understanding because of the fixation of decision-making researchers on only choice response time distributions, resulting in a theoretical plateau.

Applications are often restrictive

EAMs have been a useful measurement tool for answering theoretically relevant questions within cognitive psychology. As discussed earlier in the example from the ageing literature, the EAM framework can be used to test theories that make predictions about the latent variables of the decision-making process, such as drift rate or threshold, instead of attempting to infer changes in these latent variables from changes in observed variables (e.g., mean response time). Although the EAM framework has the potential to answer a wide variety of theoretically interesting questions, EAM applications are often restrictive. Specifically, we argue that most applications of EAMs fall into a category colloquially referred to as “EAM account of task X”, where researchers (1) find a cognitive task that EAMs are not commonly applied to, but response time and choice can be measured in, (2) fit an EAM to data from this task, and (3) estimate what parameters vary over experimental conditions or groups (e.g., Leite & Ratcliff, 2011; Yap, Balota, Sibley, & Ratcliff, 2012; Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012; Philiastides & Ratcliff, 2013; Gomez, Perea, & Ratcliff, 2013; van Ravenzwaaij, Boekel, Forstmann, Ratcliff, & Wagenmakers, 2014; Ratcliff & Strayer, 2014; Ratcliff & Smith, 2015; Gomez, Ratcliff, & Childers, 2015; Thompson, Ratcliff, & McKoon, 2016; Aschenbrenner, Balota, Gordon, Ratcliff, & Morris, 2016). These applications lead to inferences about (1) whether or not the chosen EAM provided a good account of the data, and (2) what parameters varied over the conditions or groups, and conclusions about (1) the ability of EAMs to explain data in this task, and (2) how “manipulation Y causes an effect in parameter Z”. Although these applications can provide interesting information about the task and provide theoretically relevant inferences about cognitive processes, they represent a small subset of the poten-



tial questions that EAMs could be used to answer; something that we discuss in more detail within our “Future Directions for EAMs as Measurement Tools” section. Although recent research has attempted extend EAM applications to new realms, such as *joint modelling* approaches that attempt to link multiple sources of data (Turner et al., 2013; Evans, Rae, et al., 2017; Turner, Rodriguez, Norcia, McClure, & Steyvers, 2016; Evans, Steyvers, & Brown, 2018; Turner, Van Maanen, & Forstmann, 2015; Knowles et al., 2019; Turner, Wang, & Merkle, 2017; Krajbich, Armel, & Rangel, 2010), these applications appear to be exception, rather than the rule. This suggests that the EAM framework may encourage researchers to answer restrictive questions, with EAM applications mostly revolving around the question “does drift rate or threshold vary between these conditions or groups?”.

What alternatives exist?

As discussed above, there are several benefits and drawbacks to having the EAM framework as the standard paradigm of decision-making. However, before considering the removal of EAMs as the standard paradigm, there must be a suitable framework to replace it. Importantly, this new framework would need to be able to account for the broad range of empirical data as well as EAMs, and also provide the same usefulness as EAMs in answering theoretically relevant questions that test cognitive theories. However, as far as we are aware, there are currently no other modelling frameworks that are able to match EAMs in either their explanatory power or their broad applicability. In addition, even if another modelling framework was developed in the not-too-distant future that could improve on EAMs in these areas, the framework would also need to develop the same level of infrastructure that currently exists for EAMs, such as the many software packages available for fitting these models. Therefore, EAMs appear to still be the most suitable framework for the standard paradigm of speeded decision-making, and will likely to continue to be for at least the next decade of decision-making research.

Although we believe EAMs are still the most suitable standard paradigm of decision-making, we also believe that there are two key problems that are the result of how EAMs are currently used: that we have reached a theoretical plateau in our understanding of the decision-making process, and that the EAM applications are often restricted to a narrow range of questions. In the remainder of this article, we discuss different possible future directions for EAMs that could help overcome each of these limitations. However, we note that EAMs are not necessarily the best possible framework, and that future research should dedicate more time and resources to attempting to develop

alternative frameworks to EAMs, which could challenge their position as the standard paradigm.

Future Directions for EAMs as Theories

One of the key issues we highlighted with EAMs in the previous section is that recent theoretical progress in understanding the decision-making process has reached a plateau. Although the general EAM framework appears to provide a good description of the general decision-making process, gaining a more intricate understanding of the dynamics of the process has become difficult due to the high levels of mimicry between the models in explaining the choice response time distributions. Here, we attempt to provide several future directions for EAMs that would allow further theoretical insight into the decision-making process, allowing EAMs to continue to increase our understanding of how the human decision process operates.

Moving beyond response time and choice

The field of decision-making – and in many cases, cognitive psychology – has been fixated on two sources of behavioural data: the response choices that people make, and the response times that they make them at. As discussed earlier, EAMs are able to account for these observed variables in unison, which allows them to capture the speed-accuracy tradeoff. EAMs are also able to capture the distribution of choices and response times over an entire experiment, which is more difficult to accurately predict than the summary statistics of mean response time and accuracy (Luce, 1986). The fixation on response time and choice is likely a result of their sufficiency in teasing apart the effects of task ease (i.e., drift rate) and task caution (i.e., decision threshold), as the choice response time distributions alone can be decomposed into the latent parameters of the decision-making process. However, as discussed previously, EAMs that propose different specific dynamics for the decision-making process display a high level of mimicry when assessed in their ability to account for the choice response time distributions, meaning that the choice response time distributions do not provide adequate constraint to distinguish between different theoretical perspectives of decision-making. Therefore, from the theory-based perspective of trying to understand the specific dynamics of the decision-making process, it appears that we may have reached the limit of what response time and choice alone can tell us, and it may be time to move beyond the choice response time distributions and begin forcing these models to account for additional sources of data.

However, if researchers are going to constrain EAMs to account for additional sources of data, then what sources of data should they use? Ideally, any additional source



