A Single Process Model of the Same-Different Task

Bradley Harding

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
M.A./Ph.D. in Experimental Psychology

School of Psychology
Faculty of Social Science
University of Ottawa

Abstract

The Same-Different task has a long and controversial history in cognitive psychology. For over five decades, researchers have had many difficulties modelling the simple task, in which participants must respond as quickly and as accurately as possible whether two stimuli are the "Same" or "Different". The main difficulty in doing so stems from the fact that "Same" decisions are much faster than can be modelled using a single process model without resorting to post-hoc processes, a finding since coined the *fast-same phenomenon*. In this thesis, I evaluate the strengths and shortcomings of past modelling endeavours, deconstruct the fast-same phenomenon while exploring the role of priming as its possible mechanism, investigate coactivity as a possible architecture underlying both decision modalities, and present an accumulator model whose assumptions and parameters stem from these results that predicts Same-Different performance (both response times and accuracies) using a single-process, a finding deemed near impossible by Sternberg (1998).

Keywords: Same-Different task, matching task, comparison task, priming, fast-same phenomenon, coactivity, cognitive architecture, modelling.

Acknowledgements

To my teammate and wife, Ranjita Padalia, I thank you for all of your unconditional love, support, and cheerleading throughout these years. You were there for every significant event and have shared my every experience. I always look forward to your thoughts and positive attitude, and with the amount of work you put alongside me, there should be a degree waiting for you too! I love you.

To my parents, late grand-parents, and family (new and old), you have provided me with endless moral and emotional support and I cannot express my gratitude enough. Your enthusiasm, love, and encouragement are truly appreciated.

To my friends, I thank you for everything. Matt, I love you as a brother and to Steve (and Nicole) and Kendrick, your couches have saved me thousands in late night taxis.

A very special thank you to NSERC, OGS, and the Pierre Baron Scholarship for providing me with the funding to conduct my research and to the University of Ottawa School of Psychology administrative staff for all their help throughout my time at the University of Ottawa.

With special mentions to Christophe Tremblay, Marc-André Goulet, and Nareg Berberian, I would like to thank all members of the Quibb research group (past and present) for their comments, feedback, encouragements, as well as being great friends and an amazing support system. I would also like to personally acknowledge the overwhelming number of volunteers that have come through our lab to help build our research.

Lastly, I would like to thank my supervisor, Denis Cousineau, for all the support, enthusiasm, motivation, and guidance that he has given me. Denis, without you taking a chance on me all those years ago and helping me get into the program, I wouldn't be where I am today. I am forever grateful. Thank you.

Table of Contents

Tables and Figures

Introduction

At any given moment, we are faced with an incredible number of stimuli, forcing a large number of decisions. Sorting out objects that remain unchanged from one moment to the next is an efficient way to minimize the number of operations. While it is known that comparison processes can be performed very efficiently, little is known regarding how people detect "sameness" between stimuli with split-second speed and near-perfect accuracy (see Farell, 1985, Sternberg, 1998, for extensive reviews). The objective of this thesis is to shed light on this mystery and offer a mechanism for this fundamental cognitive process.

The Same-Different task

The Same-Different task – sometimes called the comparison task or the matching task – is commonly used to explore the concepts of "sameness" and "difference". In this task, participants judge as accurately and rapidly as possible whether two presented stimuli are the "Same" or "Different".

Many variants of this task exist, including the comparison of letters (e. g., Nickerson, 1965; Bamber, 1969, 1972; Bamber & Paine, 1973; Krueger, 1973; Taylor, 1976a), numbers (e. g., Snodgrass, 1972; Silverman & Goldberg, 1975; Van Optstal & Vergut, 2011), words (e. g., Well, Pollatsek, & Schindler, 1975; Farell, 1977), faces (e. g. Tversky, 1969; Megreya & Burton, 2006), abstract patterns (e. g., Egeth, 1966; Nickerson, 1967a, 1967b; Bindra, Donderi, & Nishisato, 1968; Taylor, 1969; Link & Tindall, 1971; Snodgrass, 1972; Hock, 1973; Nickerson & Pew, 1973; Dyer, 1973), motion direction (Petrov, 2009), and tones (Bindra, Williams, & Wise, 1965; Bindra et al., 1968; Nickerson, 1969).

There also exist two variants regarding its decision rule. 1) In the conjunctive, or "all-Same", task (Bamber, 1969; Derks, 1972), participants answer "Same" when a criterion stimulus

(S_1) and a test stimulus (S_2) match on all attributes; “Different” responses are to be made when at least one attribute differs. 2) In the disjunctive, or “all-Different” task, participants answer “Same” as soon as a single match between S_1 and S_2 is found and answer “Different” if, and only if, all attributes mismatch (Nickerson, 1967a; Sekuler & Abrams, 1968; Silverman & Goldberg, 1975; Taylor, 1976a; Farell, 1977; reviewed thoroughly in Farell, 1985).

Herein, all presented Same-Different tasks will have a conjunctive decision rule and the compared stimuli will be successive strings of letters sampled from the Roman alphabet. Thus, each experiment replicates Bamber’s (1969) seminal experimental design that sparked decades of research and debate that followed from the discovery of the task’s most notable and robust result, the *fast-same phenomenon*.

The fast-same phenomenon

The *fast-same phenomenon* (expression first found in Bamber, 1972, p. 321 but with allusions found in Egeth, 1966; Bamber, 1969; Nickerson, 1967a, 1967b, 1968; it was reviewed in Farrell, 1985, St. James & Eriksen, 1991; and Sternberg, 1998) is the observation that “Same” response times (RT) are reliably much faster than the RT for most “Different” conditions and are always faster than the slowest “Different” condition. This effect is completely counter-intuitive from a modelling standpoint as “Same” responses should be based on an exhaustive examination of all attributes whereas “Different” responses can be self-terminating; this holds whether processing is serial or parallel (Taylor, 1976a; Townsend and Ashby, 1983; Townsend & Nozawa, 1995; Harding, Goulet, Jolin, Villeneuve, Tremblay, & Durand, 2016).

This effect is further characterized by (1) “Same” RTs being as fast or faster than all “Different” RTs and (2) very high “Same” accuracies (often 95% and over) surpassing accuracy of most “Different” conditions (this component of the fast-“Same” effect is sometimes referred to the *False-"Different" effect*; Beller, 1970; Krueger 1978). In addition to its faster-than-expected speed, some researchers have also noted a shallower slope for "Same" RTs as a function of the total number of attributes composing the stimulus (Bamber, 1969, 1972; Taylor, 1976a; Sternberg, 1998).

The fast-same phenomenon is robust to variations in experimental design and has become a staple finding of the task. Yet, to this day, there is no agreed-upon explanatory model of the mechanism(s) behind the phenomenon nor has there been any model that can also predict both "Same" and "Different" RT and accuracy with a single process.

Typically observed results for a Same-Different task

Herein, I will use the "*d*D*l*L" notation, where *d* represents the number of differences (D) within a string's length (L) of *l* characters. For example, the condition where a single difference in a string of three 3 letters is presented will be referred to as the 1D3L condition. A "Same" condition has zero differences and will therefore be referred to as 0D*l*L.

To be considered a faithful replication of Bamber's (1969) seminal work, one must be able to match the following properties: First, "Same" responses are faster or as fast as all "Different" judgements and follow an upwards trend as the stimuli's length increases. Second, accuracy of "Same" responses are typically very high (95% or better) and are generally unaffected by letter string length. Third, "Different" RTs decrease as the number of differences within the letter string increases, with the all-"Different" conditions being the fastest for strings of all lengths (Bamber, 1969; Taylor, 1976a). For example, a 4D4L stimulus will yield faster responses than a 4L stimulus with one, two or three differences within the string. Fourth, the matching to mismatching ratio of letters affects the mean accuracy rate of all "Different" conditions; the more matching letters there are in a "Different" test stimulus, the less accurate overall decisions tend to be. The condition where there is just a single difference in a string of four letters is the least accurate condition (Bamber, 1969; Silverman & Goldberg, 1975; Sternberg, 1998). Fifth and finally, as noted by Sternberg (1998), the RT slope as a function of stimulus length must "fan-out" as the number of differences increases (1D has the steepest slope and is steeper than the 2D condition and so on). As for "Same" responses, they have the shallowest slope of all. Expected trends for a Same-Different are exemplified in Figure I.1. These data stem from Same-Different task control conditions that will be introduced in Chapter 2.

INSERT FIGURE I.1 ABOUT HERE

Thesis architecture

This thesis is divided into four parts. In Chapter 1, I introduce various Same-Different task models that have been proposed, as well as their respective strengths and shortcomings. This literature review will serve the purpose of noting novelties and gaps that must be filled to highlight the importance of the model that I propose in a later chapter. In Chapter 2, I delve into the possible workings of the fast-same phenomenon and offer a simple intuitive explanation of its possible mechanism. In Chapter 3, I present a series of analyses results from two variants of the Same-Different task that provide evidence for a possible underlying mechanism that unifies an understanding of both "Same" and "Different" decisions. Finally, in Chapter 4, I introduce a model of Same-Different task performance that leverages the findings and insights gathered from the preceding chapters. This single process model that I will present predicts both RT and accuracy for both "Same" and "Different" decisions for all three types of Same-Different task designs. While this model is a predictive model of decisions with free parameters, the overall model's predictions are not tied to specific parameter values and return identical trends regardless of their values – the selected parameters are for scaling purposes only. I intend to publish Chapters 2 to 4 as separate manuscripts. Details for my publication plans are appended to each chapter.

Chapter 1: Modelling the Same-Different task

While it is apparent that the Same-Different task is simple in nature, there is controversy surrounding its result: there is no model that can predict the speed and accuracy of both “Same” and “Different” judgements using a single process, while also accounting for the fast-same phenomenon. In this chapter, I introduce past models of the Same-Different task and catalogue their respective strengths and weaknesses for predicting peoples’ behaviour over all 5 benchmarks of the task. This literature review serves to give an overview of past endeavours as well as show the novelty of the model I present in Chapter 4.

Holistic matching

The *holistic matching*, or *template matching* model (Egeth, 1966) is the first and simplest model to be proposed as an account of peoples’ behaviour in the Same-Different task. In this model, if S_1 matches the shape of S_2 (there is no analytical treatment of stimulus sub-parts), a fast “Same” response is triggered; slower “Different” responses are triggered as an alternative when the templates do not match.

Strengths

This model explains why “Same” responses are faster than “Different” responses. As “Different” decisions can only occur after a template mismatch, they are necessarily predicted to be slower.

Weaknesses

This model does not explain why there is a discrepancy in speed for all conditions within “Same” responses. If the template matches, all conditions should have the same overall speed (i. e. why is the 0D4L condition slower than the 0D1L condition if all “Same” RTs are simply the result of a square-peg/square-hole situation?). This issue also extends to “Different” decisions; template models predict no RT differences for the various “Different” conditions (i. e., it has been repeatedly found that the all-“Different” condition is systematically the fastest condition amongst “Different” decisions while increasing the number of matches within the string decreases the overall decision’s speed). Finally, there are no propositions as to the information gathering mechanism, nor how accuracy is predicted.

Analytical models

To overcome the template matching model’s evident shortcomings, more analytical models were proposed, the most popular of which is the serial self-terminating model (Egeth, 1966; Bamber, 1969). In this model, it is assumed that S_2 is broken down into its individual components (letters in Bamber’s, 1969, view) and each is processed sequentially (Townsend & Ashby, 1983). A “Different” response is triggered as soon as a mismatch between S_1 and S_2 is found. If no mismatch is found between stimuli, a “Same” response is triggered. While most analytical models in Same-Different research have focused on serial models, parallel models have also been proposed and reviewed to show their capabilities in predicting “Different” decisions (Hawkins, 1969; Hawkins & Shigley, 1972; Taylor, 1976a).

Strengths

This approach has since been established as the “gold standard” because it can predict “Different” responses very well and in a very parsimonious way (Bamber, 1969; Silverman & Goldman, 1975; Taylor, 1976a; Sternberg, 1998). If one were to abandon parsimony, Taylor (1976a) has shown that a limited capacity parallel self-terminating architecture, with exponentially distributed processing times, can offer a better fit for “Different” decisions.

Weaknesses

Issues arise however with this model's predictions of "Same" judgements. Serial self-termination forcibly assumes that identical strings require that S_2 be treated in its entirety before it is possible to elicit a "Same" decision, making the final decision necessarily exhaustive (the self-termination would occur once the end of the string has been processed rather than somewhere in the middle, as would be the case with "Different" responses). This of course predicts that "Same" decisions are always slower than "Different" decisions for stimuli of the same length – opposing what has been empirically observed and replicated time again. For parallel, unlimited capacity models, all individual letters of S_2 are treated at once resulting in no RT discrepancies between all "Same" conditions (Egeth, 1966; Taylor, 1976a; Sternberg, 1998). Moreover, RTs of the various "Different" conditions should not differ either as the length of the stimuli has no relevance regarding processing speed (Snodgrass & Townsend, 1980; Townsend & Ashby, 1983; Sternberg, 1998). Once more, this model offers no explanation to the accuracy rates of either decision modality.

The Identity Reporter

The issues with single-process models (holistic and serial self-termination) noted above inspired Bamber (1969) to lean towards a dual-process mechanism. In his approach, Bamber (1969) proposed that a serial-self terminating module is indeed at play to make both "Same" and "Different" judgements. However, a second decision module, dubbed the *Identity Reporter*, a process solely specialized at making "Same" responses, is also present. According to Bamber (1969), this Identity Reporter is a cognitive mechanism that only detects matching visual information much like the template matching model. When the Identity Reporter detects a physical match, its fast decision-making module is activated and outruns the serial self-terminating module's eventual "Same" response to a decision. This dual-process explanation has had many supporters (Tversky, 1969; Derks, 1972; Krueger, 1973; Nickerson & Pew, 1973; Decker, 1974; Bamber, Herder & Tidd, 1975; Silverman & Goldberg, 1975; Taylor, 1976b) and has inspired alternative attention models (Farell, 1984).

Strengths

This approach can explain why "Same" judgements are faster than "Different" judgements and why they do not follow the exhaustive response prediction made by serial processes. Furthermore, Bamber (1969) argued that the dual-process mechanism can attest for the "error awareness" that some participants have reported (where participants realized they made an error on certain trials after they had already recorded their decision). This awareness, according to Bamber, was key evidence that a serial process, whilst slower, is present and much more accurate than the faster Identity Reporter.

Weaknesses

Unfortunately, the fact that the Identity Reporter is directly tied to the holistic treatment of templates is its biggest weakness. For Same-Different tasks in which the physicality between S_1 and S_2 is altered while keeping the nominal identity constant (e. g. "J" and "j" would be considered "Same"), the fast-same disappears, yet RT remains faster than the slowest "Different" condition (1D; Posner & Mitchell, 1967; Beller, 1970; Bamber, 1972; Well & Green, 1972; Bamber & Paine, 1973; Pachella & Miller, 1976; Proctor, 1981; Eviatar, Zaidel, & Wickens, 1994; Ben-David & Algom, 2009). It would have been predicted that since the holistic process cannot act accordingly, the serial process would take over. At this point, predictions would be identical to those presented in the Analytical Models section above and "Same" responses would be slower than all “Different” decisions, including the 1D condition. The fact that this is not the case is direct evidence against this dual-process approach.

The Noisy Operator

Another model of the fast-same phenomenon is the idea that faster “Same” responses might be due to “Different” trials requiring a thorough treatment of how stimuli differ from one another (alluded to in Nickerson, 1965, but later formalized in Krueger, 1978). In his research, Krueger (1978, 1979) proposed the *Noisy Operator* model, a mechanism where participants make their decision after they have sequentially checked and rechecked all features of S_2. This model's core idea stems from the fact that matching attributes could be perceived as mismatching if perception is noisy; however, the converse is far less likely. This notion leads to the consequence that more confirmations are required to answer “Different” than “Same” (Egeth, 1966). According to the Noisy Operator, each alternative decision has a specific threshold and the associated response triggers as soon as one of the thresholds has been breached. Much like serial self-termination, this model also posits that the participants sequentially scan, or check, each feature of the stimuli to identify where the difference is located. However, unlike the serial model, each feature can be rechecked any number of times, predicting many RT results. As the name implies, the model also assumes that stimuli possess a certain amount of noise, which leads to imperfect stimuli processing. As said above, this internal noise will more likely lead “Same” stimuli to appear “Different” than “Different” stimuli appear to be the “Same”. Thus, according to Krueger (1978), therein lies the speed difference between “Same” and “Different” decisions: “Same” decisions are not quicker, it is rather that “Different” decisions are slower because the participant must recheck mismatches and identify the location of the differing dimension (also proposed as a possibility in Eriksen, O’Hara, & Eriksen’s, 1982, response-competition model).

While the simulations performed by Krueger are novel, analytical search and rechecking had been proposed as possibilities earlier (Howell & Stockdale, 1975; Taylor, 1976b).

Strengths

Krueger's (1978) Noisy Operator is the first single process model to offer an explanation to the fast-same phenomenon all while predicting how "Different" decisions are made. Furthermore, the Noisy Operator is the first model to predict accuracy rates of Same-Different data and the first to explain false-"Different" errors.

Weaknesses

Unfortunately, while making accurate predictions of RT and accuracy, the Noisy Operator has been criticized for generating unrealistic parameter estimates (Ratcliff, 1981). The model assumes that re-checking is at the core of the fast-same phenomenon but cannot explain why "Same" responses are still quicker than situations in which S_1 and S_2 are very different from one another. For example, consider that "Q" vs. "W" would be an easy trial and "E" vs. "F" would be a hard trial, the Noisy Operator predicts that the easy trial should take significantly less time to answer as re-checking is not necessary. However, we still observe a discrepancy between "Same" and "Different" RT in empirical data, regardless of how easy the "Different" trials can be. Furthermore, the model cannot account for Same-Different tasks in which the inter-stimulus interval is shorter than the time it requires to encode and re-check all pixels of a letter (200 ms/pass). Finally, as pointed out by Townsend and Ashby's (1983) comments on Same-Different models: while the analytical rechecking model is intuitive and makes many accurate predictions for both RT and accuracy, the successes are overshadowed by the glaring complexity of the model itself – for every additional dimension, the model needs to check and re-check all aspects of the stimuli which should in turn significantly increase the processing time. Townsend and

Ashby (1983) further noted that the Noisy Operator's increasing number of parameters, assumptions, and moving parts compared to simpler, more traditional models, make its testability and falsification difficult, and therefore, make its results hold less weight.

Priming

Another alternative to model the task's results is based on the *Name-Physical Disparity effect* (Proctor, 1981; Proctor & Rao, 1982, 1983) or as Krueger & Shapiro (1981) reframed it, the *priming effect*. While seemingly novel, priming had been first brought forth as a possibility by Nickerson, (1978) where the results of comparison tasks (Donderi & Zelnicker, 1969; Posner & Boies, 1971) were contrasted to those stemming from stimuli repetition tasks (Bertelson, 1961; Kornblum, 1969). In his work, Proctor (1981) composed a series of experimental conditions where in one case the participant had to answer "Same" or "Different" to physically matching stimuli and in another, to physically mismatching stimuli (much like Bamber's, 1972, research in which the stimuli cases were altered). He posited, and successfully found, that both cases presented faster responses for repeated stimuli, more so for the physically matching trials. Thus, it is posited that priming benefits physically matching stimuli because a physically identical stimulus has already been encoded just moments prior (possibly resulting in residual activation, Huber, 2008). His results also seemingly shed light on the faster-than-expected decision speed of physically mismatching "Same" stimuli (also found in Bamber 1972; Bamber & Paine, 1973; Posner & Mitchell, 1967); regardless of the letter's presentation identity, there still exists a semantic link and a phonological link (both "j" and "J" are considered the same letter), resulting in faster recognition of the target. This finding led him to posit that there is an encoding bias for "Same" responses regardless of physicality and that "Different" response necessitate a "from-scratch" encoding on every trial no matter the experimental manipulation.

Strengths

As an explanation for the fast-same phenomenon, priming is both a parsimonious an elegant approach. The model presented in Chapter 4 implements this aspect and Chapter 2 further explores the relation between priming and the fast-same phenomenon.

Weaknesses

While the priming model could explain why “Same” decisions are so much faster than “Different” decisions in the observed experimental conditions, it provided no insight on the other expected effects of a Same-Different task, such as error rates and why there is a shallower slope for “Same” responses (Taylor, 1976a) – in a simple priming model, the slope of “Same” responses should be simply shifted downwards. Furthermore, there are no insights on the decision mechanism that returns the expected RT pattern for both “Same” and “Different” results. Proctor’s (1981) work also received much criticism from other researchers such as Ratcliff and Hacker (1981; see Response bias section below) as well as Kruger and Shapiro (1981), who claimed that priming alone could not account for the task’s results and undermined the propositions that Proctor (1981) put forth.

Response bias

While stimuli-based priming could be at the core of fast-same responses, there is also the possibility of being inherently biased towards “Same” responses (Taylor, 1977; Ratcliff, 1978; Ratcliff & Hacker, 1981; Ratcliff, McKoon, & Verwoerd, 1989; Irwin, Hautus, & Francis, 2001). This modelling approach led largely by the work of Ratcliff and Hacker (1981), using Ratcliff’s Diffusion Model (RDM; Ratliff, 1978), explored the integration of the speed-accuracy tradeoff phenomenon (first introduced by Henmon, 1911; see Heitz, 2014, for a review) within the Same-

Different task by varying the levels of "cautiousness" the participants must exercise before answering "Same".

Strengths

Ratcliff and Hacker (1981) hypothesized and successfully found that with tailored instructions that elicit caution towards either response, one could create a bias against that decision (Ratcliff & Hacker, 1981). For example, instructions could entice very careful processing of "Same" responses, (ensuring that "Different" responses are as fast as possible regardless of performance) that return "Same" responses that are slower than "Different" responses. Furthermore, Ratcliff (1985) was able to show that the task (including data from his critics, Proctor & Rao, 1983) can be modelled with the RDM (notably using threshold manipulations, expanding on the works of Kruger, 1978, 1979; see also Howell & Stockdale, 1975, and Taylor, 1976b).

Weaknesses

While the response bias approach supported the hypothesis of Ratcliff and Hacker (1981), there are some fundamental issues with the reported interpretations, the most glaring of which is the misunderstanding of expected results (Farrell, 1985): Ratcliff and Hacker (1981) interpreted the fast-"Same" effect as "Same" decisions are always faster than all "Different" decisions, when it should be understood as "Same" decisions are always faster than decisions from the *slowest* "Different" condition (1D; Farrell, 1985; Sternberg, 1998). Unsurprisingly, taking this into account, Ratcliff and Hacker's (1981) results have been found before (see, Bamber, 1972; Bamber & Paine, 1973; Posner & Mitchell, 1967) and are not as novel as Ratcliff and Hacker (1981) suggested. Notably, in the 4L conditions (the only stimulus length chosen in Ratcliff & Hacker's, 1981, experiment), all-"Same" and all-"Different" RTs are often close and their confidence intervals frequently overlap. Furthermore, there have been critiques of Same-Different response patterns being solely caused by response biases in the past (Taylor, 1976b). Ratcliff and Hacker's (1981) results do not go to the core of why there is an upward slope for "Same" decisions as string length increases and fail to explain why the speed of "Same" decisions are affected when the stimuli's physical identities are modified. Also, it took extreme instructions to flip the RTs between "Same" and "Different" judgements. Finally, as noted by Proctor and Rao (1982): why are "Same" responses so quick when no experimental manipulation is enforced on participants?[1]

Response competition

During this fertile exchange between Ratcliff and Hacker (1981) and Proctor (1981), another stochastic modelling approach was introduced, the *response competition* model (Eriksen et al., 1982). While similar to Ratcliff's diffusion model approach (Ratcliff, 1978; Ratcliff & Hacker, 1981), this model not only accumulates evidence towards a particular decision but also backpropagates to accelerate the detection of other relevant information present within the stimulus (Eriksen et al., 1982; Krueger, 1987; St. James & Eriksen, 1992; Pan & Eriksen, 1993). Therefore, for "Same" responses, matching information sends a signal back through the accumulator and increases the overall speed of the process. However, when it comes to "Different" responses, its detection process resembles that of the Noisy Operator (Krueger, 1978); as "Different" responses often contain both matching and mismatching information. In

---

[1] It should be noted that Ratcliff and Hacker (1983) replied to this comment in a short note which sparked a fertile exchange (in chronological order: Proctor, 1981, Ratcliff & Hacker, 1981, Proctor & Rao, 1982, Ratcliff & Hacker, 1983, Proctor & Rao, 1983, Proctor, Rao & Hurst, 1984, Ratcliff, 1985, Proctor, 1986). They disapproved of Proctor and Rao's (1982) criticisms that the difference between both "cautious-Same" and "cautious-Different" conditions' RTs roughly equal to the RT discrepancy typically found between "Same" and "Different" responses. They also argued that the absolute differences between conditions should not be taken seriously as there is no way to discern if bias manipulation affected "Same" and "Different" responses in the same way.

such trials, the backpropagation causes a response competition that hinders the ability to accumulate evidence towards a “Same” versus “Different” decision and slows down the “Different” response.

Strengths

The response competition model introduced the concept of interactivity between channels and was able to explain the fast-same phenomenon with backpropagation.

Weaknesses

Much like the Noisy Operator model (Krueger, 1978), response competition is a very post-hoc model. Furthermore, if the number of mismatching dimensions is equal or greater to the number of matching dimensions, there should be little or no response competition. This would in turn benefit “Different” decisions only. The only other alternative to alleviate this noted issue would be to assume that the “Same” channel’s backpropagation holds more weight than that of the “Different” channel, an asymmetry which is unlikely (Farrell, 1985) considering that both decisions are equally possible a priori.

Literature reviews: No new information

Shortly after the debate between Ratcliff and Proctor (see Footnote 1), Farrell (1985) reviewed modelling approaches published at that point. In his review, he breaks down each individual approach and offers criticisms for each, similar to what I have done here. Sadly, no advances were proposed and modelling the Same-Different task remained at a stalemate.

Research on the Same-Different task halted for over a decade before another review was carried out in 1998, this time by esteemed researcher Saul Sternberg, (Sternberg, 1998). His chapter presents, reviews, and criticizes several novel and established models, but to no avail. According to Sternberg, despite a growing body of evidence against dual-process models and, “as unappealing as it is to introduce such complexity, we are forced to conclude that the two responses are generated by different processes” (Sternberg, 1998, p. 435).

Chapter 2: Controlling the fast-same phenomenon

Physical identity and priming

Although there are different levels of priming, the most relevant form for Same-Different decisions is identity priming where it is assumed that residual processing activation benefits a presentation of an identical stimulus within a brief time interval (Huber, 2008; Huber & O'Reilly, 2003; Jacob, Breitmeyer, & Trevino, 2013). It also posits the presence of a hierarchy through which any visually-presented stimulus must travel. The bottom levels are visual, the middle levels process phonological information, and the top levels treat the semantics of the stimuli. See Huber (2008), Huber and O'Reilly (2003), Eviatar, Zaidel, and Wickens (1994), and Lupker, Nakayama, and Perea (2015), for work pertaining to priming that targets a specific level in the processing hierarchy. When faced with identical stimulus, the network quickly reactivates, and the stimulus is "fast-tracked" through the hierarchy. As discussed in Chapter 1, if one were to remove physical priming benefits (by altering the physical aspect of S_2), phonological and semantic priming benefits (both "j" and "J" are the same letter and are pronounced identically) would remain, resulting in higher level forms of priming and consequently, faster recognition of the target. In other words, stimulus processing would not benefit from residual activation from the lower, perceptual, levels even though there may be priming influences at the upper phonological or semantic levels.

Cancelling low-level priming

As previously discussed, the priming model is a parsimonious and elegant mechanism to explain the fast-same phenomenon. However, to validate this hypothesis, it is necessary to create experiments in which the fast-"Same" responses are cancelled, or at least attenuated by manipulating the strength of this identity priming. As noted above, one way to do so is by altering the physical appearance of the compared stimuli so that a different processing pathway is taken to the upper processing levels of a decision. Such manipulations will be found in Experiments 1 and 2 described in this chapter. Alternatively, it should be possible to take different processing pathways by changing stimulus modality. For example, one could present the criterion audibly so that participants can still create a mental construction of S_1 without benefiting from a primed physically identical stimulus. This experimental manipulation is found in Experiment 3. Finally, one can avoid identity priming altogether by not presenting a criterion stimulus at all. Instead, cues can be presented that retrieve S_1 from long-term memory (LTM). This ensures that any activation resulting from the cues in the perceptual levels of processing are completely unrelated to the test stimulus. From LTM retrieval (and within the small interval of time given to the participants), only activation of the semantic level is probable considering that the mental representation of S_1 is not constructed from bottom-up pathways. This variant is found in Experiment 4. In this experiment, I asked the participant to memorize four stimuli (one for each of the experiment’s four possible stimuli lengths) and make all their subsequent Same-Different judgements based on cues showing the stimuli’s length only.

In this chapter, I explore these priming cancellation techniques and their effect on fast-“Same” results. If the fast-same phenomenon is indeed caused by priming, a cancellation or an attenuation of speed for “Same” responses in all these experimental manipulations should be observed whereas “Different” decision times should remain unaffected.

In all experiments, I analyze accuracy and RTs. I also examine slopes for “Same” and 1D conditions as a function of stimulus length; I chose these two conditions to address Sternberg’s (1998) claim that these conditions should be located at both extremities of the “fan-out” effect typically observed in a Same-Different task. While at opposite ends, they are theoretically the most similar. It is expected that the usual RT and accuracy trend (summarized in Chapter 1) will be found in all conditions except those where “Same” RT are intended to be altered. Furthermore, I analyze standard deviations and skewness in all conditions to observe whether the fast-“Same” and attenuated-”Same” results operate with qualitatively different underlying mechanisms; priming should influence speed of processing but otherwise show no qualitative differences between conditions. If the fast-same phenomenon can be attenuated or abolished, without these latter analyses yielding qualitative differences across experimental manipulations, one could conclude that the mechanism underlying the decision is unaffected by the experimental manipulation. As the scope of this chapter is centered on explaining the possible mechanism behind the fast-same phenomenon, I will focus mostly on “Same” decisions throughout these analyses.

Experiment 1: Case manipulation

In this first experiment, the string’s letter cases were varied to see whether changes in stimulus appearance between S_1 and S_2 affected the speed of “Same” responses in the context of a standard Same-Different task. This study is similar to Bamber’s (1972) study where uppercase and lowercase letters were intermixed randomly within a stimulus. He found that physical mismatches led to a reduction of the fast-“Same” effect, yet RTs remained faster than the slowest “Different” condition. In my experiment, it is expected that a fast-“Same” will occur when the stimuli match by letter identity regardless of case, and that it will be attenuated (like Bamber, 1972) when the stimuli otherwise mismatch. It is also expected that “Different” RTs will be unaffected by letter case as priming should only benefit “Same” responses.

Methodology

Participants

Participants were undergraduate and graduate students recruited at the University of Ottawa. All participants were between 18 and 30 years of age, had normal or corrected vision, and were informed of the experiment's procedure as well as the protocol and ethical rules of the University of Ottawa. They gave written and verbal consent to participate in this task. Finally, all participants were compensated $10 for their time (approximately 1 hour for briefing, testing, and debriefing). In this and all subsequent experiments, I aimed to recruit 20 participants per condition. This is five times more than in Bamber's (1969) article, and more than in most of the articles reviewed in Chapter 1; thus, statistical power should be satisfactory.

Stimuli

Stimuli were displayed on a calibrated CRT display having a resolution of 1024 × 768 pixels and a screen refresh rate of 85 Hz. The screens' displays were also calibrated to ensure a luminance and RGB standard across participants. Participants were seated approximately 50 cm from the front of the screen with the computer keyboard placed on the desk in front of them. Participants could adjust the latter to ensure comfortable testing conditions. Twelve consonants (B, C, D, F, J, K, L, N, S, T, V, and Z) were selected to serve as stimuli, matching as best as possible Bamber's (1969) original study. The stimuli were presented within a 10° × 10° visual angle centered on the screen with the first string (S_1) shown 4° above the center of the computer screen and the second (S_2) shown 4° below the center of the screen. Stimuli were always presented as white letters on a black background.