of data should still allow response time and choice to be recorded, so the models can be constrained by all of these variables simultaneously, and the additional source of data should clearly map onto the current EAM framework, so that fairly uncontroversial extensions can be made from the current models to these new sources of data. One possibility is to change the types of decisions that people make, so that they provide a rating on how much they prefer one alternative over another, rather than just a dichotomous choice on which alternative they prefer. This is commonly implemented as a confidence rating (Vickers, 1979; Van Zandt & Maldonado-Molina, 2004; Ratcliff & Starns, 2009; Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2013), where participants decide how confident they are in their chosen alternative (though also see “best-worst choice”; Finn & Louviere, 1992; Hawkins et al., 2014) either during the initial decision (Ratcliff & Starns, 2009, 2013), or in a subsequent decision (Van Zandt & Maldonado-Molina, 2004; Pleskac & Busemeyer, 2010). Choice confidence provides a rich source of information to researchers, providing some measurement of the preference for one response alternative over another, and has led to a range of models being developed to explain these confidence responses, especially within the area of recognition memory (Van Zandt & Maldonado-Molina, 2004; Ratcliff & Starns, 2009; Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2013). However, as far as we are aware, no studies have attempted to use confidence responses to distinguish between the more intricate dynamics of the complete decision-making process. One reason for this may be that recording choice confidence requires changing the type of decisions that participants make, meaning that researchers need to alter their experimental paradigms in order to obtain this information. Although obtaining the additional information is certainly beneficial, it could also be argued that changing the type of decisions that participants make from dichotomous choice to a scale qualitatively changes the way that people make these decisions. Indeed, several models of choice confidence assume that the process is changed to at least some extent, such as *RTCON* (Ratcliff & Starns, 2009), which assumes that each confidence category is now a completely separate response alternative with a separate accumulator. Furthermore, researchers must also be wary of how they measure confidence, as the time of measurement (e.g., during the initial decision [Ratcliff & Starns, 2009] or in a subsequent decision [Pleskac & Busemeyer, 2010]) and the type of measurement (e.g., ordinal [Ratcliff & Starns, 2009] or a bounded continuum [Vaghi et al., 2017]) determines how confidence manifests at an observed level. Therefore, although choice confidence is an interesting additional source of data that can help us better understand certain aspects of decision-making, inferences made

in these paradigms may lose direct applicability to the standard two-alternative forced choice paradigms used in most decision-making tasks.

Another possibility for obtaining additional sources of data is to change the amount of information recorded from the decision-making process, attempting to integrate information from several points in time during the decision. One simple method would be to record additional responses that occur after the initial response is made, which reflect some change in preference (e.g., a “change of mind”; Resulaj, Kiani, Wolpert, & Shadlen, 2009; Yeung, Botvinick, & Cohen, 2004; Hasbroucq, Burle, Akamatsu, Vidal, & Possamai, 2001) in the decision-making process. Rabbitt and colleagues (Rabbitt, 1966, 1967, 1968, 1969; Rabbitt & Rodgers, 1977; Rabbitt, Cumming, & Vyas, 1978; Rabbitt & Vyas, 1981; Maylor & Rabbitt, 1987; Rabbitt, 2002) provided initial investigations of these post-decision responses in isolation, using paradigms with “error-correcting responses”, where participants are explicitly instructed to correct their errors if they believe they have made them. Even when participants are not instructed to correct their errors, second responses for the opposing alternative – something that we termed “double responding” in previous work – are often anecdotally observed, though are rarely studied (Evans, Dutilh, Wagenmakers, & van der Maas, *in press*). Importantly, Evans, Dutilh, et al. (*in press*) showed that double responses can be used to further constrain EAMs, with data from two experiments showing strong evidence for lateral inhibition being required to simultaneously account for the choice response time distributions and double responding behaviour. Post-decision behaviour – such as double responses or error correct responses – provide a source of data that can further constrain the different variants of EAMs and require no changes (or in the case of error correcting responses, minimal changes) to the experimental paradigm. However, future research could aim to obtain even more measurements of the decision-making process, such as measurements at several discrete points in time, or even using continuous measurement.

From encoding to responding

The EAM framework has been focused on explaining the human decision-making process. As discussed earlier, EAMs have provided an accurate account of the choice response time distributions from a range of cognitive tasks, displaying their success in accounting for the decision-making process. However, when thinking more broadly in terms of all of the mental processes required to make a decision, EAMs *only* attempt to explain the point of the process where the response is being chosen, leaving all other aspects of the complete process undefined. For example,



EAMs do not attempt to explain how the information from the stimulus is converted into the accumulated evidence that drives the decision process, how the evidence from the previous decision “resets” for the new decision, or how reaching the decision threshold results in the motor response being made; these components are simply placed into the “non-decision time” bracket. Although being agnostic to the other parts of the complete process helps to make EAMs a computationally tractable and generalizable measurement tool, leaving these parts undefined also means that EAMs cannot provide a complete, mechanistic explanation of the process from stimulus input to response completion. Therefore, although EAMs provide an accurate account of the response selection part of decision-making, they do not provide an explanation for the complete process involved in making a decision.

One problematic consequence of EAMs ignoring the other parts of the complete decision process is that they make the implicit assumption that these sub-processes are *separate* and *independent*. For example, EAMs assume that moment-to-moment samples of evidence are independent, meaning that there is no interaction between encoding and accumulation, and that the motor action does not begin until the response selection process is completed, which cannot be changed once the decision has been triggered. However, these assumptions seem somewhat implausible given research in related areas. For example, the EMG research discussed earlier suggests that people occasionally make initial movements towards one alternative before changing their response to the other alternative, suggesting that people begin their motor responding before the final response is selected. It also seems likely that there can be some level of dependency in the moment-to-moment samples of evidence taken from the environment, as either actual auto-correlation in the environment or perceived auto-correlation by participants, meaning the independent samples assumption seems somewhat implausible. Therefore, it seems that to understand the complete process of decision-making, researchers need to extend EAMs beyond their current explanation of only the response selection process to account for the other sub-processes involved in decisions.

One promising method that could provide a more complete understanding of the process is the recent wave of joint modelling techniques within cognitive modelling. Joint modelling involves creating a single modelling framework that attempts to capture multiple sources of data that would traditionally be assessed separately with separate analyses or modelling frameworks. Joint modelling techniques vary greatly from study to study, such as creating a full covariance matrix across all parameters from two models of two different types of data (Turner et al.,

2013), estimating the correlation between model parameters and a different source of data (Knowles et al., 2019), estimating the correlation between the same parameters across different people or experimental conditions (Evans, Steyvers, & Brown, 2018), or restricting a specific parameter of a model to be a function of another source of data (Evans, Rae, et al., 2017). The types of data also varies greatly from study to study, with previous research having combined EAMs with neural data (Turner et al., 2013; Turner et al., 2016; Turner et al., 2015; Turner et al., 2017), personality data (Evans, Rae, et al., 2017), eye-tracking data (Krajbich et al., 2010), genetics data (Evans, Steyvers, & Brown, 2018), EMG data (Servant et al., 2016) and developmental data (Knowles et al., 2019). Importantly, these methods can be combined with psychophysiological data, such as neural recordings (e.g., single cell recordings, electroencephalography [EEG], functional magnetic resonance imaging [fMRI]) and motor recordings (e.g., EMG), which would provide further insight into how these data that reflect other sub-processes relate to the latent parameters within EAMs. However, these joint modelling techniques would only provide a measurement of the relationship between the different sub-processes that make up a decision, and would not provide a complete, mechanistic model of decision-making.

Another method for gaining a more complete explanation of the decision process is to combine knowledge across fields. Just as the response selection process is well studied and understood within the area of decision-making, perceptual encoding and motor responding are also well studied and understood within their respective fields of research, with theories and models for how these processes operate. However, these fields of research have traditionally been separate from decision-making, where different fields each attempt to understand their part of the process. One clear method for obtaining a more complete understanding of the entire process would be combining these separate research traditions by integrating the different models – and sources of data analysed – in order to create a complete, mechanistic account of the entire decision process. This could potentially be achieved through methods that already exist within cognitive modelling, such as the “front-end” model approaches discussed earlier. Another possible initial step would be building more complex models of the decision-making, which involve interactions between different sub-process and/or processing channels (e.g., Loftus, Busey, & Senders, 1993; Townsend & Wenger, 2004; Wenger & Townsend, 2006; Eidels, Houpt, Altieri, Pei, & Townsend, 2011; Townsend, Houpt, & Silbert, 2012) at different stages of the entire process (e.g., the decision input, decision process, or decision output). These models could still remain within the EAM framework, but



would contain more specific predictions about both the decision-making process itself, and the process that occur before (e.g., perception) and after (e.g., motor) the decision. However, given that the different existing EAM variants already strongly mimic each other in their predictions for choice response time distributions, which often prevents researchers from distinguishing between their different proposals for dynamics of the decision-making process, it seems likely that these increasing complex models would require novel experimental paradigms – or new sources of data, as discussed in the previous subsection – to properly constrain their predictions and make them distinguishable. One other – though more difficult – solution would involve developing EAMs into a complete cognitive architecture (Newell, 1990), similar to the ACT-R (Anderson, 1993; Anderson et al., 2004; Anderson, 2007; ; see Ritter, Tehrani, & Oury, 2019, for a recent review), SOAR (Milnes et al., 1992; Laird, 2008), and EPIC (Kieras & Meyer, 1994; Meyer & Kieras, 1994; Meyer et al., 1995) architectures used in others areas cognitive research. These architectures each aim to provide a universal theoretical framework based on how these cognitive processes may be implemented in the brain, while also allowing the development of sub-theories for specific tasks that researchers may be interested in. Although the development of a complete cognitive architecture based on the EAM framework would likely be a difficult and lengthy process, an easier goal may be attempting to connect EAMs with existing cognitive architectures, which may help to provide EAMs with a more complete explanation of the entire decision-making process.

Explaining “random” variability

Every model within the EAM framework incorporates at least one source of “random” variability within its architecture, either as moment-to-moment fluctuations in parameter values, or decision-to-decision fluctuations in parameter values. These random fluctuations make the process with EAMs stochastic, allowing different responses to be made at different times for identical decisions. Specifically, models have been proposed within the EAM framework that include one or more of the following four types of random variability: (1) moment-to-moment variability in evidence accumulation (Stone, 1960; Laming, 1968; Ratcliff, 1978; Ratcliff & Rouder, 1998; Usher & McClelland, 2001; Ratcliff & Tuerlinckx, 2002), (2) decision-to-decision variability in evidence accumulation (Laming, 1968; Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff & Tuerlinckx, 2002; Brown & Heathcote, 2005, 2008), (3) between decision variability in the amount of starting evidence (Ratcliff & Rouder, 1998; Ratcliff & Tuerlinckx, 2002; Brown & Heathcote, 2005, 2008), and (4) between decision vari-

ability in the time for non-decision processes (Ratcliff & Tuerlinckx, 2002). Each type of random variability is governed by some moment-to-moment or decision-to-decision distribution, with these distributions usually being simple, convenient, analytically tractable distributions, such as the normal, truncated normal, or uniform (though see Terry et al., 2015, for other potential distributions; and Jones & Dzhafarov, 2014, for the importance of distributional assumptions). The number of random variability parameters differs between different EAMs, ranging from the LCA that only contains a single source of random variability – moment-to-moment variability in drift rate – to the full diffusion model that contains all four sources of random variability mentioned above. Importantly, these differences in the number of random variability parameters results in different explanations for benchmark response time phenomena. While the full diffusion model provides a general explanation, suggesting that these benchmark phenomena arise due to random variability between decisions in drift rate and starting point, the LCA provides a more specific explanation about the dynamics of the decision-making process, suggesting that these benchmark phenomena arise due to lateral inhibition between alternatives and leakage of evidence over time.

One of the key reasons for inclusion of random variability parameters is their ability to account for benchmark response time phenomena while remaining computationally tractable. For example, the full diffusion model is able to explain these benchmark phenomena while still retaining a computationally tractable likelihood function, which with modern frameworks can be implemented within a reasonable amount of time (e.g., Voss & Voss, 2007), whereas the addition of lateral inhibition and leakage dynamics make the LCA computationally intractable, meaning that the model is mostly fit through simulation-based methods (Evans, 2019a). However, this computational tractability comes at the expense of explanation. From a theoretical perspective, it seems unlikely that the variability captured by the random variability parameters is actually due to some non-deterministic process, and instead the variability is likely due to sources that are either unknown, difficult to quantify, or difficult to explicitly model. However, the random variability parameters provide little explanation about how the process operates, and what exactly causes these sources of variability (Evans, Tillman, & Wagenmakers, *in press*).

One method to explain these random sources of variability is to measure more factors that we believe may be the sources of the variability. Theoretically, if all factors that contributed to the variability in a parameter – such as drift rate – could be measured, then all variability would be accounted for, and there would be no need to include ran-



dom variability parameters in the model. Some of these factors are easy to capture and are already modelled in most applications of EAMs, such as item effects (e.g., Evans, Hawkins, Boehm, Wagenmakers, & Brown, 2017, where accumulation was constrained to be a linear function of the stimulus coherence). Other factors may be easy to measure, but difficult to incorporate into EAMs, such as sequential effects in decision-making (though see Usher & McClelland, 2001; Brown et al., 2008; Jones, Curran, Mozer, & Wilder, 2013). There are also factors that may be more difficult to measure, such as fluctuations in attention over the course of the experiment (though see Hawkins, Mittner, Boebel, Heathcote, & Forstmann, 2015), which may require novel experimental manipulations to measure and assess their impact on the process. However, regardless of the difficulty, future research should strive to uncover systematic sources of variability that are captured by these random variability parameters, in order to gain a more complete understanding of the decision-making process.

Future Directions for EAMs as Measurement Tools

The other key issue that we highlighted with EAMs is that recent uses of the framework as a measurement tool have become restrictive, and that different approaches may be able to provide new, interesting insights into cognition. Here, we attempt to provide several future directions for EAMs that would allow greater insight into cognition to be gained from applications of the framework.