The letters were randomly selected on every trial. String length also varied randomly from 1 to 4 letters on every trial. No letter was presented twice within the same stimulus and matching letters would appear in the same position for both S_1 and S_2. For “Different” conditions, the differing letter(s) were different from those already used in both S_1 and S_2; S_2 could have no differences (“Same”) or a number of differences between 1 and L.

The main experimental manipulation, the priming manipulation, is that on half of the trials, S_1 could be shown using uppercase letters only; on the other half, S_1 was composed only of lowercase letters. Same occurred orthogonally for S_2. Thus, half of the trials presented physically matching stimuli while the other half showed physically mismatching stimuli. Unlike Bamber (1972), the entirety of the string’s composition was uppercase or lowercase letters; Bamber's original experiment could have stimuli resembling “JcvD”, or “jCvD” (a pilot study using this manipulation proved to be difficult for participants to complete and mean accuracy rates plummeted below what are typically observed). In the control condition, there was no change in case (384 trials; half using lowercase letters for both S_1 and S_2, half using uppercase letters for both stimuli). In the other half of the trials, the cases always mismatched between S_1 and S_2 so that on 192 trials a lowercase S_1 was presented with an uppercase S_2, and on 192 trials, an uppercase S_1 was presented along with a lowercase S_2. Participants were specifically instructed to pay no attention to the case of the stimuli and to respond solely on the nominal identity of the letter.

Procedure

During the on-screen instructions, the participants were instructed to respond by pressing the "*CTRL*" key located on the far left of the keyboard using their left hand, and the "*ENTER*" key located on the far right of the keyboard (on the numeric pad) using their right hand. The “Same” or “Different” decision associated with each button was counterbalanced based on the participant number. Therefore, half pressed “Same” with their left hand and half with their right.

The experiment began once the participant was ready and verbally acknowledged that he or she understood the procedure.

The timeline of a typical trial is shown in Figure 2.1. As shown, a fixation cross was presented for 500 ms, followed immediately by S_1, which was presented for 400 ms. Afterwards, a blank screen was presented for 400 ms followed by S_2, the test stimulus. This test stimulus was shown for 5000 ms or until a decision was made. Feedback was given for 500 ms on non-responses and on errors only to avoid diverting the gaze of the participant when they correctly answered. For correct answers, the screen was simply blank for 500 ms. Afterwards, there was a 500 ms blank screen before the subsequent trial began. While the task is easy, and participants rarely made mistakes, a message in red was shown if the participant made five mistakes in a row. Additionally, participants were offered short breaks to stretch their legs and rest their eyes after every 192 trials (one quarter of the experiment's total number of trials).

INSERT FIGURE 2.1 ABOUT HERE

Following testing, all participants were given a debriefing to answer any queries and to explain the goal of the study.

Experimental design

One session consisted of 768 trials. Both priming conditions consisted of 384 trials each, of which half were "Same" strings and half were "Different" strings. Additionally, strings of all lengths had an equiprobable chance of presentation, meaning that there was an equal number of 1L, 2L, 3L, and 4L stimuli. Additionally, within a given string length, differences had the same probability of occurrence. For example, when 4 letters are shown, 1D, 2D, 3D, and 4D each occurred an equal number of times, with serial position of the differences assigned at random. All trials were presented in a random order.

Table 2.1 summarizes these conditions (string length × number of differences) with the number of trials in each for a total of 384 trials in each of the two priming conditions.

INSERT TABLE 2.1 ABOUT HERE

Results

Screening of the data

Data from 14,592 total trials was gathered (768 trials × 19 participants). Twenty total participants were initially recruited but one was excluded prior to analysis for having very slow RT (mean RT of 1220 ms); it is suspected that the participant did not understand the instruction to respond as quickly as possible because most of the RTs were well above 1000 ms, an abnormality in the task. For the remaining participants, there were 27 RT below 200 ms and 17 RTs above 2500 ms that were excluded; all remaining trials had responses recorded within the 2500 ms allowed. For analyses of the response times, erroneous trials were filtered out to arrive at a total of 13,870 correct trials (678 errors were recorded, that is, 4.6% of errors). These screening procedures will be the same for all subsequent experiments.

Effect of upper vs. lower case presentation

Because we are not interested in the identity of the case but rather in the overall physicality of the stimuli, I tested if there was a significant difference in RTs when both stimuli (S_1 and S_2) were shown entirely as uppercases or entirely as lowercases; this represents half of the total trials. A 14 × 2 ANOVA (0D1L to 4D4L × all-uppercase vs. all-lowercase) showed non-significant results for the effect of case ($F(1, 529) = 0.017$, $p = 0.897$; there were no observations for 1D4L for one participant, hence 529 rather than 530 degrees of freedom). An identical 14 × 2 ANOVA was performed for both conditions in which the stimuli were mismatching on case which also showed non-significant results for case ($F(1, 530) = 0.009$, $p = 0.926$). Consequently, both matching and both mismatching conditions were combined and analyzed irrespective of the actual case's identity.

Mean response times and accuracy

Mean RT and accuracy rates for each condition are presented in Figure 2.2 as a function of string length. Error bars denote the difference and correlation adjusted 95% confidence intervals of the mean (CI; Cousineau, 2005; Morey, 2008) as recommended by Baguley (2012, see Cousineau, 2017).

INSERT FIGURE 2.2 ABOUT HERE

Regarding the "Different" RTs, the results show the typical trend in both priming conditions (case matching and case mismatching): The RTs are slower as the number of letters increased and as the number of differences between S_1 and S_2 diminished. We clearly see the fan-out effect involving both D and L. The "Different" results fit the expectations as there are no notable differences between experimental manipulations for all "Different" conditions: all error bars and mean values almost completely overlap. In fact, a 10 × 2 ANOVA of "Different" conditions as a function of physicality conditions shows no significant results ($F(1, 378) = 1.988$, $p = 0.159$). The overall mean RT for "Different" in the case matching condition is 547 ms whereas it is 558 ms in the case mismatch condition.

Regarding the "Same" RTs, in the case matching condition, the "Same" RTs are fast, being below the fastest "Different" responses. This pattern of result is typical of a fast-"Same" effect. However, the same cannot be said for "Same" results in the case mismatch condition. These RTs are slower than the physically matching condition and return "Same" RTs that are no longer among the fastest of all responses. Thus, the fast-"Same" effect is attenuated in this condition.

The overall mean RTs for “Same” responses in the matching case condition is 502 ms whereas it is 537 ms in the mismatching case condition.

Regarding accuracy rates, there are no differences between physically matching cases and mismatching case trials for all “Same” and “Different” responses; all mean accuracy rates and error bars overlap almost completely. The mean accuracy for physically matching cases is 96.5% and 95.2% for “Same” and “Different” respectively; for physically mismatching cases, they are 95.1% and 94.6% for “Same” and “Different” respectively.

The only condition where many errors occurred is in the 1D condition, which is also the condition where responses take the longest to be made. Hence, it suggests a speed-accuracy trade-off where errors are committed to avoid response times that are too long. The RTs in the 1D condition are possibly underestimated, more so for larger L.

RT slopes

To see if the difference between primed and non-primed experimental manipulations had deeper roots, I measured the slope and intercept for “Same” and 1D conditions for both physically matching and mismatching case trials by running a regression weighed by the condition’s number of trials. Only these conditions were analyzed because the 1D condition is the closest to an exhaustive process and so its characteristics should resemble a “Same” condition most (3 checks and 1 self-termination vs 4 checks and termination). Additionally, these conditions are located at the fan-out extremities as noted by Sternberg (1998). This slope analysis will also be able to assess if priming generates a general decrease in processing time (an intercept effect) or if the effect is letter-based, which would flatten the slopes only. Slopes and intercepts were measured and averaged per participants as well as their corresponding standard errors (standard error of the intercept and standard error of the slope; SE). The average SE was then divided by the square root of the number of participants, an approach formalized by Jeffreys (1931, p. 61, eq. 1), where $\overline{\mathrm{SE}}$ is the average of the estimated standard errors, SE_i is the individual standard errors for participant *i* and *n* is the sample size:

$$\overline{\mathrm{SE}} = \frac{1}{\sqrt{n}} \times \sqrt{\frac{\sum_{i=1}^{n} (\mathrm{SE}_i)^2}{n}} \tag{1}$$

This type of standard error is the within-subject standard error. It is focused on the estimation error within the participant, not from the errors of estimation across participants.

The results of Experiment 1's slope and intercept analyses are presented in the first two rows of Table 2.2. This table shows the average slopes and intercepts as well as the within-subject error of estimation SE (in parenthesis) for each measure. The columns "Experimental manipulation 1" refer to when both stimuli were physically matching and "Experimental manipulation 2" refer to when there is a physical mismatch, a case mismatch in this experiment.

INSERT TABLE 2.2 ABOUT HERE

Regarding intercepts, those for the 1D conditions are higher than that of "Same" decisions in both experimental manipulations, an average slowdown of 30 ms (with a SE of 9 ms). However, there is no difference between matching and mismatching case conditions (i. e. the two 1D intercepts are almost identical to one another and so are the two "Same" intercepts).

Regarding the slopes, "Same" responses are the shallowest regardless of experimental manipulation. However, there is a strong increase in slope when cases physically mismatched. The 1D to "Same" slope ratio goes down from 2.38:1 (39 ms/L vs. 17 ms/L) to 1.45:1 (44 ms/L to 30 ms/L) when the stimuli physically mismatched. This means that the "Same" slopes are almost twice as large relative to "Different" when the stimulus pair physically mismatched compared to when they physically matched.

In sum, the fast-“Same” effect is entirely a slope effect in this experiment. Slope of “Same” responses are reduced whereas slope of “Different” responses are roughly unchanged. This could imply a processing rate that is accelerated when stimuli physically match as well as a different processing mode between experimental manipulations. To eliminate the latter possibility, we must turn to other aspects of the RT distributions.

Higher statistical moments of RTs

To see whether “Same” decisions were processed in a qualitatively different manner when physicality was altered, I examined two additional aspects of the RT data: the standard deviation and the skewness. Similarly to the mean RT analysis, data were aggregated by participants and averaged for all conditions. Results from this analysis are shown in Figure 2.3. The error bars denote the within-subject 95% CI using the appropriate SE and CI estimator (Harding, Tremblay, & Cousineau, 2014) for each descriptive statistic. Note that the error bars for standard deviation are asymmetrical as they are taken from the χ distribution, an asymmetrical distribution. Additionally, the error bars for skewness are all the same size because the SE measure for skewness depends only on the sample size and are therefore identical across conditions.

INSERT FIGURE 2.3 ABOUT HERE

As is seen, both conditions follow extremely similar patterns in terms of standard deviation and skewness and the small visual differences that exist between conditions are unimportant; error bars overlap almost completely between experimental manipulations. Differences between conditions are non-significant for both measures ($F(1, 530) = 3.176$, $p = 0.075$ for standard deviation and $F(1, 530) = 1.107$, $p = 0.293$ for skewness)

Discussion

In this experiment, we can observe that overall trends in the results for physically matching stimuli are identical to those reported in other Same-Different research. Furthermore, the results were observed independent from if the physical match was in uppercase or lowercase letters. This confirms that the classic results are not a by-product of stimulus specificity but on the contrary, are quite robust to changes in materials. Most importantly, we saw that physical matches are essential for the presence of the fast-same phenomenon, a finding matching Bamber's (1972) results that this task aimed to replicate. "Same" response RTs were slowed when the cases mismatched while "Different" results were almost unaffected. This last result is evidence towards the necessary from-scratch processing hypothesis posited above. Likewise, accuracies remained unchanged between experimental manipulations; visually they are practically identical and error bars for all conditions overlap almost completely.

Slopes and intercepts offered compelling information regarding the mechanism at play. The slope is shallower for "Same" responses than the slowest (and steepest slope) difference condition when the stimuli are physically matching. This is congruent with the priming hypothesis that should benefit these conditions only. Nevertheless, when physicality mismatched, the "Same" slopes grew steeper (by a factor of 1.82) and more similar to the 1D slope (although the 1D slope is maybe lowered by a high error rate). This transition in slope also has implications on the various conditions composing "Same" responses. When physical priming is removed, the encoding of stimuli requires a deeper treatment to identify what the stimulus represents. As the priming annulment for longer strings results in a slower response time, it is apparent that stimulus complexity (the letter strings' length) plays a factor as well; there may be some sort of exhaustive construct at play to process all "Same" letters. Finally, the relatively unchanged intercept between conditions suggests that a bias or priming for “Same” decisions remains. As Huber (2008), Eviatar et al. (1994), and Lupker et al. (2015), have noted, priming could be absent on the physical level but still be present phonetically and semantically – this would lead to the observed discrepancy between intercepts of “Same” and 1D.

When it comes to the standard deviation and the skewness of “Same” and “Different” judgements, both show almost identical patterns across experimental manipulation and does not appear to differ from one another (in fact, ANOVAs show that there are no significant differences between experimental manipulations for the mean of these statistics). Therefore, one could conclude that the underlying process, regardless of physical identity, is unchanged. This implies that a change in physicality does not trigger a qualitative change in treatment.

The many similarities of “Same” and “Different” results across experimental manipulations provide insight into the possible mechanism at play. Both matching and mismatching stimuli have indistinguishable trends for intercept, accuracy, standard deviation, and skewness for all conditions composing both “Same” and “Different” decisions. The only notable change stems from the mean RT slopes, prompting us to posit the notion that there is an identical underlying process for both case matched and mismatched stimuli; discrepancies in RT, and therefore the fast-same phenomenon, could be due to an additional factor in the process chain, namely, physical priming. As discussed, even if physical priming is present in the matching condition, it can only benefit “Same” trials due to a residual activation in the processing pathways.

While the letter-case experiment supports a priming view, it is unclear whether the observed results are generalizable to other stimuli or to mismatches of less extreme change. Uppercase and lowercase variants of the same letter can be physically very different from one another and may even have no single feature in common. This high-level of discrepancy may be the cause of the fast-same's attenuation. I therefore replicated this task in Experiment 2 with physical mismatches of a smaller magnitude, changing font and typeface only.

Experiment 2: Font manipulation

To see if the results found in Experiment 1 were caused by important physical changes (such as cases), or whether they are generalizable to minor changes in stimuli physicality, another Same-Different experiment was carried out with subtler physical changes. To do so, font and typeface of the letters were varied. As the attributes used in the present experiment had only minor changes in physicality, results of this task should indicate whether the fast-"Same" effect in the Same-Different task can be attenuated in a gradual manner or if it follows an all-or-nothing rule.

Methodology

Participants included 20 new consenting adults aged 18 to 30 with normal or corrected vision. They were all informed of the experiment's procedure as well as the protocol and ethical rules of the University of Ottawa. They gave written and verbal consent to participate in the experiment. Finally, all participants were compensated $10 for their time.

All stimuli, procedures, and experimental design for this experiment are identical to Experiment 1 except for what follows: Within the experimental manipulation, rather than having matching/mismatching cases, only the font and typeface are varied. For the first typeface, stimuli were the same twelve capital consonants as Experiment 1 written in a non-italic Arial font (ex: JCVD). In the second typeface, letters were italic and written in the Bondoni MT font (ex: *JCVD*). The Arial font was selected as it does not include serifs (the small lines at the end of a stroke) whereas the Bondoni MT font includes serifs, adding additional changes in the overall presentation of the stimuli. Once more, the physically matching group had matching typeface between S_1 and S_2 of which half were presented with each typeface variant. The second half of trials, exactly like the Experiment 1, presented mismatches between physical attributes of S_1 and S_2; all trials were randomly presented. The numbers of trials for each condition are presented in Table 1. Again, participants were instructed to pay no attention to the physical changes of the stimuli and make their decisions solely on the letter identities only.