A standard focus on latent parameters

In most fields within cognitive psychology, research questions focus on whether specific experimental conditions/groups are either (1) faster/slower than one another by assessing the mean correct response time, or (2) more/less accurate than one another by assessing the task accuracy. As discussed previously, these observed variables do not directly translate to latent parameters of the underlying cognitive process, and changes in either of these observed variables could reflect changes in task ability or task caution. These assessments of observed variables can result in misleading conclusions being drawn, such as Salthouse's 1996 cognitive slowdown theory, which then need to be corrected by subsequent "EAM account of task X" studies, such as Ratcliff et al. (2001). Therefore, we argue that the continuing focus of cognitive psychology on observed variables – which are used to make indirect inferences about latent variables – creates a need for subsequent EAM studies (i.e., "EAM account of task X" applications) to properly test the conclusions of these original studies using latent parameters, and distracts researchers in decision-making from exploring other types of questions that may provide interesting and informative results.

One potential method for increasing the number of novel applications of EAMs is to remove the void left by mean response time and accuracy analyses, and make the latent parameters with EAMs the *standard focus for analyses within cognitive psychology*. Specifically, we believe that the latent parameters of EAMs should be made the default variables for analyses within cognitive psychology, where researchers continue to perform the same standard statistical analyses that they have previously, but using the latent parameters of drift rate and decision threshold instead of mean response time and accuracy. Although many previous studies – which we cite throughout our article – have taken this approach for their specific research question, our proposal is a broader one; that the entire mindset of cognitive psychology should shift from the current default focus on observed variables, to a default focus on theoretically meaningful, latent parameters. We believe that if EAMs were the standard focus for cognitive psychology studies that made inferences about the underlying cognitive process, then subsequent EAM studies would no longer be required to properly test the original claims, removing the current burden on decision-making researchers to heavily invest their time into "EAM account of task X" applications. Using the ageing example discussed earlier, researchers initially focused on the observed variable of mean response time, finding that older participants are slower than younger participants, and inferring that this reflected a cognitive slowdown. Essentially, this standard focus on mean response time in cognitive psychology created the need for the subsequent study of Ratcliff et al. (2001), to properly test the cause of the slower performance in older participants. However, if EAMs were the default method of analyses for cognitive psychology, then previous research would have found that older participants are generally more cautious than younger participants, the *cognitive slowdown* theory (Salthouse, 1996) – which is still commonly cited – would never have been proposed, and the study of Ratcliff et al. (2001) would not have been required. Importantly, if these latent parameters were the default for cognitive psychology, then researchers would need to justify their use of mean response time and accuracy instead of these parameters, especially in cases where researchers attempt to use these observed variables as a proxy to the cognitive process.

In order to make the latent parameters from EAM the default analysis for cognitive psychology analyses, further research is needed to better integrate EAMs with the statistical techniques currently used by cognitive psychology researchers (and potentially psychometric researchers; see Batchelder, 2010; Pe, Vandekerckhove, & Kuppens, 2013, for the concept of "cognitive psychometrics"). Currently, the methods of statistical inference within the EAM frame-



work is highly variable between studies, with different studies using different methods – which can lead to different conclusions (Evans, 2019b, 2019c) – to answer similar types of questions. Therefore, for EAMs to become the default analysis within cognitive psychology, some default, uniform method of applying EAMs must first be created within decision-making. One potential reason for this lack of uniformity across studies is that common statistical analysis methods are not always feasible in the context of more complex cognitive models. For example, Bayes factors (Kass & Raftery, 1995) have become one of the more common methods of statistical analysis for observed variables (Rouder, Morey, Speckman, & Province, 2012). However, when applied to EAMs and other complex cognitive models, Bayes factors become analytically intractable to calculate. However, recent research has proposed several methods for calculating Bayes factors for cognitive models, which reduce the computational burden required to estimate Bayes factors for these complex models (Evans & Brown, 2018; Gronau et al., 2017; Annis, Evans, Miller, & Palmeri, 2019; Evans & Annis, 2019), and an *R* package has been developed to implement one of these methods (*bridge sampling*; Gronau, Singmann, & Wagenmakers, 2019). Conceptually, the combination of Bayes factor approximation methods with EAMs could be thought of as a latent variable version of linear mixed-effect models, where a general linear model (e.g., t-test, ANOVA) operates *on top* of the EAM, with the latent parameters are estimated in a Bayesian manner and the uncertainty in these parameters (i.e., the posterior distributions) replacing the measurement error in the observed variable used in standard linear mixed-effects models. Furthermore, to be practically accessible for most cognitive psychology researchers, these integrated analysis methods for EAMs would need to be implemented in easy-to-use packages for common statistical programming languages.

Another important issue is establishing the consistency between different EAM variants in their conclusions as measurement tools. Although most EAMs provide close mimicry of one another when accounting for the choice response time distributions, this does not mean that these different variants will provide similar estimates of the latent parameters – or, as is most often the question of interest, come to similar conclusions about which latent parameters vary between experimental conditions and/or groups. Previous research has found that several different variants of the diffusion model show a close mapping with the parameters of the LCA (van Ravenzwaaij & Oberauer, 2009), and that the diffusion model and the linear ballistic accumulator (LBA) provide close agreement on esti-

ated parameters when the accumulation process in the LBA is constrained to more closely reflect the diffusion model (Donkin, Brown, Heathcote, & Wagenmakers, 2011), which suggests that the conclusions drawn from EAMs may be fairly independent of the specific EAM applied. However, the recent many-lab study of Dutilh et al. (2018) – where each research team had to decide which latent parameters varied across experimental conditions in 14 different experiments¹ – found reasonable discrepancies in conclusions between teams that used the diffusion model and teams that used the LBA, which were much larger than discrepancies between teams that used the same model. Furthermore, recent parameter recovery studies have provided a pessimistic perspective of the measurement properties of more complex EAMs, showing poor recovery for models with decision urgency that varies over the course of a decision (Evans, Trueblood, & Holmes, 2019), the LCA (Miletić, Turner, Forstmann, & van Maanen, 2017), and even the diffusion model when all random between-trial variability parameters are included (Lerche & Voss, 2016; Boehm et al., 2018). However, the inability to recover absolute parameter values does not necessarily mean that the relative differences between conditions in the values of specific parameters cannot be recovered, or that these more complex EAMs will differ from simpler EAMs in their conclusions about which latent parameters vary between experimental conditions and/or groups. In order to make EAMs a more robust default analysis for cognitive psychology, future research should focus on more clearly determining the consistency (or lack thereof) between different EAMs in their conclusions about cognitive processes in realistic situations, as well as assessing the measurement properties of more complex EAMs in the context of relative changes in parameter values.

Answering different types of questions

As discussed previously, most applications of EAMs appear to take the form of “does parameter *Y* differ over condition/group *Z*”, and assess whether there is some reliable population difference between the conditions or groups. In some cases these applications answer important theoretical questions, such as the ageing example discussed previously, where Ratcliff et al. (2001) found older adults to be more cautious than younger adults. However, focusing most applications on these types of questions seems restrictive, and likely to ignore a many potentially interesting questions that EAMs could be used to answer.

One different potential application involves asking questions about *how many* participants display specific effects. Recent research within cognitive psychology has be-

¹It should be noted that Dutilh et al. (2018) did not actually use different experiments, and instead collected data from a single, large experiment, which they split into 14 “pseudo-experiments”.



gun to explore whether certain classic cognitive effects apply to all individuals, and whether any individuals display no effect, or reverse effects. For example, Haaf and Rouder (2018) developed a series of Bayesian hierarchical models to answer these questions for the Stroop effect (Stroop, 1935) – the classic cognitive effect whether people are slower to name the colour ink that a word is in if the word is the name of a different colour – finding that all participants appear to display the Stroop effect. These questions are no longer focused on whether there is some reliable population difference between the conditions or groups, but instead whether the entire group appears to behave in a homogeneous manner, providing a better understanding of potential individual differences. These types of questions could also provide valuable insights for cognitive models, such as assessing whether all participants become worse at tasks as difficulty increases, or whether all participants become less cautious when instructed to respond with an emphasis on speed. In cases where not all participants display an effect, these questions could also be extended to assessing how many participants display the effect, such as whether most people show a decrease in drift rate (Rae, Heathcote, Donkin, Averell, & Brown, 2014) and/or non-decision time (Voss, Rothermund, & Voss, 2004) when instructed to respond with an emphasis on speed. These types of questions seem particularly relevant to EAMs, where a small number of participants with strong evidence in favour of an model/effect can provide overwhelming evidence in favour of it, meaning that the population difference between the conditions or groups is only reflective of a few participants. Future research using EAMs may benefit from a lesser focus on population difference between the conditions or groups, and a greater focus on whether effects appear to occur in all, or most, people.

Another potential application involves using EAMs as *diagnostic tools* in applied settings. Most applications of EAMs focus on better understanding cognition through novel theoretical and/or empirical insights, assessing whether there is some relative difference between experimental conditions or groups. However, applications of EAMs rarely focus on the absolute values of the parameters and how they may be potentially useful, such as for diagnostics in applied setting. For example, previous studies have found strong, robust links between drift rate and IQ (Ratcliff et al., 2010, 2011), suggesting that there is a strong link between drift rate on basic decision-making tasks and people's intelligence. Importantly, this finding is not only theoretically interesting and an external validation of the meaning of the drift rate parameter; it also shows that drift rate can serve as a complement, or substitute, to intelligence testing in clinical diagnostics. IQ tests can often be time consuming to implement, and can suffer from poten-

tial culture-based biases and practice effects. Based on previous findings linking drift rate to intelligence, the measurement of drift rate in basic cognitive tasks appears to present a promising complement to standard intelligence testing, especially in situations where many intelligence tests may not be appropriate, such as the measurement of cognitive decline over a period of time where constant re-testing is required. Another example is the measurement of cognitive control in applied fields, which is commonly measured through the interference effect in the Stroop task (Stroop, 1935) or the flanker task (Eriksen & Eriksen, 1974), or through specifically designed inhibition tasks such as the stop-signal task (Lappin & Eriksen, 1966). However, EAMs have been specifically developed to account for these conflict tasks (Hübner, Steinhauser, & Lehle, 2010; White, Ratcliff, & Starns, 2011; Ulrich, Schröter, Leuthold, & Birngruber, 2015; Matzke, Dolan, Logan, Brown, & Wagenmakers, 2013), and can be used to provide a latent measurement of cognitive control. Therefore, researchers in these applied fields would be better served using the measurement of cognitive control from one of these models as a default, instead of the raw response times, providing any instance where the absolute parameter values in EAMs can serve as a useful diagnostic tool.

Concluding Remarks

Although evidence accumulation models continue to provide a useful standard paradigm for decision-making research, we believe that there are several limitations regarding how EAMs are currently tested and applied. Specifically, we believe that a theoretical plateau has been reached in our understanding of human decision-making through EAMs, and that applications of EAMs to better understand cognition have become restrictive and of limited value. Within the current article we propose several potential future directions for EAMs, and believe that exploring some of these future directions will allow EAMs to again provide researchers with valuable insights into the decision-making process, and more generally, cognition.

References

- Anderson, J. R. (1993). *Rules of the mind*. Psychology Press.
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111, 1036–1060.
- Annis, J., Evans, N. J., Miller, B. J., & Palmeri, T. J. (2019). Thermodynamic integration and steppingstone sampling methods for estimating Bayes factors: A tutorial. *Journal of Mathematical Psychology*, 89, 67–86.