Results

Screening of the data

I gathered data from 15,360 total trials (768 trials per participant × 20 participants) of which 19 RTs were below 200 ms, 59 were above 2500 ms and 37 were non-answers. Only correct RTs were retained for RT analysis, for a total of 14,521 trials. There were 724 errors in total (5.0% of errors).

Effect of typeface presentation

Once more, to ensure that both physically matching trials and physically mismatching trials were respectively identical, two 14 × 2 ANOVAs (0D1L to 4D4L × both matching conditions; 0D1L to 4D4L × both mismatching conditions) were conducted which both returned non-significant results ($F(1, 558) = 0.013$, $p = 0.908$ and $F(1, 558) = 0.108$, $p = 0.742$ for the matching and mismatching conditions respectively). Therefore, all matching conditions were merged together, and all mismatching conditions were merged together.

Mean response times

Mean RT and accuracy rates for each of the experimental manipulations are presented in Figure 2.4 in the same format as Figure 2.2. Error bars denote the difference and correlation adjusted 95% CI of the mean.

INSERT FIGURE 2.4 ABOUT HERE

As shown, the results match what is expected from a standard Same-Different task when there is a physical match. There is the presence of the fast-same phenomenon with physically matching stimuli, and “Different” decisions fall where they are expected. The overall mean RT for “Different” in the physically matching condition is 602 ms whereas it is 597 ms in the physically mismatching condition. A 10 × 2 ANOVA shows no significant results between both physicality conditions for “Different” RT ($F(1, 398) = 0.177$, $p = 0.674$).

Much like Experiment 1 above, “Same” decisions RT were severely attenuated when the physicality between stimuli did not match while “Different” decisions were seemingly unaffected by stimulus physicality. Error bars of both experimental manipulations for all conditions overlap other than for “Same” decisions implying that they are not different from one another. Overall “Same” RT for the physically matching trials is 542 ms compared to 566 ms for the physically mismatching trials.

There are no differences for any of the accuracies (“Same” nor “Different”) across experimental manipulations; all error bars and mean values overlap considerably. Overall accuracies for “Different” decisions are 94.5% and 95.0% for both physically matching and mismatching conditions respectively and 96.0% and 95.5% for “Same” decisions in the same order. By the analyses thus far, results are identical to those observed in Experiment 1.

RT slopes

To see whether the attenuated slope effect observed in Experiment 1 are also is present in this experiment, the overall slopes and intercepts of the 1D and “Same” conditions were measured for both experimental manipulations as well as their respective SE by using a regression weighed by condition's number of trials. The results of this analysis are presented in the second section of Table 2.2. Experimental manipulation 1 refers to physically matching trials whereas Experimental manipulation 2 refers to physically mismatching trials.

As seen, once more, the intercepts for 1D are considerably higher than those of “Same” regardless of experimental manipulation. All intercepts are also quite comparable within their respective condition and SE are small, indicating unidentical baseline processing rate for both “Same” and “Different”. The slope analysis yields the same general results as well. The 1D:Same slope ratio for physically matching trials is 2.04:1 (51 ms/L vs 25 ms/L) whereas the same ratio for physically mismatching trials is 1.20:1 (44 ms/L vs 36 ms/L), which represents a slope shift of a magnitude of almost one and half times for “Same” slopes (the physically mismatching “Same” trials have a slope that is 1.44 times larger than those of the physically matching “Same” trials). Once more, the SE of these values are quite small and strongly suggests that the slope relation is conditional to identity priming, at least at the orthogonal level.

Higher statistical moments of RTs

To further posit a comparable processing mode across experimental manipulations, results regarding standard deviations and skewness are shown in Figure 2.5. Once more, the error bars denote the correlation and difference adjusted 95% CI using the appropriate SE and CI estimator (Harding et al., 2014)

INSERT FIGURE 2.5 ABOUT HERE

Results of this experiment give little supplemental information to the results of Experiment 1. All trends are nearly identical and the statistics of both “Same” and “Different” decisions within both experimental manipulations are visually indistinguishable. In fact, the error bars for both decisions overlap so much that it would be impossible to conclude that any of the conditions are different from one another. For both standard deviation and skewness, there are no significant differences between experimental manipulations (*F*(1, 558) = 0.024, *p* = 0.876 for standard deviation and *F*(1, 558) = 0.131, p = 0.717 for skewness).

Discussion

Results from this experiment show that slight changes in the physicality results in identical trends to those of more extreme change that we saw in Experiment 1 above. As is the case with Experiment 1, RT trends for mismatching “Different” decisions are not significantly different from their matching counterparts and fit what is expected from a Same-Different task replication. Moreover, the accuracy trends for both matching and mismatching stimuli match what is expected from the task as well as match the results from Experiment 1. Most importantly, when stimuli are subtly physically mismatched, the fast-“Same” effect is abolished as is the situation, conditional on case, in Experiment 1. These experiments lead me to believe that a physical match, and the priming mechanisms associated with that factor, is necessary for the operation of the fast-same phenomenon. Further in-depth analyses on distribution spread and shape revealed no new information compared to the previous experiment and the same goes for slope and intercept analyses.

I conclude that a change in physicality, as minor/major as it may be, results in an attenuation of “Same” RT whereas nothing else seems to be affected (as is corroborated by the non-significant differences for “Different” mean RT and statistics of higher moments). Changes in “Same” responses times seem to be at the surface level and the mechanisms, processing the stimuli are unchanged.

Experiment 3: Visual vs. auditory stimuli

In this experiment, I explored the role of priming and the mental representation hierarchy by constraining processing to the phonological level and up. By presenting the first stimuli aurally to the participants, it is possible to force their mental representation up the mental representation hierarchy and bypass any influence of visual priming altogether. If the priming hypothesis holds true, the fast-same phenomenon should be cancelled, and we should observe the same patterns found in Experiment 1 and 2's physically mismatched condition. As in the previous experiments, there is also a control condition in which stimuli are presented visually as I have done in all other Experiments presented so far.

Methodology

Participants included 20 adults aged 18 to 30 with normal or corrected vision. All participants were informed of the experiment's procedure as well as the protocol and ethical rules of the University of Ottawa. They gave written and verbal consent to participate in this task. Finally, all participants were compensated $10 for their time.

All stimuli, procedures, and experimental design for this experiment's control condition are identical to Experiment 1. The only difference in both priming conditions was the insertion of an additional stimulus consisting of underscore(s) presented between the fixation point and S_1. The underscore(s) was/were presented for 500 ms and their number matched the length of the string. In the auditory condition, S_1 was broken down into its individual letters and pronounced serially (1 s per letter; for example, 4L stimuli took 4 s to present) through desktop speakers in their left-to-right visual order. The presentation language was assigned non-randomly based on the participant's primary language, either French or English. There was an example stimulus prior to the testing phase to ensure that participants could hear and identify the presented stimuli. The auditory stimuli were generated using a voice synthesizer (DSpeech, version 1.56 and the voices of Microsoft Anna and ScanSoft Virginie for English and French respectively). As usual, S_2 was presented visually and participants were instructed to issue decisions by the same conjunctive decision rule used in Experiments 1 and 2.

Figure 2.6 presents a schematic of the updated procedure used only for this experiment.

INSERT FIGURE 2.6 ABOUT HERE

Results

Screening of the data

Data was gathered from 14,592 total trials (768 trials × 19 participants; 1 participant was removed during the analysis for having very slow RT with mean of 1238 ms) of which 3 were below 200 ms, 86 were above 2500 ms, and 11 were non-answers. For RT analyses, I filtered out erroneous trials as well as non-answers, resulting in a total of 13,690 trials. There were 802 total errors (5.5% error rate for total trials).

Mean response times

Mean RT and accuracy rates for each of the experimental manipulations are presented in Figure 2.7. Error bars denote the correlation and difference adjusted 95% CI of the mean.

INSERT FIGURE 2.7 ABOUT HERE

For visual trials, the usual Same-Different mean RT and accuracy patterns are found. Mean RT is 530 ms and 608 ms for “Same” and “Different” decisions respectively and 94.0% and 95.4% accurate for both decisions in the same order. For auditory trials, all the response times are much slower with a grand mean of 737 ms. Despite this main effect, “Different” conditions all fall where they are expected and the fast-same phenomenon was abolished – identical to the previous experiments where the fast-“Same” was cancelled, “Same” responses are still faster than the 1D condition. This result holds considerable weight because it shows that comparing a phonological mental representation to a visual mental representation is faster for “Same” stimuli than for “Different” stimuli; this will be further evidenced by the slope and intercept analyses below. Mean “Same” RT for auditory trials are 728 ms while they are 745 ms for “Different”. Participants are 93.7% accurate for “Same” decisions and 94.8% accurate for “Different” decisions; overall mean accuracy rates are congruent with what was found in previous experiments.

RT slopes

As for slope analyses, results are presented in the third section of Table 2.2. Condition 1 refers to the visually presented S_1 and condition 2 refers to the auditory S_1. For intercepts of the visual trials, there is no new information worth noting. Intercepts for “Same” stimuli are lower than those for “Different” stimuli by a magnitude of ~90 ms. However, the slopes are almost at a 1:1 ratio (20 ms/L:20 ms/L for Same:1D). This lackluster slope ratio contradicts what we have observed for fast-“Same” responses so far in Experiment 1 and 2 and is not representative of the ~2:1 ratio we expect to see. Nevertheless, because there is a task-switching demand in this paradigm, it is unsurprising that some results have not carried over. As for the auditory stimuli, intercepts for “Same” are still quicker than the 1D condition and the slopes follow a ~1:1 ratio which is expected for attenuated-“Same” stimuli. However, we observe a very steep slope compared to the other tasks (140 ms/L vs. 40 ms/L). This means that stimuli are processed at a much slower rate. This could be due to serial nature of S_1 that might force a more serial mental representation/retrieval.

Higher statistical moments of RTs

As for standard deviation and skewness analyses, results are presented in Figure 2.8. Error bars denote the correlation and difference adjusted 95% CI constructed using the appropriate SE and CI estimator (Harding et al., 2014).

INSERT FIGURE 2.8 ABOUT HERE

As shown, results from the visually presented stimuli do not differ from all those seen up to date. With this, there is clearer evidence that the processing type for visually presented Same-Different trials is identical regardless of experimental manipulation (i. e., that the patterns we observe are correlated with a visual presentation of the stimuli). Regarding the auditory stimuli, the standard deviation follows an upward trend as the letter string length grows. Although this is an abnormality considering all the previous results, this could be due to the serial presentation of S_1, and thus its mental representation. This could also be due to interference of S_1 (the presentation of another stimuli hampers the recall of previous stimuli) and decay (stimuli that have been held longer in memory may be harder to recall) of the mental representation over time. As for skewness, there are no abnormalities worth noting. This should therefore limit any reservations we have towards the standard deviation measures. Nevertheless, there exists a significant difference between both experimental manipulations for both standard deviation and skewness ($F(1, 530) = 47.135$, $p < 0.000$ for standard deviation and $F(1, 530) = 20.797$, $p < 0.000$ for skewness).

Discussion

In this task, I presented S_1 to participants through two different sensory modalities. For visual stimuli, there is no unexpected finding; the results concur with all the physically matching results found in Experiments 1 and 2 above. The RT, accuracy, standard deviation, and skewness graphs as well as the intercept measures all follow the same trend. The only difference stems from the RT slope ratio of 1:1. As mentioned, this could be due to task switching but also backwards masking of S_1 due to the underscores presented just moments prior. Regarding auditory trials, they further the hypothesis that perfectly matching visual information is necessary for a fast-“Same”. Surprisingly, almost all results (barring the standard deviation) are identical to visual trials. This leads one to believe that a mental representation of S_1 is created; it is this representation that S_2 is compared to.

To continue this line of thought, one must continue up the representation hierarchy and explore mental representations that are restricted to the semantic level. If the attenuation of “Same” responses persists, one could easily attribute the fast-same phenomenon and its earliest cancellation to physical matching.

Experiment 4: Long-term memory associations

In this last experiment, the role of priming in the Same-Different task is further explored by comparing results from strictly bottom-up stimuli (as is the case with the three experiments presented above) with strictly top-down stimuli (stimuli held in LTM). In the present experiment, and for both experimental manipulations, participants were required to memorize four stimuli of specific lengths (a single stimulus was memorized for each of the four possible letter conditions per participant). In the “letters shown” condition (S_1-Shown henceforth), the letters were presented directly to the participant as usual. In the LTM retrieval condition (S_1-Absent henceforth), the only visual stimuli presented was a cue to the length of the stimuli; from this cue, the participant had to access LTM to retrieve the memorized string before it can be compared to S_2. In this condition, neither the physical priming, nor the phonological levels are activated; the only mental representation comes from the semantic link existing between the number of underscores and the memorized strings. It is believed that a fast-“Same” will not be observed in the S_1-Absent condition as priming cannot operate on and benefit from physical matches, while the effect will be seen in the S_1-Shown manipulation (this is a standard Same-Different task). I also predict that overall decisions will be more accurate in both experimental manipulations due to the very small number of stimuli presented over and over again, possibly yielding training effects. Additionally, due to these training effects and the atypical experimental design, it was uncertain that the typically observed RT and accuracy patterns would be elicited from this task. Finally, I was unsure regarding the results skewness and standard deviation analyses would yield as the task is shifting from a comparison task to a LTM retrieval + comparison task.

Methodology

Once more, all stimuli, procedures, and the experimental design for the control condition are identical to Experiment 1 except that the letters are not chosen randomly on every trial (see below). Participants included 21 adults aged 18 to 30 with normal or corrected vision. However, only 19 were retained for analysis; one participant was removed after post-testing analyses showed a below 50% accuracy performance in the task. Another participant was removed during the analysis period for having both unusually and consistently slow RTs (a mean of 1128 ms).

All participants were informed of the experiment's procedure as well as the protocol and ethical rules of the University of Ottawa. They gave written and verbal consent to participate in this task. Finally, all participants were compensated $10 for their time.

In a preliminary phase, participants were asked to memorize four separate strings of letters (randomly generated for each participant), each corresponding to a specific string length. For example, if the memorized four-letter string was "JCVD", all 4L S_1 would be presented as "_ _ _ _" or "J C V D"; if the memorized two-letter string was "CN", all 2L S_1 would be presented as "_ _" or "C N". To ensure that the four memorized strings were learned adequately, this was followed by a memory test of all four target stimuli: 1 to 4 dashes were shown in a random order and the participant had to type the appropriate string. The preliminary phase was repeated until participants could type all four memorized items without error twice in a row.

In the second phase of the experiment, the experimental design was the same as in the previous experiments. However, rather than randomly generating a different S_1 on every trial, only the four memorized strings were used. On half of the trials, S_1 was presented as underscores matching the length of the stimulus, and on the other half, the actual memorized stimulus was presented. Therefore, there were only four stimuli in the entire task and training results are expected.

Results

Screening of the data

I gathered data from 14,592 total trials (768 trials × 19 participants) of which 17 were below 200 ms, 98 were above 2500 ms, and 22 were non-answers. For RT analyses, I filtered out erroneous trials as well as non-answers for a total of 13,463 trials. There were 1000 total errors (6.9% error rate).

Mean response times

Mean RT and accuracy rates for each of the experimental manipulations are presented in Figure 2.9. Error bars denote the correlation and difference adjusted 95% CI of the mean.

INSERT FIGURE 2.9 ABOUT HERE

For S_1-Shown trials, RT and accuracy patterns for both “Same” and “Different” trials resemble the classic Same-Different trend that we have observed throughout this manuscript – mean RTs are 543 ms and 609 ms for “Same” and “Different” decisions respectively and accuracy is 96.4% and 96.1% for both decisions respectively. As for the S_1-Absent experimental manipulation, RT results show a much different perspective: “Same” results are much slower than their S_1-Shown counterparts and are much slower than most “Different” RT except the 1D conditions leading me to conclude that the fast-“Same” is attenuated considerably (error bars do not overlap for both “Same” conditions across experimental manipulations). Moreover, although “Different” conditions fall where they are expected, they are much slower than their S_1-Shown counterparts. Mean “Same” RT are 661 ms while they are 689 ms for “Different”. However, accuracy shows almost identical patterns across experimental manipulations; mean accuracy results are almost identical and error bars for all conditions overlap – participants are 93.5% accurate for “Same” decisions and 94.9% accurate for “Different” decisions.