- Aschenbrenner, A. J., Balota, D. A., Gordon, B. A., Ratcliff, R., & Morris, J. C. (2016). A diffusion model analysis of episodic recognition in preclinical individuals with a family history for alzheimer's disease: The adult children study. *Neuropsychology*, *30*, 225–238.
- Audley, R. J., & Pike, A. R. (1965). Some alternative stochastic models of choice 1. *British Journal of Mathematical and Statistical Psychology*, *18*(2), 207–225.
- Batchelder, W. H. (2010). Cognitive psychometrics: Using multinomial processing tree models as measurement tools.
- Boehm, U., Annis, J., Frank, M. J., Hawkins, G. E., Heathcote, A., Kellen, D., ... Wagenmakers, E.-J. (2018). Estimating across-trial variability parameters of the diffusion decision model: Expert advice and recommendations. *Journal of Mathematical Psychology*, *87*, 46–75.
- Brinley, J. F. (1965). Cognitive sets, speed and accuracy of performance in the elderly. In A. T. Welford & J. E. Birren (Eds.), *Behavior, Aging and the Nervous System* (pp. 114–149). Springfield, IL: Thomas.
- Brown, S. D., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review*, *112*, 117–128.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, *57*, 153–178.
- Brown, S. D., Marley, A. A. J., Donkin, C., & Heathcote, A. (2008). An integrated model of choices and response times in absolute identification. *Psychological Review*, *115*, 396–425.
- Burle, B., Spieser, L., Servant, M., & Hasbroucq, T. (2014). Distributional reaction time properties in the Eriksen task: Marked differences or hidden similarities with the Simon task? *Psychonomic Bulletin & Review*, *21*, 1003–1010.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459.
- Coles, M. G., Gratton, G., Bashore, T. R., Eriksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(5), 529–553.
- Cox, G. E., & Shiffrin, R. M. (2017). A dynamic approach to recognition memory. *Psychological Review*, *124*(6), 795.
- Donkin, C., Brown, S., Heathcote, A. J., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: Different models for response time, same conclusions about psychological mechanisms? *Psychonomic Bulletin & Review*, *55*, 140–151.
- Donkin, C., & Nosofsky, R. M. (2012). A power-law model of psychological memory strength in short-and long-term recognition. *Psychological Science*, *23*(6), 625–634.
- Dutilh, G., Annis, J., Brown, S. D., Cassey, P., Evans, N. J., Grasman, R. P., ... Donkin, C. (2018). The quality of response time data inference: A blinded, collaborative assessment of the validity of cognitive models. *Psychonomic Bulletin & Review*.
- Eidels, A., Houpt, J. W., Altieri, N., Pei, L., & Townsend, J. T. (2011). Nice guys finish fast and bad guys finish last: Facilitatory vs. inhibitory interaction in parallel systems. *Journal of mathematical psychology*, *55*(2), 176–190.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search task. *Perception & Psychophysics*, *16*, 143–149.
- Evans, N. J. (2019a). A method, framework, and tutorial for efficiently simulating models of decision-making. *Behavior Research Methods*, 1–15.
- Evans, N. J. (2019b). Assessing the practical differences between model selection methods in inferences about choice response time tasks. *Psychonomic Bulletin & Review*, 1–29.
- Evans, N. J. (2019c). What factors are most important in finding the best model of a psychological process? Comment on Navarro (2018). *PsyArXiv*.
- Evans, N. J., & Annis, J. (2019). Thermodynamic integration via differential evolution: A method for estimating marginal likelihoods. *Behavior Research Methods*, 1–18.
- Evans, N. J., Bennett, A. J., & Brown, S. D. (2018). Optimal or not; depends on the task. *Psychonomic Bulletin & Review*, 1–8.
- Evans, N. J., & Brown, S. D. (2017). People adopt optimal policies in simple decision-making, after practice and guidance. *Psychonomic Bulletin & Review*, *24*, 597–606.
- Evans, N. J., & Brown, S. D. (2018). Bayes factors for the linear ballistic accumulator model of decision-making. *Behavior Research Methods*, *50*, 589–603.
- Evans, N. J., Brown, S. D., Mewhort, D. J., & Heathcote, A. (2018). Refining the law of practice. *Psychological Review*, *125*, 592–605.
- Evans, N. J., Dutilh, G., Wagenmakers, E.-J., & van der Maas, H. L. J. (in press). Double responding: A new constraint for models of speeded decision making. *Cognitive Psychology*, 1–50.
- Evans, N. J., Hawkins, G. E., Boehm, U., Wagenmakers, E.-J., & Brown, S. D. (2017). The computations that support simple decision-making: A comparison between



- the diffusion and urgency-gating models. *Scientific Reports*, 7, 16433.
- Evans, N. J., Holmes, W. R., & Trueblood, J. S. (2019). Response-time data provide critical constraints on dynamic models of multi-alternative, multi-attribute choice. *Psychonomic Bulletin & Review*, 1–33.
- Evans, N. J., Rae, B., Bushmakin, M., Rubin, M., & Brown, S. D. (2017). Need for closure is associated with urgency in perceptual decision-making. *Memory & Cognition*, 45, 1193–1205.
- 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, 1925–1944.
- Evans, N. J., Tillman, G., & Wagenmakers, E.-J. (in press). Systematic and random sources of variability in perceptual decision-making: Comment on Ratcliff, Voskuilen, and McKoon (2018). *Psychological Review*, 1–50.
- Evans, N. J., Trueblood, J. S., & Holmes, W. R. (2019). A parameter recovery assessment of time-variant models of decision-making. *Behavior research methods*, 1–14.
- Finn, A., & Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: The case of food safety. *Journal of Public Policy and Marketing*, 11, 12–25.
- Gomez, P., Perea, M., & Ratcliff, R. (2013). A diffusion model account of masked versus unmasked priming: Are they qualitatively different? *Journal of Experimental Psychology: Human Perception and Performance*, 39, 1731–1740.
- Gomez, P., Ratcliff, R., & Childers, R. (2015). Pointing, looking at, and pressing keys: A diffusion model account of response modality. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 1515–1523.
- Grasman, R. P. P., Wagenmakers, E.-J., & van der Maas, H. L. J. (2009). On the mean and variance of response times under the diffusion model with an application to parameter estimation. *Journal of Mathematical Psychology*, 53, 55–68.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., . . . Steingroever, H. (2017). A tutorial on bridge sampling. *Journal of Mathematical Psychology*, 81, 80–97.
- Gronau, Q. F., Singmann, H., & Wagenmakers, E.-J. (2019). Bridgesampling: An R package for estimating normalizing constants. *Journal of Statistical Software*.
- Haaf, J. M., & Rouder, J. N. (2018). Some do and some don't? Accounting for variability of individual difference structures. *Psychonomic Bulletin & Review*, 1–18.
- Hasbroucq, T., Burle, B., Akamatsu, M., Vidal, F., & Possamai, C.-A. (2001). An electromyographic investigation of the effect of stimulus–response mapping on choice reaction time. *Psychophysiology*, 38, 157–162.
- Hawkins, G. E., Marley, A. A. J., Heathcote, A., Flynn, T. N., Louviere, J. J., & Brown, S. D. (2014). The best of times and the worst of times are interchangeable. *Decision*, 1, 192–214.
- Hawkins, G. E., Mittner, M., Boekel, W., Heathcote, A., & Forstmann, B. U. (2015). Toward a model-based cognitive neuroscience of mind wandering. *Neuroscience*, 310, 290–305.
- Heathcote, A., Lin, Y.-S., Reynolds, A., Strickland, L., Gretton, M., & Matzke, D. (2018). Dynamic models of choice. *Behavior Research Methods*, 1–25.
- Hübner, R., Steinhauser, M., & Lehle, C. (2010). A dual-stage two-phase model of selective attention. *Psychological review*, 117(3), 759.
- Jones, M., Curran, T., Mozer, M. C., & Wilder, M. H. (2013). Sequential effects in response time reveal learning mechanisms and event representations. *Psychological review*, 120(3), 628.
- Jones, M., & Dzhafarov, E. N. (2014). Unfalsifiability and mutual translatability of major modelling schemes for choice reaction time. *Psychological Review*, 121, 1–32.
- Kass, R. ., & Raftery, A. E. (1995). Bayes factors. *Journal of American Statistical Association*, 90(430), 773–795.
- Kieras, D. E., & Meyer, D. E. (1994). *The EPIC architecture for modeling human information-processing and performance: A brief introduction*. MICHIGAN UNIV ANN ARBOR DIV OF RESEARCH DEVELOPMENT and ADMINISTRATION.
- Knowles, J., Evans, N. J., & Burke, D. (2019). Some evidence for an association between early life adversity and decision urgency. *Frontiers in Psychology*, 10, 243.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and comparison of value in simple choice. *Nature Neuroscience*, 13, 1292–1298.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- LaBerge, D. A. (1962). A recruitment theory of simple behavior. *Psychometrika*, 27, 375–396.
- Laird, J. E. (2008). Extending the soar cognitive architecture. *Frontiers in Artificial Intelligence and Applications*, 171, 224.
- Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London: Academic Press.