RT slopes

As for slope analyses, results are presented in the fourth and last part of Table 2.2. Experimental manipulation 1 refers to S_1-Shown stimuli while Experimental manipulation 2 refers to S_1-Absent stimuli.

As is seen, the intercepts for 1D are all higher than those for “Same”, typical of the results shown thus far. Additionally, as noted with the mean RT analysis above, the S_1-Absent trials are much slower than the S_1-Shown trials; the intercepts increase by a magnitude of roughly 100 ms. As for the slopes, it is 1.3: 1 for the S_1-Shown condition but below 1:1 in the S_1-Absent condition. The slope increment from 3L to 4L in both 1D and “Same” trials for the S_1-Absent trials flattens out and affects the overall measure. Therefore, the sub-1:1 ratio for slopes may be biased even if the regression is weighed by the number of trials in the conditions with lots of errors.

Higher statistical moments of RTs

As for standard deviation and skewness analyses, results are presented in Figure 2.10. Error bars denote the correlation and difference adjusted 95% CI constructed using the appropriate SE and CI estimator (Harding et al., 2014).

INSERT FIGURE 2.10 ABOUT HERE

As is seen, results from the analyses are just as conclusive as results presented in Experiments 1 and 2 above. Standard deviation and skewness are stable across L and D conditions. Although all conditions are almost identical across experimental manipulation and error bars overlap quite considerably, a 14 × 2 ANOVA for standard deviation yielded significant differences across experimental manipulation ($F(1, 530) = 10.211$, $p = 0.001$) while an identical 14 × 2 ANOVA for skewness did not ($F(1, 530) = 3.205$, $p = 0.074$).

Discussion

As seen, RT in the S_1-Shown condition are similar to typical Same-Different tasks (physically matching conditions of Experiment 1 and 2 and the visually presented stimuli of Experiment 3; the only experimental manipulations thus far which allows the presence of fast-"Same" responses). This can be attributed to the possible training effects of S_1: as the participant is repeatedly shown the S_1 stimuli of each string length, encoding the latter is streamlined because the individual components of each of the four possible S_1 stimuli are identical. Therefore, participants were only required to attribute their attentional resources to detecting matches and mismatches presented in S_2 – as soon as a mismatch is present, all attentional resources can be attributed to it. While the overall pattern for "Different" results matches what is expected, they are slower than all S_1-Shown trials. All "Different" conditions from the S_1-Absent experimental manipulation are significantly slower than their S_1-Shown counterparts. As the training effects are the same, the only difference between experimental manipulations is the presence of the stimulus itself in the S_1-Shown trials. This last remark is of importance as it further highlights the fact that physical stimuli seem to be at the forefront of the functioning of the Same-Different task. While changes in “Same” RT have been explored to this point, “Different” results do not have any speed decrements other than in this experiment. The discrepancy in “Different” RTs here could be attributed to the extra process required of S_1: a retrieval of stimuli stored in LTM (top-down processing). It is therefore apparent that training effects, regardless of top-down or bottom-up stimuli, will lead to an identical pattern of results.

This novel study offers interesting insight into the workings of the Same-Different task. In this task, no priming was possible in half the trials because no visual information was presented. In these trials, the fast-same was severely attenuated and RT slope was much closer to the 1D line. Consequently, the results show that visual information is not necessary to complete the task but is necessary for the presence of the fast-same phenomenon.

General discussion

In all the above experiments, I manipulated the amount of priming in four Same-Different task variants. These variants included two in which S_1 and S_2 could be physically matching or mismatching, a third where mental representation was limited to phonology, and a fourth where S_1 was not presented and thus long-term memory was involved. Throughout these experiments, when priming was not intended to be altered (physically matching trials, visually presented trials, and S_1-Shown trials, the “control” conditions if you will), RT and accuracy for all “Same” and “Different” conditions fit exactly where they are expected when considering the literature (see the expected results section above for a detailed overview). As for trials intended to cancel priming, all experimental manipulations (physically mismatching trials, auditory trials, and S_1-

Absent trials), tell the same story: Accuracy and RT is consistent in “Different” conditions compared to their “control” counterparts whereas the speed of “Same” decisions is severely attenuated. This change in speed also entails a change in slope representing “Same” speed attenuation. Fast-“Same” decisions have a shallow slope at the bottom of the fan effect, as observed by Sternberg (1998) whereas attenuated-“Same” decisions have a longer RT as well as a steeper slope approaching the top of the fan effect (the 1D conditions in all experiments). Accuracies for the attenuated-“Same” decisions are not different from their fast-“Same” counterparts.

It is my belief that the attenuated-“Same” effect is equivalent to a “Same” decision with no added benefits whatsoever. This impression stems from the fact that the attenuated-“Same” follows the expected slope from the fan-out noted by Sternberg (1998). The results also strongly suggest the necessity of a physical match between S_1 and S_2 to obtain fast-“Same” responses. This change in slope between fast-“Same” and attenuated-“Same” RT also offers insight into the nature of the fast-same phenomenon. It seems to be the case that “Same” responses elicit some sort of “pop-out” effect (Treisman, 1985; Wolfe, 1991) when S_1 is primed in any of the three ways experimented above. Although the pop-out effect is often defined as having participants quickly identify a target in a sea of identical foils, one could believe the same idea applies here; “Same” responses seem to be quite evident to the participant. This elicits an advantage to press “Same”. This hypothesis agrees somewhat with Bamber’s (1969) position of “Same” responses necessitating a template matching process. However, rather than have a dual process *template vs analytical model* with the template model dictating decisions, one could have an analytical model with a matching detector. In short, if the template benefits the *detection* of S_2, the *decision* is processed analytically with a higher rate of accrual. If the template does not match, no benefit is given to the accrual of “Same”.

This template-gate hypothesis is supported by the nROUSE model of residual activation (Huber, 2008; Huber & O’Reilly, 2003) where stimuli are compared and processed in a series of hierarchical levels. The lower physical matching level acts as a gate-keeper as to whether the stimuli can be fast-tracked or not. If the stimuli are physically mismatching, then they must go through all levels analytically as is the case for all “Different” decisions; residual activation can therefore not operate or benefit the response. This fast-track makes for a speeded response that seem obvious for “Same” decisions whereas “Different” responses operate “from-scratch” on every trial. For “Same” trials that must be fully analytically processed, with no benefit in accrual rate, string length is directly proportional to RT as it is the case with many cognitive architectures (Townsend & Ashby, 1983), hence the steeper slope.

As for the other measures of RT distribution, spread and skew, they are conclusive in their inconclusiveness. There are no changes whatsoever between experimental manipulations; it suggests that the way the participants process stimuli and compare elements is rigid across experimental manipulations (other than the auditory stimuli, but this change could be due to the serial, presentation of S_1). All experiments offer the same, relatively flat slope across string length. A dual-process model, or any change in processing mode across manipulation would more than likely have resulted in some sort of modification to either descriptive statistic, especially if we consider that “Sameness” can be assessed by a fast-parallel process whereas “Differentness” must be assessed by a serial process (as proposed with the Identity Reporter vs. serial self-termination model of Bamber, 1969). This lack of change offers the evidence that both “Same” and “Different” decisions operate with the same, single, underlying process.

Nevertheless, while there are no clear differences, this offers a glimpse into the cognitive architecture at play for either decision. A decrease in skewness typically leads one to believe that the distribution normalizes as stimuli size increases owing to the central limit theorem. The latter is strongly associated with serial architectures. However, the fact that it is constant (no significant differences, equivalence tests would need to be performed to assess whether they are identical) across string length implies parallel processing. As for standard deviation, there are no changes which implies a parallel network as well. Standard deviations should increase if the number of comparisons is increased (as is the case for serial architectures) when processing is not exhaustive.

In all experiments, the only manipulated factor was priming and the only factor affected is “Same” mean RT and RT slope. This implies a strong relationship between the two to observe such a targeted effect that does not change how “Different” decisions are taken. Priming is a very surface level process modulated directly by the stimulus itself. Therefore, changes in its capability to operate do not affect the deeper levels of processing.

Conclusion

Hampering the ability of the criterion stimulus to prime the test stimulus abolishes (or at least attenuates) the fast-same phenomenon altogether without having any impact on the speed of “Different” decisions; the slopes for “Same” responses were also attenuated whereas no change was observed for “Different” conditions. The evidence shows, for the first time (to the best of my knowledge and extensive literature review), that this change in “Same” speed not only has no effect on the accuracy patterns of either decision, but also has no effect on measures of spread and skew across experimental manipulations. The fact that the only variable affected is priming and the only observed change was in “Same” RT and slope offers an elegant and parsimonious approach to explaining the fast-same phenomenon. I believe, and concur with the research brought forth by Nickerson (1978) and Proctor (1981), that priming seems to be at the core of performance that has eluded modelling.

While many other processes could be at play (of which a list is presented in the introduction above), none have been as diligently assessed as I have done here and has offered as many positive results. Future research should therefore expand the analyses herein to other experimental manipulations and, ideally, highlight a particular modelling approach to see if they can reproduce the strong effects observed in my experiments. Future research should also expand on the supplementary analyses presented here and strive to model the core process behind both “Same” and “Different” decisions; although I do not delve into the processing debate for either decision in this Chapter, it seems as if the comparison process is done with some sort of parallel process. This will be further explored in Chapter 3. Nevertheless, priming should be integrated into whatever models are developed henceforth to account for both fast-“Same” RT and attenuated-“Same” RT. Finally, the fact that “Same” RT can be attenuated is important but there is no information on why they remain faster than that of the slowest “Different” condition. I believe this to be the next step in this research. Measuring stopping-rule and capacity as a function of string length may be the key to make progress. Alternatively, it is necessary to test if there is an inherent “Same” bias as noted by Ratcliff and Hacker (1981). However, intuitions gathered from this Chapter lead one to posit that behavior reflects either residual activation in the upper semantic levels or possibly threshold differences.

Publication intentions

Research for the Same-Different task has been stagnant for decades and modelling the task remains an obstruction in fundamental cognition research. The experiments presented here aim to overcome this research hurdle and attempts to explain the elusive fast-same phenomenon, a feat that has yet to be surmounted in over half a century of research. Herein lies the impact of this work and I intend to submit this Chapter in a journal that reflects this importance. I believe that these four variants of the Same-Different task offer compelling evidence in identifying priming as the primary mechanism behind this thoroughly replicated phenomenon.

Chapter 3: Evidence of coactivity within the Same-Different task

What is coactive processing and how do we identify it

There exist many different types of information processing architectures that can model cognitive phenomena of which a few were presented in Chapter 1. Notably, there are serial architectures where information is treated sequentially and parallel architectures where it is treated all at once. However, there also exists another option that is much less studied and researched, coactive architectures of information processing. Coactive architectures can be considered as a parallel, *collaborative*, architecture of information processing where a decision is made when a sufficient amount of information has been gathered, regardless of its source (Miller, 1982; Schwarz, 1989; see Townsend & Nozawa, 1995, and Harding, Goulet, Jolin, Tremblay, Villeneuve, & Durand, 2016 for reviews). This means that relevant evidence stemming from redundant sources of information jointly accumulate towards a single threshold; the more redundant sources of information there are, the faster the overall process will be (Kinchla & Collyer, 1974; Holmgren, Juola, & Atkinson, 1974; van der Heijden, 1975; Miller, 1982; van der Heijden, La Heir & Boer, 1983; van der Heijden, Schreuder, Maris, & Neerincx, 1984; Proctor & Healy, 1985; Mordkoff & Yantis, 1991, 1993; Gottlob, 2007; Ben-David & Algom, 2009; Gondan, Götze, & Greenlee, 2010; Castro, Wasserman, & Young, 2012; Engmann & Cousineau, 2013).

Coactive architectures are often characterized by specific response time and accuracy distribution signatures. First, mean RT becomes quicker as redundancy increases. Second, RT distributions increase in skewness as more redundant sources of information accumulate together (see the *Extreme Limit Theorem*, Cousineau, Goodman & Shiffrin, 2002). Third, with the addition of redundant information, standard deviation for RT distributions reduces, which opposes the central limit theorem associated with serial architectures. Fourth, unlike the well-studied speed accuracy trade-off (first introduced by Henmon, 1911; see Heitz, 2014, for a review), coactive architectures become more accurate as more sources of relevant information are present (Engmann & Cousineau, 2013; Grice & Gwynne, 1987; Miller, 1982). This is because with more sources of redundant information, it becomes easier to identify a target dimension (the probability of seeing relevant information increases). For example, within a redundant target detection task (the task classically used to exemplify coactivity), if the target dimensions are “red” and “square”, the presentation of a red square (redundant condition) will yield faster and more accurate responses than simply a square (single-target condition) or a red stimulus (single-target condition; Engmann & Cousineau, 2013; Feintuch & Cohen, 2002; Miller, 1982; Mordkoff & Yantis, 1991, 1993).

While coactive architectures can be inferred indirectly by observing the distribution characteristics noted above, they can be quantitatively identified by measuring the *Race Model Inequality* (RMI; often colloquially referred to as the *Miller inequality*, or *Miller bound*; Miller, 1978, 1982). In this analysis, the cumulative density functions (CDF) of the two RT distributions trials with a single target dimension are plotted. Afterwards, these two conditions are combined to generate a third, composite, CDF that simulates the fastest theoretical processing speed of a standard parallel race model. This composite bound is then compared to the empirical CDF of the redundant target condition. If the empirical CDF dominates (is above) the CDF at some points along the composite function, one can assume that the addition of a second target improves overall decision speed more than what a parallel unlimited capacity model could predict, indicating redundancy gains. This inequality measure, if breached, offers strong evidence against standard parallel models and suggests coactivity.

Redundancy effects within the Same-Different task

Unsurprisingly, redundancy effects and coactive architectures have been present, or at least alluded to, in past Same-Different research. Egeth (1966) noted that with the addition of relevant target attributes, “Same” decisions are faster whereas the addition of irrelevant target attributes lead to identical result trends; “Different” decisions are also unchanged. This was later corroborated in a similar experimental design (Miller & Bauer, 1981; which, along with Miller’s 1978 work, directly inspired his research on coactive models, Miller, 1982) where relevant information also translated into speed gains in a comparison task.

Nevertheless, what is surprising is that a proper assessment of coactivity in the Same-Different task has not been conducted nor has coactivity been formally explored in any literature reviews noted in Chapter 1 (Farrell, 1985; Sternberg, 1998; both of which were published after Miller and Bauer’s, 1981, work). This is especially surprising considering the seemingly intuitive link between the paradigm and the architecture – for “Different” responses, decisions are faster when more mismatches are present within strings of the same length (i. e., the addition of relevant targets leads to faster responses). This begs the question: Is this speed increase, conditional on the number of differences, a consequence of redundancy gains?

In this Chapter, I will explore the possible existence of coactivity within the Same-Different task with a twofold approach: First, I will test the race model inequality for eligible “Different” conditions using the appropriate inequality bound. More thoroughly, the composite bound will be created with the combination of both 1D2L conditions (difference on the first letter, difference on the second letter; both single-target conditions). Afterwards, the 2D2L (the redundant condition with differences on both letters) condition's CDF will be compared to this composite bound. If coactivity is present in this condition, it is expected that dominance of the all-"Different" condition on the composite bounds will be observed. I will also conduct the same analysis with 3L stimuli with the relevant bounds.

Second, the "Same" responses cannot be analyzed with inequality bounds because they do not have the conditions necessary to create the composite bound: there are no trials in which only some information points to a "Same" response. Instead, I will investigate RT distributions for coactive evidence. This approach will allow me to see if coactivity can underlie both "Same" and "Different" decisions, and thus, give evidence towards a single process model of the Same-Different task (of which I present a possibility in Chapter 4). However, because standard "Same" responses might be affected by threshold variations tied directly to string length (4L stimuli require 4 times more information than 1L stimuli for an identical "Same" response), I introduce a variant to the Same-Different task. In this novel task, participants will have two possible S_2 conditions requiring a "Same" response. The first condition is the standard conjunctive "Same" and the second is called the *wildcard*-"Same" response. This wildcard is made entirely of the letter "X" and is the same length as S_1. Participants are instructed that as soon as an "X" is presented during S_2, they should treat it identically as a "Same" response (hence the wildcard name – akin to wildcard characters in the UNIX operating system). This new decision condition not only limits the possible thresholds effects (regardless of *l* condition, participants only need to detect a single "X") but keeps the redundant nature from the standard-"Same" responses. The more information presented, the more information relevant to answer "Same" is available and should lead to an increase in speed and more accurate responses. Standard deviation and skewness will also be assessed to see whether the coactivity artifacts noted above are present.