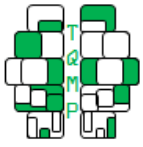
- Lappin, J., & Eriksen, C. W. (1966). Use of delayed signal to stop a visual reaction-time response. *Journal of Experimental Psychology*, 72, 805–811.
- Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception, & Psychophysics*, 72, 246–273.
- Leite, F. P., & Ratcliff, R. (2011). What cognitive processes drive response biases? A diffusion model analysis. *Judgment & Decision Making*, 6(7), 651–687.
- Lerche, V., Christmann, U., & Voss, A. (2019). Impact of context information on metaphor elaboration. *Experimental Psychology*, 65, 370–384.
- Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in psychology*, 7, 1324.
- Link, S. W. (1975). The relative judgement theory of two choice response time. *Journal of Mathematical Psychology*, 12, 114–135.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, 40, 77–105.
- Loftus, G. R., Busey, T. A., & Senders, J. W. (1993). Providing a sensory basis for models of visual information acquisition. *Perception & Psychophysics*, 54(4), 535–554.
- Luce, R. D. (1986). *Response times*. New York: Oxford University Press.
- Matzke, D., Dolan, C. V., Logan, G. D., Brown, S. D., & Wagenmakers, E.-J. (2013). Bayesian parametric estimation of stop-signal reaction time distributions. *Journal of Experimental Psychology: General*, 142, 1047.
- Matzke, D., Hughes, M., Badcock, J. C., Michie, P., & Heathcote, A. (2017). Failures of cognitive control or attention? The case of stop-signal deficits in schizophrenia. *Attention, Perception, & Psychophysics*, 79, 1078–1086.
- Matzke, D., Love, J., & Heathcote, A. (2017). A Bayesian approach for estimating the probability of trigger failures in the stop-signal paradigm. *Behavior Research Methods*, 49, 267–281.
- Maylor, E. A., & Rabbitt, P. (1987). Effects of alcohol and practice on choice reaction time. *Perception & Psychophysics*, 42, 465–475.
- Meyer, D. E., & Kieras, D. E. (1994). *EPIC computational models of psychological refractory-period effects in human multiple-task performance*. MICHIGAN UNIV ANN ARBOR DIV OF RESEARCH DEVELOPMENT and ADMINISTRATION.
- Meyer, D. E., Kieras, D. E., Lauber, E., Schumacher, E. H., Glass, J., Zurbriggen, E., ... Apfelblat, D. (1995). Adaptive executive control: Flexible multiple-task performance without pervasive immutable response-selection bottlenecks. *Acta Psychologica*, 90(1-3), 163–190.
- Miletić, S., Turner, B. M., Forstmann, B. U., & van Maanen, L. (2017). Parameter recovery for the leaky competing accumulator model. *Journal of Mathematical Psychology*, 76, 25–50.
- Milnes, B. G., Pelton, G., Doorenbos, R., Hucka, M., Laird, J., Rosenbloom, P., & Newell, A. (1992). A specification of the Soar cognitive architecture in Z.
- Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience*, 32, 2335–2343.
- Newell, A. (1990). *Unified theories of cognition*. Harvard University Press.
- Nosofsky, R. M., Little, D., Donkin, C., & Fific, M. (2011). Short-term memory scanning viewed as exemplar-based categorization. *Psychological Review*, 118, 280–315.
- Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review*, 104, 266–300.
- Osth, A., & Farrell, S. (2018). Using response time distributions and race models to characterize primacy and recency effects in free recall initiation. *PsyArXiv*.
- Osth, A., Jansson, A., Dennis, S., & Heathcote, A. (2018). Modeling the dynamics of recognition memory testing with an integrated model of retrieval and decision making. *Cognitive psychology*, 104, 106–142.
- Pe, M. L., Vandekerckhove, J., & Kuppens, P. (2013). A diffusion model account of the relationship between the emotional flanker task and rumination and depression. *Emotion*, 13(4), 739.
- Philiastides, M. G., & Ratcliff, R. (2013). Influence of branding on preference-based decision making. *Psychological Science*, 24, 1208–1215.
- Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: A theory of choice, decision time, and confidence. *Psychological Review*, 117, 864–901.
- Rabbitt, P. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, 71, 264–272.
- Rabbitt, P. (1967). Time to detect errors as a function of factors affecting choice-response time. *Acta Psychologica*, 27, 131–142.
- Rabbitt, P. (1968). Three kinds of error-signalling responses in a serial choice task. *Quarterly Journal of Experimental Psychology*, 20, 179–188.
- Rabbitt, P. (1969). Psychological refractory delay and response-stimulus interval duration in serial, choice-response tasks. *Acta Psychologica*, 30, 195–219.
- Rabbitt, P. (2002). Consciousness is slower than you think. *The Quarterly Journal of Experimental Psychology Section A*, 55, 1081–1092.



- Rabbitt, P., Cumming, G., & Vyas, S. (1978). Some errors of perceptual analysis in visual search can be detected and corrected. *The Quarterly Journal of Experimental Psychology*, 30, 319–332.
- Rabbitt, P., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Quarterly Journal of Experimental Psychology*, 29, 727–743.
- Rabbitt, P., & Vyas, S. (1981). Processing a display even after you make a response to it. How perceptual errors can be corrected. *The Quarterly Journal of Experimental Psychology Section A*, 33, 223–239.
- 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, 1226–1243.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347–356.
- Ratcliff, R., & Rouder, J. N. (2000). A diffusion model account of masking in two-choice letter identification. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 127–140.
- Ratcliff, R., & Smith, P. (2015). Modeling simple decisions and applications using a diffusion model. In J. R. Busemeyer, Z. Wang, J. T. Townsend, & A. Eidels (Eds.), *The Oxford Handbook of Computational and Mathematical Psychology* (pp. 35–62). Oxford University Press.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20, 260–281.
- Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, 116, 59–83.
- Ratcliff, R., & Starns, J. J. (2013). Modeling response times, choices, and confidence judgments in decision making. *Psychological Review*, 120, 697–719.
- Ratcliff, R., & Strayer, D. (2014). Modeling simple driving tasks with a one-boundary diffusion model. *Psychonomic Bulletin & Review*, 21, 577–589.
- Ratcliff, R., Thapar, A., & McKoon, G. (2001). The effects of aging on reaction time in a signal detection task. *Psychology and Aging*, 16, 323–341.
- Ratcliff, R., Thapar, A., & McKoon, G. (2010). Individual differences, aging, and IQ in two-choice tasks. *Cognitive Psychology*, 60(3), 127–157.
- Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, 140, 464–487.
- Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, 9, 438–481.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, 102, 261–300.
- Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, 461, 263–266.
- Ritter, F. E., Tehranchi, F., & Oury, J. D. (2019). ACT-R: A cognitive architecture for modeling cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(3), e1488.
- Roe, R. M., Busemeyer, J. R., & Townsend, J. T. (2001). Multi-alternative decision field theory: A dynamic artificial neural network model of decision-making. *Psychological Review*, 108, 370–392.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56(5), 356–374.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Schweickert, R. (1989). Separable effects of factors on activation functions in discrete and continuous models: d' and evoked potentials. *Psychological Bulletin*, 106(2), 318.
- Servant, M., White, C., Montagnini, A., & Burle, B. (2016). Linking theoretical decision-making mechanisms in the simon task with electrophysiological data: A model-based neuroscience study in humans. *Journal of Cognitive Neuroscience*, 28(10), 1501–1521.
- Servant, M., White, C., Montagnini, A., & Burle, B. (2015). Using covert response activation to test latent assumptions of formal decision-making models in humans. *Journal of Neuroscience*, 35(28), 10371–10385.
- Smith, P. L., & Van Zandt, T. (2000). Time-dependent poison counter models of response latency in simple judgment. *British Journal of Mathematical and Statistical Psychology*, 53, 293–315.
- Smith, P. L., & Vickers, D. (1988). The accumulator model of two-choice discrimination. *Journal of Mathematical Psychology*, 32, 135–168.
- Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin & Review*, 19, 139–145.
- Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, 25, 251–260.



- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Teodorescu, A. R., & Usher, M. (2013). Disentangling decision models: From independence to competition. *Psychological Review*, 120, 1–38.
- 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.
- Thompson, C. A., Ratcliff, R., & McKoon, G. (2016). Individual differences in the components of children's and adults' information processing for simple symbolic and non-symbolic numeric decisions. *Journal of Experimental Child Psychology*, 150, 48–71.
- Tillman, G., & Logan, G. D. (2017). The racing diffusion model of speeded decision making. *PsyArXiv*.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. CUP Archive.
- Townsend, J. T., & Fikes, T. (1995). A beginning quantitative taxonomy of cognitive activation systems and application to continuous flow processes (tech. rep. no. 131). indiana university bloomington. *Cognitive Science Program*.
- 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(6), 476–494.
- Townsend, J. T., & Wenger, M. J. (2004). A theory of interactive parallel processing: New capacity measures and predictions for a response time inequality series. *Psychological review*, 111(4), 1003.
- Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multi-attribute linear ballistic accumulator model of context effects in multi-alternative choice. *Psychological Review*, 121, 179–205.
- Tsetsos, K., Usher, M., & Chater, N. (2010). Preference reversal in multiattribute choice. *Psychological Review*, 117(4), 1275–1293.
- Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Sederberg, P. B., & Steyvers, M. (2013). A Bayesian framework for simultaneously modeling neural and behavioral data. *NeuroImage*, 72, 193–206.
- Turner, B. M., Rodriguez, C. A., Norcia, T. M., McClure, S. M., & Steyvers, M. (2016). Why more is better: Simultaneous modeling of eeg, fmri, and behavioral data. *NeuroImage*, 128, 96–115.
- Turner, B. M., Van Maanen, L., & Forstmann, B. U. (2015). Informing cognitive abstractions through neuroimaging: The neural drift diffusion model. *Psychological Review*, 122, 312–336.
- Turner, B. M., Wang, T., & Merkle, E. C. (2017). Factor analysis linking functions for simultaneously modeling neural and behavioral data. *NeuroImage*, 153, 28–48.
- Ulrich, R., Schröter, H., Leuthold, H., & Birngruber, T. (2015). Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. *Cognitive Psychology*, 78, 148–174.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, 108, 550–592.
- Usher, M., & McClelland, J. L. (2004). Loss aversion and inhibition in dynamical models of multialternative choice. *Psychological Review*, 111, 757–769.
- Vaghi, M. M., Luyckx, F., Sule, A., Fineberg, N. A., Robbins, T. W., & De Martino, B. (2017). Compulsivity reveals a novel dissociation between action and confidence. *Neuron*, 96(2), 348–354.
- Van Zandt, T., & Maldonado-Molina, M. M. (2004). Response reversals in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 1147–1166.
- van Ravenzwaaij, D., Boekel, W., Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2014). Action video games do not improve the speed of information processing in simple perceptual tasks. *Journal of Experimental Psychology: General*, 143, 1794–1805.
- van Ravenzwaaij, D., & Oberauer, K. (2009). How to use the diffusion model: Parameter recovery of three methods: Ez, fast-dm, and dmat. *Journal of Mathematical Psychology*, 53(6), 463–473.
- Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods*, 40, 61–72.
- Verdonck, S., & Tuerlinckx, F. (2014). The Ising decision maker: A binary stochastic network for choice response time. *Psychological Review*, 121(3), 422–462.
- Vickers, D. (1979). *Decision processes in visual perception*. London: Academic Press.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, 32, 1206–1220.
- Voss, A., & Voss, J. (2007). Fast-dm: A free program for efficient diffusion model analysis. *Behavior Research Methods*, 39, 767–775.
- Wagenmakers, E.-J., Ratcliff, R., Gomez, P., & McKoon, G. (2008). A diffusion model account of criterion shifts in the lexical decision task. *Journal of Memory and Language*, 58(1), 140–159.
- Wagenmakers, E.-J., van der Maas, H. J. L., Dolan, C., & Grasman, R. P. P. (2008). EZ does it! Extensions of the



- EZ-diffusion model. *Psychonomic Bulletin & Review*, 15, 1229–1235.
- Wagenmakers, E.-J., van der Maas, H. J. L., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, 14, 3–22.
- Wenger, M., & Townsend, J. T. (2006). "on the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli". *Journal of Experimental Psychology: Human Perception and Performance*, 23(3), 755–779.
- White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive Psychology*, 63, 210–238.
- Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the drift-diffusion model in Python. *Frontiers in Neuroinformatics*, 7.
- Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 53–79.
- Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, 111, 931–959.

Citation

Evans, N. J., & Wagenmakers, E.-J. (2020). Evidence accumulation models: Current limitations and future directions. *The Quantitative Methods for Psychology*, 16(2), 73–90. doi:[10.20982/tqmp.16.2.p073](https://doi.org/10.20982/tqmp.16.2.p073)

Copyright © 2020, *Evans and Wagenmakers*. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 17/05/2019 ~ Accepted: 19/08/2019