Experiment 1: Coactivity bounds in a classic Same-Different task

To properly assess the RMI within the Same-Different task, a variant of the original task is used to ensure the cleanest source of data possible. In this first experiment, S_1's length was fixed per quarter block to remove variability participants might have by adjusting their attention to the random letter string length on every trial. By removing this complication, the results of the RMI tests should be the most representative of each L condition.

Methodology

Participants

Twenty-four undergraduate and graduate students aged between 17 and 30 years of age were recruited via word of mouth and internal advertising at the University of Ottawa to participate in this task. All participants were informed of the experiment's procedure as well as the protocol and ethical rules of the University of Ottawa; all gave written and verbal consent to participate in this task. Upon the completion of the task, participants were compensated $10 for their time and effort. Following the task, each participant was debriefed on the objective of the study.

Stimuli

Stimuli were displayed on a calibrated CRT display that had a resolution of 1024 × 768 pixels, a 85 Hz screen refresh rate, and a luminance and RGB standard. Twelve capital consonants (B, C, D, F, J, K, L, N, S, T, V, and Z) were selected as stimuli to replicate as best as possible Bamber's (1969) original study. The stimuli could be presented within a 10° × 10° visual angle centered on the screen with the first string shown 4° above the center of the computer screen and the second shown 4° below the center of the screen (assuming the participant was seated roughly 50 cm from the center of the screen). Stimuli were always presented as white letters on a black background to create a strong visual contrast that ensured stimuli were clearly perceived.

The letters were randomly selected on every trial and string length could vary from 1 to 4 letters. No letter was presented twice within the same stimulus and matching letters would appear in the same position for both S_1 and S_2. For “Different” conditions, the differing letter(s) in S_2 were mutually exclusive with letters included in S_1; S_2 could have 0, 1, 2, 3, or 4 differences to S_1.

Procedure

Prior to the experimental phase, participants were informed of the general goals of the experiment without inducing any bias towards “Same” or “Different” responses. They were then led to a secluded room that housed the computer where they were seated for the remainder of the experiment (unless it was instructed on screen that they could take a short break, or if they chose to leave the experiment). A written copy of the experiment’s instructions was presented on screen. The on-screen instructions instructed participants that their decisions would be recorded by pressing the "CTRL" key located on the far left of the keyboard using their left hand and the "ENTER" key located on the far right of the keyboard (on the numeric pad) using their right hand. The “Same” or “Different” decision associated with each button was counterbalanced based on the participant number with half pressing “Same” with their left hand. Once the participant was ready and they had verbally acknowledged that they understood the procedure, they began the experiment. The timing of a trial was identical to the timing of a trial in Experiments 1, 2, and 4 from Chapter 2 and is graphically shown in Figure 2.1.

As shown in Figure 2.1, a fixation cross was presented for 500 ms, followed immediately by the presentation of S_1, which was presented for 400 ms. Afterwards, a blank screen was presented for 400 ms followed by the presentation of S_2, the test stimulus. The test stimulus was shown for 5 s or until a decision is made. Feedback was given for 500 ms in case of either a non-response or an error in order to avoid diverting the gaze of the participant when they correctly answered; for correct answers, the screen remained blank for 500 ms. Afterwards, there was a 500 ms blank screen before the next trial began. Although the task was easy, the participant was given a message in red to pay attention if he/she made more than 5 mistakes in a row. Additionally, the participant was offered short breaks after every 192 trials (1/4 blocks of the experiment's total number of trials).

Following testing, all participants were given a debriefing to answer any queries and to explain the goal of the study. Once they felt sufficiently informed, participants left the testing area.

Experimental Design

This experiment included 768 trials; the experimental design is shown in Table 3.1.

INSERT TABLE 3.1 ABOUT HERE

As shown, just like the design for experiments in Chapter 2, half of the total trials were "Same" and half were "Different". Additionally, within each decision, there was an equal number of 1L, 2L, 3L, and 4L stimuli blocked in quarters of 192 trials (768 trials total). Within each of these blocks, "Same" and "Different" responses had an equiprobable chance of occurrence (1D, 2D, 3D, and 4D each occurred the same number of times). The positions of the differences within the strings were also random. There were 24 possibilities (${}_4Pr_4$) of length block order and each was associated to a participant number that ensured that the task was as balanced as possible. This explains why this experiment recruited more than the usual 20 participants.

Results

Screening of the data

Data was gathered from 18,432 total trials (768 stimuli × 24 participants; one for each block order permutation). There were 16 RT slower than 200 ms that were removed as well as 55 non-answers; no answer was slower than 2500. For all RT analyses, I filtered out erroneous trials to arrive at a total of 17,448 valid data (913 errors were recorded representing a 5.0% error).

Mean response times

Mean RT and accuracy rates for each condition are presented in Figure 3.1. Standard deviation and skewness are also presented. Error bars denote the difference and correlation adjusted 95% confidence interval (CI; Cousineau, 2005; Morey, 2008) as recommended by Baguley (2012) using the appropriate SE and CI estimator (Harding, Tremblay, & Cousineau, 2014)

INSERT FIGURE 3.1 ABOUT HERE

When compared to data from similar experiments (Bamber, 1969, Sternberg, 1998) as well as the control conditions of Chapter 2, this Same-Different task variant returns the expected trend. “Different” RTs are slower as the number of letters increases and as the number of differences between S_1 and S_2 diminishes. The “Same” RTs are fast, faster than the slowest “Different” responses (1D). This pattern of results is prototypical of a fast-“Same”. Mean RT for “Same” are 493 ms with a 95.2% accuracy rate, while the mean RT for “Different” is 534 ms with a 94.8% accuracy rate. This is of note as it shows that a change in experimental design (blocked vs. mixed trial design based on stimulus length) did not introduce any unexpected artifacts other than slightly faster RTs.

As for the measures of spread and shape, they are identical to those of Experiments 1, 2, and 4 of Chapter 2. They are all mostly flat across string length (except for skewness of “Different” decisions which seem to have a slowly descending trend) and provide corroborating evidence that the change in experimental design had no measurable influence on people’s performance in the task.

Race Model Inequalities

To begin the race model inequality analyses, one must first select the proper conditions: 1D2L and 2D2L for the first inequality, and 1D3L, 2D3L, and 3D3L for the second inequality. 4L stimuli cannot yet be properly assessed using the RMI as these theoretical bounds are limited to two or three targets (Diederich & Colonius, 2004; Engmann & Cousineau, 2013). Furthermore, it would be inappropriate to divide the 4L stimuli into two pairs of letters and use the 2L RMIs, as a fundamental assumption of this analysis is that stimuli are processed separately and with such a large stimulus, crosstalk is inevitable (Miller, 1982). Finally, coactive architectures are often associated with a ceiling effect where redundancy gains diminish rapidly as the number of relevant targets grows (see Harding, Goulet, Cousineau, & Chartier, 2017, for a simulation exemplifying this ceiling effect in a standard coactive model).

For the first inequality, I applied the classic Miller bound (Miller, 1978, 1982) as well as another RMI bound, the *Townsend bound* (Townsend & Nozawa, 1995). The Townsend bound is similar to the Miller bound in that it is interpreted in the same way: if the redundancy condition dominates the composite condition, one can assume the presence of coactivity. However, rather than constructing the bound with the addition of single-target conditions' CDF, the Townsend bound utilises the multiplication of single-target conditions’ survivor functions, SF (which are simply 1 - CDF). As noted above, one must only select the 1D2L conditions to be compared to the 2D2L conditions. The 1D2L condition must be separated by the position of the differing letter (difference on the first letter, difference on the second letter) and binned accordingly. Afterwards one must plot all three of these conditions as CDF in a single plot. The analyses are realized on a per-participant basis. The geometric averages of the distributions are presented in the following figure (Cousineau, Thivierge, Harding & Lacouture, 2015).

The results of this inequality analysis are presented in the left panel of Figure 3.2.

INSERT FIGURE 3.2 ABOUT HERE

As shown, the 2D2L condition dominates over both 1D2L conditions, as well as the composite bound over the range of RT spanning 380 to 440 ms. This inequality between the composite and redundancy bounds signifies that the extra mismatch within the string benefits the overall decision more than what could be expected from an unlimited-capacity parallel model. Thus, the fact that threshold variations are kept at a minimum (both with the fact that the conditions are blocked and that “Different” decisions require but a single attribute to mismatch) implies that coactivity exists within the system. Moreover, the more conservative Townsend bound show the same results.

As for the second inequality, the Miller bound cannot be generated when more than two source distributions are used. There are two RMI analyses that are applicable for three dimensions: 1) the *Diederich bound* (Diederich, 1992) which is an extension of the Miller Bound, and 2) the *Engmann bound* (Engmann, 2009), an extension of the Townsend bound which not only uses the addition of redundant targets, but also the gains between intermediate conditions (3D3L vs. 2D3L vs. 1D1L compared to just 3D3L vs. all three 1D1L condition). Engmann and Cousineau (2013) showed that this last bound is more conservative than the

Diederich bound and, therefore, provides the strongest test for evidence of coactivity. The results of this second analysis are presented in the right panel of Figure 3.2.

As shown, the redundancy condition's CDF breaches both composite bounds, even the more conservative Engmann bound in the range of RT from 380 ms to 420 ms. This translates to the presence of an additional redundancy gain with the addition of a third target dimension over and above any gain that could be present for pairs of dimensions. This confirms the claim of coactivation for "Different" decisions in the Same-Different task.

Unfortunately, the RMI cannot be applied to "Same" decisions because the conditions required to build the composite bound are unavailable. Indeed, the Same-Different task does not have any single-target conditions for "Same" (the 0D1L does not qualify because it does not have the same number of total dimensions as the 0D2L condition). Therefore, to see whether there are redundancy gains possible for this decision, or whether coactivity is possible for both decisions, one must approach the issue differently. This is the purpose of Experiment 2.

Experiment 2: Coactive characteristics in a task with wildcard-"Same" judgements

In this second experiment, I show that the presence of coactivity for "Same" decisions can be assessed in the Same-Different task. Typically, "Same" responses require all dimensions to be identical to be considered correct (there is no non-redundant "Same" trial) and therefore induce the thresholds to increase with the length of S_2. While the addition of matching letters within the string should lead to redundancy gains, it is unlikely that these gains are sufficient to oppose this inherent slow-down associated with processing additional letters. To correct this, I integrated Mordkoff and Yantis's (1991) framework of using wildcard stimuli to offer alternative conditions for "Same" responses. These stimuli, although visually different from S_1 as they are composed entirely of the letter "X", must be considered as "Same". As the wildcard stimulus is composed of a unique letter, threshold adjustments for S_2 is no longer necessary because only a single dimension needs to be detected (if the participant sees the first, second, third, or fourth letter first, it doesn't matter as they are all "X"s). Hereafter wildcard stimuli will be referred to as X-0D*l*L stimuli. "Different" trials were presented as before and had no wildcard alternative.

Methodology

Participants

Twenty additional undergraduate and graduate students aged between 17 and 30 years of age were recruited to participate in this task. Once more, all participants were informed of the experiment's procedure as well as the protocol and ethical rules of the University of Ottawa; all gave written and verbal consent to participate in this task. They were compensated $10.

Stimuli

The same twelve capital consonants were selected as stimuli for the standard task, whereas "X" was selected to compose the wildcard trials. All other details are identical to Experiment 1. The letters composing the standard stimuli were randomly selected on every trial and string length could vary as 1, 2, or 4 letters (more on the 3L missing condition later).

Procedure

Prior to the experimental phase, participants were informed of the general goals of the experiment without inducing any bias towards "Same" or "Different" responses. They were also diligently instructed that while they would have to compare identical or non-identical stimuli, if a stimulus composed of "X" is presented during S_2, they would have to press the key associated with "Same". These wildcard stimuli were always composed of only the letter "X" and were always the same length as S_1. This change in procedure was explicitly repeated to participants to ensure that the demand imposed by this new variant was understood.

All other details are identical to Experiment 1 above and the timeline of trial events were unchanged for this experiment.

Experimental Design

This experiment consisted of 576 total trials and the new experimental design of the control task is shown in Table 3.2.

INSERT TABLE 3.2 ABOUT HERE

As shown, there are many changes from Experiment 1 including the omission of all 3L stimuli. This was done to ensure the easiest counter-balance between “Different” and “Same” (including both standard and wildcard) judgements and ensures a close adaptation of Mordkoff and Yantis’s (1991) contingency controls. This omission is not worthy of further mention as it was shown in Experiment 1 that modifying the presentation order of the task has no effect on the overall pattern of results (it would be as if 3L stimuli were presented in the last block and not counted during the analysis). The second change is the addition of wildcard-“Same” stimuli. The final change, related to the first two, is the adjustment of proportions of each trial type to remove any first letter bias (i. e., by which a participant could only look at the first letter of S_2) and guess above chance whether the stimuli is “Same” or “Different”. This notion of wildcard stimuli to re-balance the experimental design was the original goal of Mordkoff and Yantis’s (1991) experimental design for a task of a similar nature (the redundant target detection task). With the wildcard trials, both they and I controlled specifically for inter-stimulus response contingency and non-target response contingency in order to control for decision bias and ensure a perfectly balanced experimental design – Mordkoff and Yantis noted that in some instances of the classic experimental design (Miller, 1982), one could look at a single dimension and guess correctly above chance whether the trial was a target, circumventing the task demand that requires participants to process all dimensions.

Results

Screening of the data

Data were gathered from 11,520 total trials (576 stimuli × 20 participants). One subject was excluded from analysis for having an overall accuracy of 72.3% (the average accuracy rate for a Same-Different task is typically 95%, as seen so far, with most participants being between 90% and 100% correct). Of the remaining 10,944 trials, there were 7 RTs slower than 2500 ms and they were removed in addition to 12 non-answers; none were below 200 ms. 10,424 valid observations were retained for RT analyses (501 errors were recorded for an error rate of 4.3%).

Mean response times

Mean RT and accuracy rates for each condition are presented in Figure 3.3; error bars denote the difference and correlation adjusted 95% confidence intervals for the means (CI; Cousineau, 2005; Morey, 2008) as recommended by Baguley (2012).

INSERT FIGURE 3.3 ABOUT HERE

As shown, the experimental design had little effect on the overall pattern of “Different” responses producing a faithful overall picture of performance (noted in Chapter 1); the omission of 3L stimuli had no effect on the positioning of the other conditions. “Different” decisions have an overall mean RT of 655 ms and an overall mean accuracy rate of 95.9%.

When it comes to “Same” judgements, they are slower than expected based on the experimental conditions seen in Chapter 2. Whereas standard-“Same” decisions are still faster than the 1D condition, attenuation of that result is evident with fast-“Same” responses being less prominent than in Experiment 1 and the control conditions of Experiments 1-4 in Chapter 2.

Although there are only two 1D conditions, if we were to linearly extrapolate towards 4L, we should observe that "Same" responses remain faster than "Different" responses as the mean RT of 1D4L should fall well within the error bars of 0D4L. This attenuation is likely due to the low overall proportion of standard-"Same" trials (standard-"Same" were less than 17% of total trials); manipulating proportion of trials was shown to attenuate the fast-same phenomenon (Ratcliff & Hacker, 1981). As for the wildcard-"Same" trials, they follow a much different trend: for RT, they significantly decrease as letter string length increases ($F(2, 3578) = 21.533$, $p < 0.000$; Tukey analyses show significant differences between 1L and 2L, $p < 0.000$, 1L and 4L, $p < 0.000$, and no significant differences between 2L and 4L, $p = 0.184$). Rather than holding steady at roughly 95.7% (as is expected from a standard Same-Different task), accuracy significantly improves as string length increases ($F(2, 3773) = 12.922$, $p < 0.000$; Tukey analyses show significant differences between 1L and 2L, $p < 0.000$, 1L and 4L, $p < 0.000$, and no significant differences between 2L and 4L, $p = 0.746$). Importantly, and regardless of decision rule, fast-"Same" is attenuated for the standard-"Same" trials and un-precedent for the wildcard-"Same". "Same" decisions have an overall mean RT of 616 ms (652 ms for standard-"Same" and 599 ms for wildcard-"Same") and an overall mean accuracy rate of 93.8% (91.6% for standard-"Same" and 94.8% for wildcard-"Same").

Distribution characteristics suggesting a coactive process

Because I want to focus on evidence of coactivity for "Same" judgements, I focus in this section on wildcard-"Same" stimuli. Because the participant only needs to detect a single "X" to accurately record an accurate "Same" decision, threshold limits can remain constant in all the wildcard trials. Yet, due to the aforementioned redundant nature of the stimuli (the more "X"s that are present, the more targets are present), there should be redundancy benefits. It is not possible to compute redundancy inequalities because there are no irrelevant attributes in the wildcard stimuli. Consequently, I examine other characteristics that signal coactive processing.

RT distributions of all three X-0D conditions are presented in the left column of Figure 3.4 – this distribution shows the pooled RT across all participants for those conditions for illustrative purposes only. Mean standard deviation and skewness measures of the wildcard stimuli are presented in the right panel of the same figure, computed on a per-participant basis. Error bars denote the correlation and difference adjusted 95% CI using the appropriate SE and CI estimator (Harding, et al., 2014).

INSERT FIGURE 3.4 ABOUT HERE

As shown, the wildcard-“Same” RT distributions become faster, more skewed, and compress towards the low-RT side of the plot as the number of redundant targets (“X”) increases. Mean RT for X-0D1L is 626 ms with a mean accuracy rate of 93.0%, mean RT for X-02L is 588 ms with a mean accuracy rate of 96.6%, and mean RT for X-0D4L is 576 ms with a mean accuracy rate of 97.2%. This decrease in mean RT and increase in accuracy are characteristic of redundancy gains (Miller, 1982; Grice & Gwynne, 1987; Engmann & Cousineau, 2013). If we look to higher statistical moments, the standard deviation measures decrease while skewness increases. Both trends are telltale of coactivity and furthers the notion that this architecture underlies decisions within the Same-Different task. Nevertheless, neither of these statistical trends are significant (*F*(2, 57) = 1.402, *p* = 0.254 for standard deviation and (*F*(2, 57) = 1.505, *p* = 0.231 for skewness).

For the wildcard trials, it is not possible to compute inequalities because it requires trials in which X are mixed with non-target letters. Such stimuli however do not trigger a "Same" response.

General discussion

In this chapter, I explored the possible presence of coactivity within the Same-Different task. Although this architecture had been alluded to before (Egeth, 1966; Miller & Bauer, 1981), there has been no formal exploration until now in the Same-Different task. Given the results of both experiments, it seems plausible, if not very likely, that coactivity underlies both “Same” and “Different” judgements. This is not only of note in terms of architectural diagnosis but especially considering the fact that it gives evidence towards a single process model. Single process models in the past (explored in Chapter 1) have all been unable to account for all aspects of the task. With the results given here and those of Chapter 2, we can now explore the parsimonious possibility of a coactive model for both “Same” and “Different” with priming benefits for repeated exposure (i. e., the visually identical “Same” trials).

Granted, future work should broaden the scope of the architecture analysis of the Same-Different task to include more sophisticated diagnostic methods. One such method is *Systems Factorial Technology* (SFT; Townsend & Nozawa, 1995), a methodology suite which is specialized in diagnosing cognitive architectures (including coactive) when using a 2 × 2 experimental design (a tutorial can be found in Harding et al., 2016). While the standard Same-Different task would need some modifications to qualify for an SFT analysis, it should not cause any issues as I have shown here, and in Chapter 2, that even large changes to the experimental design results in the expected trends from Chapter 1.

Publication intentions

Just like Chapter 2, this chapter can be submitted as a standalone manuscript with the addition of more theoretical framework (already written in the introduction and Chapter 1 of this thesis, only small adaptations would be required). I intend to publish this manuscript in a journal which is specialized in architecture analysis as well as cognitive paradigms (e. g. *Journal of Mathematical Psychology*). The potential impact of this manuscript is especially highlighted by the fact that the first experiment gives concrete empirical evidence towards a coactive architecture for “Different” responses while the second experiment supports it for “Same” responses. For the first time in five decades of research, I have shown that the same overall process can underlie both decisions, a feat deemed difficult or impossible to do (see Sternberg, 1998).

Chapter 4: A new predictive model of the Same-Different task

Accumulator models

Overview of accumulator models

Sequential sampling models have had a long and successful history in modelling human behavior (starting with Audley & Pike, 1965; Townsend & Ahsby, 1983; Luce, 1986). One of its branches, composed of accumulator models (also known as counting models or race models), focuses on the perceptual mechanisms behind decision-making. In these models, *evidence* (activations from target dimensions) are sampled from the stimuli and decisions are made when a critical amount of evidence has been collected (see Fortsmann, Ratcliff, & Wagenmakers, 2016, for an extensive review of accumulator models in cognitive psychology).

At their foundation, accumulator models follow the well-known kinematic relationship of time equals distance over speed and are comprised of three main parameters: the accumulation rate (the speed at which evidence is sampled from the stimuli), the accumulation's starting point (where the accumulation process begins), and the threshold (where the accumulation process ends). The relationships between – and the variability within – these three parameters define the decision-making time. For example, easy to detect stimuli should have a faster accumulation rate and/or accumulate for a shorter overall distance (represented by a higher starting point, lower threshold, or both), while harder to detect stimuli should have the opposite parameter settings.

Implementation of accumulator models

Although there are several implementations of accumulator models, the one presented herein is influenced by the Linear Ballistic Accumulator (LBA; Brown & Heathcote, 2008), a model that has shown to be effective in modelling human RT and accuracy (Heathcote & Hayes, 2012; Donkin, Brown, & Heathcote, 2011; amongst others). While the model presented here is not specifically an LBA (some error parameters have been omitted), it can make equally well-defined predictions regarding mean RT and returns the expected positively skewed RT distribution, a regularity that is observed in many cognitive paradigms including the Same-Different task (Taylor, 1976b; Krueger, 1978; Sternberg, 1998). In the LBA, and in my model, there is 1) a varying starting-point that is randomly sampled on every trial from a uniform distribution, 2) a fixed threshold that does not vary, and 3) a rate parameter that is randomly sampled from a normal distribution prior to every trial.

In this model's most rudimentary implementation, an accumulation rate is randomly sampled from a two-parameter normal distribution and a random starting point is sampled from a uniform distribution to include two sources of variability on every trial. For all trials, the threshold is fixed and does not vary. This implementation is shown in Equation 1:

$$\mathrm{RT} = \frac{b-a}{v} + \mathrm{T}_0 \tag{1}$$

where *RT* represents the total time taken to make the decision, *b* represents the threshold, *v* represents the accumulation rate (randomly sampled on each trial), and *a* is the starting point (also randomly sampled on each trial). T_0 is the residual time it takes to issue a decision, usually a parameter that is associated to the other factors that go into decision-making (such as motor and/or encoding time).

Accumulators can sample evidence from different stimulus dimensions. These sources of evidence can then be combined to form an accumulation network forming the *information processing architecture* (reviewed in Harding, et al., 2016; Townsend & Nozawa, 1995; and Chapter 1) and include the coactive architecture discussed in Chapter 3. Implementing coactivity in sampling models is simple: all relevant accumulators gather evidence towards a single threshold. The implementation equation for coactive processing, differing slightly from the standard implementation presented above, is shown in Equation 2:

$$\mathrm{RT} = \frac{\mathrm{b} - \sum a_i}{\sum v_i} + \mathrm{T}_0 \tag{2}$$

in which the rate from a single source (*a*) is replaced by the sum of all the information sources' rates (a_i). This is based on the assumption that all the sources operate in parallel, so that their individual rates sum. The major implication of this implementation is that redundant stimuli are decided upon much more rapidly than stimuli with a single source of evidence (as introduced in Chapter 3). Indeed, the overall processing rate is increased, being the sum of all dimensions' processing rates. Finally, the starting points are also summed prior to each trial to reflect that the advantage that one source may possess advantages the system as a whole. There can be any number of redundant accumulators. Here I will simulate up to four repetitions of the dimensions representing 4L stimuli.

A new Same-Different model

Assumptions

The model that I developed to explain Same-Different performance is based on five core assumptions and empirical constraints stemming from the experiments presented in Chapters 2 and 3 as well as the wisdom of models presented in Chapter 1. Together, these assumptions predict Same-Different RT and accuracy from the standard task, the tasks in which priming for "Same" decisions is attenuated, and the wildcard-"Same" trials.

1- Coactive accumulation of evidence for both "Same" and "Different" decisions. As discussed in Chapter 3, coactivity seems to be present for both decision modalities. For the purposes of this model, stimuli are broken down into individual letters for which there exists a match and mismatch detector. These detectors gather evidence and pool it towards either a “Same” accumulator or a “Different” accumulator. Therefore, this model can account for each of the possible “Same” and “Different” conditions by summing match/mismatch detectors towards their corresponding threshold. For example, if the stimulus is 1D4L there are 3 match detectors accumulating towards “Same” and 1 mismatch detector accumulating towards “Different”. Finally, inspired by the Noisy Operator (Krueger, 1978), when a stimulus is wholly identical or wholly different, the detectors of the opposite decision accumulate “noisy” evidence (the accumulation rates are sampled from distributions with a much smaller mean representing the fact that spurious activity can occur without direct perceptual stimulation).

2- <u>Hindered capacity for both decisions</u>. As discussed in Chapter 3, typical coactive architectures arrive to decisions quickly and the addition of extra detectors makes for accelerated results. While these trends are not unheard of in similar tasks (see Miller, 1982; Engmann & Cousineau, 2013), keeping the standard implementation would lead to “Different” conditions, with the same number of mismatches, to have the same RT regardless of how many letters compose the string. Therefore, it is necessary that this aspect is corrected all whilst keeping the advantage that coactive architectures have over strictly parallel models. This comes in the form of *hindered workload capacity*. This capacity implementation limits the overall processing capability when the number of letters increases in the string, even though the overall architecture remains in an overall super-capacity environment. In other words, while the efficiency of the detectors is hindered by the increased workload, the overall system still holds the benefits from being a coactive architecture. Therefore, it is not strictly a limited capacity model, nor is it operating at its full, super-capacity, potential, hence the “hindrance”. To implement this limited capacity, one must ensure that an increased workload affects the overall accumulation rate for both decisions, regardless of the number of matches or mismatches. For this implementation, I have opted to divide the final accumulation rate by the square root of the number of letters composing the string, a fixed parameter. For example, both 0D4L and 1D4L have detectors accumulating at the summed rate for each decision divided by $\sqrt{4}$. In the 0D condition there are four summed detectors for “Same” and four summed *noise* detectors for “Different”. In the 1D condition there are three summed detectors for “Same” and a single detector for “Different”. All four of these accumulators have a final rate divided by $\sqrt{4}$.

3- <u>Exhaustive thresholds for “Same” responses and self-terminating for “Different” responses</u>. As noted by Bamber (1969), “Different” decisions require but a single identification of a difference, and “Same” decisions are forcibly exhaustive. This means that regardless of what “Different” condition the model is simulating, be it noise from a 0D condition, or the actual summed rate of “Different” detectors accumulating evidence, the threshold is fixed. As for “Same” decisions, the threshold is multiplied by S_2’s total length. This mimics the fact that all matches must be detected before one can answer “Same”. While coactive architectures typically have a fixed threshold, this slight modification ensures that even with a coactive architecture, the accumulation must happen for a longer amount of time. This discrepancy in threshold follows the successes that analytical models have had (See Chapter 1) and allows the prediction of errors (which will be discussed in greater detail below).

4- Priming benefits for “Same” detectors. As noted by Proctor (1981), Bamber (1972), Bamber & Paine (1973), and the results of Chapter 2, “Same” decisions are especially fast when the physical nature of the stimuli is wholly identical. Physically mismatching stimuli do not benefit from this manipulation, which means that the speed bonus must forcibly come from an outside mechanism, such as priming. Therefore, in this model, priming is implemented as an increase in accumulation rate that is present only in “Same” detectors. Other alternatives could be a starting-point bias, or a reduced threshold. This parameter is optional and can be removed to predict attenuated-“Same” responses. No benefit is implemented for “Different” responses.

5- Flexibility in threshold settings. In the wildcard experiment presented in Experiment 2 of Chapter 3, S_2 was replaced with a learnt stimulus (“X” could replace all letters composing S_2 and had to be considered as a “Same” stimulus). From a modelling perspective, this would translate to setting the “Same” threshold to self-termination because only a single “X” needs to be detected.

The model is schematized in Figure 4.1 to help visualize how the assumptions interact in various scenarios. In this figure, the correct accumulator reached the decision. However, due to randomness, sometimes the accumulator for the opposing decision will breach its threshold first. In the top-right panel is a 0D1L stimulus, in the top-left panel is a 1D2L stimulus, in the bottom-left panel is a 1D4L stimulus, and in the bottom-right panel is a X-0D4L stimulus

---

INSERT FIGURE 4.1 ABOUT HERE

As shown, the more redundant accumulators there are, the faster the overall process will be. In cases where there was no information of a particular dimension (all-"Different" trials have no matching information and all-"Same" trials have no mismatching information), the accumulation rate is sampled from "noise", a distribution that has a much lower mean than that of the standard rate.

Modelling RT

To model RT with this model, I expand Equation 2 to implement the five assumptions above. This new equation is given in Eq. 3.

$$\mathrm{RT} = \frac{\mathrm{b} - \sum a_i}{\frac{\sum v_i}{\sqrt{\mathrm{n}_{\mathrm{Total}}}}} + \mathrm{T}_0 \tag{3}$$

Just like the other equations above, *RT* is the time taken for one of the decisions to take place; there are two of these accumulators racing against one another on every trial and the fastest time, the smallest of RT_{Same} and RT_{Diff}, is the decision that is made. Additionally, unlike the other equations, the accumulation rate is modulated by the square root of the total number of letters; this modulation owes to the capacity hindrance, assumption number 2. All other parameters are identical.

The model's parameters are presented in Table 4.1 for the standard-"Same" task, the attenuated-"Same" task, and the wildcard-"Same" task.

INSERT TABLE 4.1 ABOUT HERE

For simplicity's sake, all accumulation rates for all detectors are sampled from a single two-parameter normal distribution with mean μ and standard deviation σ. Similarly, all starting points for all detectors are selected from the same uniform distributions with bounds low and high. The threshold is fixed on every trial, varying with number of letters for "Same" decisions but fixed for "Different" decisions and wildcard-"Same" decisions (this refers to assumption 5 and stems from the results presented in Chapter 3). One may note that there is a priming benefit in the attenuated-“Same” responses, albeit its magnitude is 1/3 of the benefit found in the standard task. This stems from the results of Chapter 2 where it was shown that there could exist some benefits at other points in the processing hierarchy. Nevertheless, this does not affect the fact that “Same” responses are always as fast or faster than the slowest “Different” condition; in Appendix A, I show that even without any priming benefits, this model return “Same” responses that are always faster than or equal to the 1D slope. Furthermore, as noted in Table 4.1, the standard deviation of the accumulation rates’ normal distribution is adjusted based on the number of detectors. As the number of detectors grow, the variability of the accumulation rate increases at a rate directly proportional to the square root of the total number of letters. This stems from the fact that because the detectors’ accumulation rates are additive, and all detectors operate within the same system, their variability must be additive as well. Finally, an additional parameter, the motor time, is set to an arbitrary value – it is mostly a scaling parameter for minimum RT. As seen, only two parameters are adjusted to make predictions appropriate for each Same-Different task variant: only the priming benefit for “Same” responses and the threshold multiplier for “Same” responses are modified to fit experimental variants.

The model’s predictions, using the parameters noted above, are presented in Figure 4.2. In the top panels are the RT predictions and in the bottom panels are accuracy predictions. The left panels show the model’s predictions for the standard-“Same” task, the middle panels show the model’s predictions for the attenuated-“Same” task, and the right panels show the model’s predictions for the wildcard-“Same” task.

INSERT FIGURE 4.2 ABOUT HERE

As seen, all trends required for “Different” decisions are simulated without issue; all conditions match what is expected from Same-Different empirical data (for which a review was written in the introduction of this thesis) and the slopes for each “Different” conditions, in relation to string length fan-out as expected (Sternberg, 1998).With the addition of the priming benefit, “Same” responses resemble the fast-same phenomenon that is observed in the control conditions of Chapter 2 and the blocked experiment of Chapter 3. When the priming benefit is removed, “Same” responses slow down significantly and are no longer the fastest condition. Nevertheless, regardless of this slow-down, “Same” decisions remain faster, or as fast, than the slowest “Different” condition (1D). As for the wildcard-“Same” trials, RTs follow the same downward slope found in empirical research (as shown in Chapter 3). However, in these simulation results, it is to note that a small priming benefit for “Same” is also implemented. If it was not, “Same” decisions would perfectly overlap the all-“Different” conditions (both have identical simulation parameters). The reasoning behind this addition comes from the idea that participants are trained to seek the letter “X” within S_2 which creates an expectation for Xs, resulting in a form of “expectation-based” priming.

The results presented above, give a clear picture that this model can predict “Different” and “Same” RTs (both standard and wildcard variants) using a single underlying mechanism. Furthermore, this model can account for changes in the physical nature of the stimuli, a novel contribution to research showing that it is not necessary to implement a dual-process mechanism, or post-hoc decision processes. Most importantly, the model shows that the elusive fast-same phenomenon can be modelled using only priming benefits for “Same” decisions.

As of now, the model is set up with pre-selected parameters and is not meant for empirical data-fitting on a per participant basis. A procedure to find best-fitting parameters remains to be implemented but will utilize maximum likelihood. The first step in the process is to construct the probability density function (PDF) for my model. Nevertheless, as the LBA already has a well-established PDF, the extensions that I propose here should not pose any fundamental issue. While the parameters utilized here were chosen to best exhibit the model's predictions, the model itself is not bound to this particular selection; whenever there is enough variation in the accumulation rate and a large enough threshold, the predictions shown here are found. Other examples of parameter selection are given in Appendix A to show that scaling is the only component of the model that is affected when parameters are modified.

Modelling accuracy

Most Same-Different models only make predictions for RT, overlooking accuracy altogether. In fact, in Sternberg's (1998) book chapter dedicated to modelling the task, he states that the RTs are so ostensibly counter-intuitive from a modelling standpoint that no time would be dedicated to accuracies (although in his closing remarks, he dedicates a sub-section to accuracy rates of the task). The fundamental issue behind the inability to predict accuracies stem from the selected models. Serial accumulators would predict perfect accuracy for "Same" decisions (unless internal noise was added to the system), as it is the only possible decision once the stimuli has been exhaustively treated. Parallel architectures would offer no discernable difference between conditions as the addition of relevant dimensions only increases the probability that one of those dimensions will breach the threshold first, not that it offers complementary evidence towards a decision.

The Same-Different model presented here can account for accuracy rates using nothing but the innate functions of the model itself. In this model, as there are constantly two concurrent decision modules accumulating at once (one for "Same", and one for "Different") using the same fundamental mechanism, I can verify whether the correct accumulator breached the threshold first by observing whether a “Same” or “Different” condition was simulated. For errors in a “Different” condition, consider for example a 1D2L condition where a “Different” detector races against a “Same” detector to a response, with both rates being damped by a factor equal to the square root of two. If the “Same” accumulator happens to breach the threshold first, the overall decision is considered an error. Therefore, as can be inferred, more redundant “Same” detectors, even with the increased threshold, make for a higher probability of “Same” breaching the threshold first, especially considering that “Different” decisions' accumulation efficiency is also affected by the total number of letters composing the string. For errors in a “Same” condition (and those in an all-“Different” condition), as detectors of an opposite decision accumulate noise at a slower rate (akin to Krueger’s, 1978, Noisy Operator), a race still exists between detectors. This allows for the erroneous accumulator to still have a chance to breach the threshold first (albeit rarely), which translates to higher accuracy rates for those conditions (this is representative of the high empirical accuracy rates typically observed in those conditions).

The model’s predictions for accuracies in all three variants of the Same-Different task are presented in the bottom panels of Figure 4.2. As shown, “Different” accuracy rates fall where appropriate and differ very little from one experimental variant to another – as was shown in Chapter 2, this is what is expected. For “Same” accuracy rates, when there are priming benefits, the overall pattern matches what is expected of the empirical data (overviewed in the introduction and shown in Figure 1). This also goes for the accuracy rates of wildcard-“Same” trials. However, when “Same” decisions are simulating the attenuated trials task, accuracy rates substantially differ from what is seen, especially for the “Same” (notably the 0D2L, 0D3L, and 0D4L) conditions. As “Same” decisions are the only component that suffers a change in implementation (a lack of priming benefits make for a much slower response) “Different” decisions have no problem beating them in a race most of the time. Consequently, it is the only aspect that is affected. A possible solution to this shortcoming is to add a post-hoc selection parameter that temporarily affects both decision’s thresholds (raise “Different” and/or lower “Same”) if all “Same” detectors are fully activated – this would reflect Ratcliff and Hacker’s (1981) notion that “Same” decisions have an inherent bias and give an additional advantage for “Same” responses to reach the threshold first. While this extra parameter could potentially benefit the model fit, the model would lose parsimony and results would no longer be directly interpretable as a consequence of the race between accumulators.

Modelling higher statistical moments

Figure 4.3 shows simulated standard deviation and skewness.

INSERT FIGURE 4.3 ABOUT HERE

At first glance, the model’s predictions are mediocre: standard deviations for “Different” conditions increase as the string length increases while it remains flat for “Same” responses. As seen in Chapter 2, all conditions should have a flat standard deviation across all conditions (except the atypical phonetic condition of Experiment 3). However, when it comes to the measures of skewness, the model returns rather flat slopes as would be expected from empirical data. The only caveat for its full success would be the fact that there are meaningful simulated differences between conditions (the error bars are not close to overlapping whatsoever) whereas empirical data shows skewness measures that are rather constant across all conditions. Potentially, with fewer data per condition, error bars would grow enough to overlap and offer evidence towards non-significant differences between conditions as is the case with the empirical data. Another solution would be to run the model with variable parameters and graph the higher statistical moments with these results. Because parameters are usually set on a per-participant basis, the simulation results presented herein reflect but a "single" simulated participant fit to averaged results that lead to very little between-subject variability – with more variability within each condition in the results (stemming from many different “participants” doing the task), the more reflective the model will be of empirical results.

Conclusion

This model can account for mean RT and accuracy rates for both “Same” and “Different” responses and demystifies the mechanism behind the fast-same phenomenon using a single underlying mechanism. Out of 84 possible conditions (14 conditions × 3 variants × 2 RT/Accuracy; we can disregard higher moment statistics for now as the model is in evident need of tailoring to fit these conditions as well), the model accurately predicts the mean of 81 conditions. The only 3 that cannot be predicted adequately are the accuracy rates for the 0D2L, 0D3L, and 0D4L conditions of the attenuated-“Same” task. All this is accomplished with 9 parameters (7 free parameters for the standard task plus the adjusted priming benefits for the non-standard tasks as well as the threshold change for the wildcard-“Same” experiment). Whilst 7 (or 9) parameters may seem too much, other similar accumulator models have included more parameters to explain fewer conditions (Ratcliff, 1985, has 21 free parameters to explain 6 conditions; Gomez, Ratcliff, & Perea, 2007, has 25 free parameters to explain 8 conditions). It therefore provides strong evidence that the five core assumptions (four if we are to ignore the wildcard-“Same” variant) are important for modelling performance in the Same-Different task.

Publication intentions

My objective for this Chapter is to publish it along with the literature review given in Chapter 1 as its introduction. However, as this model stems from the results presented in Chapter 2 and 3, it would benefit from being submitted for publication after their acceptance. This is especially important in order for the assumptions presented herein to hold their weight. Finally, as a model of performance, the predictions are quite accurate considering that the parameters were hand-selected to best assess the empirical results. However, a parameter fitting procedure (maximum likelihood) using more sophisticated methods applied separately to each participants' data would be necessary. As parameters are usually fit on a participant basis, the manuscript would also benefit from having the same participants do all three task variants many times over to ensure a clean source of data with the same parameter values across tasks

Concluding remarks

Modelling the Same-Different task has been difficult for half a century due to past endeavors' limitations in explaining the fast-same phenomenon while being able to predict RTs and accuracies of both "Same" and "Different" decisions using a single underlying process.

In this thesis, I show new Same-Different results for the first time in over two decades and offer compelling evidence towards a new model of the Same-Different task. This model has been built using intuitions from past successes (reviewed in Chapter 1) as well as those stemming from results of experiments presented within this thesis. Notably, in Chapter 2, I show that the seemingly impossible problem of modelling the fast-same phenomenon is plausibly solved as the product of priming, an insight from Proctor (1981) and Nickerson (1978). In these tasks, through a systematic approach, I show that the mental representation of S_1 needs to be constructed using physically identical, visual, stimuli. As soon as this dimension is altered, the fast-same phenomenon disappears yet remains faster than the slowest "Different" condition. Furthermore, in Chapter 3, using the results of the RMI and noting the characteristics of RT distributions, I show compelling evidence that a coactive architecture can underlie the decision mechanism for both "Same" and "Different" decisions. Considering that the RMI is both a confirmatory tool for coactivity as well as refutation tool for parallel processing, the results of Experiment 3.1 show that all future models should be able to explain redundancy gains through either coactivity or other capacity-altered models. This finding also refutes past models in which a serial or a parallel architecture was suggested, further showing the impact of this thesis. Finally, the model presented here is good at predicting the RT of all three same-different task variants presented herein and is capable of predicting the accuracies for both wildcard-"Same" variants and standard-"Same" variants. In its current iteration, the model cannot predict the accuracy for the 0D3L nor the 0D4L conditions of the attenuated-“Same” task nor can it attest for the statistics of higher moments. Nevertheless, by fitting it with the hundreds of thousands of data amassed in our lab, I believe that these mediocre predictions are temporary.

While this thesis offers advances with promise of closure to some aspects of the task, there still is much ground to cover. Most importantly, confirmatory analyses of architecture should take place including the SFT method (Townsend & Nozawa, 1995), a analysis tool specialized at detecting the underlying architecture of a system. Additionally, a Same-Different task that affects priming and not the mental representation hierarchy should also be conducted in order to best assess the type and mechanism of priming in fast-“Same” response. Nevertheless, above all else, the model’s parameters should be fit on a participant basis to best assess the model’s ability as a predictor of Same-Different performance.

References

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. Doi: 10.1111/j.2044-8317.1965.tb00342.x

Baguley, T. (2012). Calculating and graphing within subject confidence intervals for ANOVA. *Behavior Research Methods, 44,* 158-175. Doi: 10.3758/s13428-011-0123-7

Bamber, D. (1969). Reaction times and error rates for “Same” and “Different” judgments of multidimensional stimuli. *Perception & Psychophysics, 6*(3), 169– 174. Doi: 10.3758/BF03210087

Bamber, D. (1972). Reaction times and error rates for judging nominal identity of letter strings. *Perception & Psychophysics, 12*(4), 321–326. Doi: 10.3758/BF03207214

Bamber, D. & Paine, S. (1973). Information retrieval processes in “Same”-“Different” judgments of letter strings. In S. Kornblum (Ed.), *Attention and performance IV*. New York, NY: Academic Press.

Bamber, D., Herder, J., & Tidd, K. (1975). Reaction times in a task analogous to “Same”-“Different” judgment. *Perception & Psychophysics*, *18*(5), 321-327. Doi: 10.3758/BF03211207

Beller, H. K. (1970). Parallel and serial stages in matching. *Journal of Experimental Psychology, 84*(2), 213-219. Doi: 10.1037/h0029065

Ben-David, B. M., & Algom, D. (2009). Species of redundancy in visual target detection. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 958–976. Doi: 10.1037/a0014511

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology, 13*, 90-102. Doi: 10.1080/17470216108416478

Bindra, D., Donderi, D. C. & Nishisato, S. (1968). Decision latencies of “Same” and “Different” judgements. *Perception & Psychophysics) 3*(2b), 121-136. Doi: 10.3758/BF03212780

Bindra, D., Williams, J. A., & Wise, J. S. (1965). Judgments of sameness and difference experiments on decision time. *Science, 150*(3703), 1625–1627. Doi: 10.1126/science.150.3704.1699-a

Heathcote, A., Brown, S., Wagenmakers, E. J., & Eidels, A. (2010). Distribution-free tests of stochastic dominance for small samples. *Journal of Mathematical Psychology*, *54*(5), 454–463. Doi: 10.1016/j.jmp.2010.06.005.

Castro, L., Wasserman, E. A., & Young, M. E. (2012). Variations on variability: effects of display composition on same–different discrimination in pigeons. *Learning & Behavior, 40*(4), 416–426. Doi: 10.3758/s13420-011-0063-1

Cousineau, D. (2005). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology, 1*(1), 42-45. Doi: 10.20982/tqmp.01.1.p042

Cousineau, D., Goodman, V. W., & Shiffrin, R. M. (2002). Extending statistics of extremes to distributions varying in position and scale, and the implications for race models. *Journal of Mathematical Psychology, 46*, 431-454. Doi:10.1006/jmps.2001.1399

Cousineau, D., Thivierge, J.-P., Harding, B., Lacouture, Y. (2016). Constructing a group distribution from individual distributions. *Canadian Journal of Experimental Psychology, 70*(*3*), 253-277. Doi: 10.1037/cep0000069

Decker, L. R. (1974). The effect of method of presentation, set, and stimulus dimensions on “Same” - “Different” reaction times. *Perception & Psychophysics, 16*(2), 271–275. Doi: 10.3758/BF03203941

Derks, P. L. (1972). Visual recognition of similarity and identity. *Journal of Experimental Psychology, 95*,237-239. Doi: 10.1037/h0033282

Diederich, A. (1992). Probability inequalities for testing separate activation models of divided attention. *Perception and Psychophysics, 52*, 714-716. Doi: 10.3758/BF03211708

Diederich, A., & Colonius, H. (2004). Bimodal and trimodal multisensory enhancement of reaction time: Effects of stimulus onset and intensity. *Perception and Psychophysics, 66*, 1388-1404. Doi: 10.3758/BF03195006

Donderi, D. C., & Zelnicker, D. (1969). Parallel processing in visual same-different decisions. *Perception & Psychophysics, 5*(4), 197-200. Doi: 10.3758/BF03210537

Donkin, C., Brown, S., Heathcote, A. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review, 18*(1). 61-69. Doi: 10.3758/s13423-010-0022-4

Dyer, F. N. (1973). Same and different judgments for word-color pairs with “irrelevant” words or colors: evidence for word-code comparisons. *Journal of Experimental Psychology, 98*(1), 102–108. Doi: 10.1037/h0034278

Egeth, H. E. (1966). Parallel versus serial processes in multidimensional stimulus discrimination. *Perception & Psychophysics, 1*(4), 245–252. Doi: 10.3758/BF03207389

Engmann, S. (2009). *Redundancy gain: manifestations, causes, and predictions*. Université de Montréal: unpublished doctoral dissertation.

Engmann, S., & Cousineau, D. (2013). Triple redundant signals effect in the visual modality. *Universitas Psychologica, 12*(5), 1473–1488. Doi: 10.11144/Javeriana.UPSY12-5.trse

Eriksen, C. W., O'Hara, W. P., & Eriksen, B. (1982). Response competition effects in same-different judgments. *Perception & Psychophysics, 32*(3), 261–270. Doi: 10.3758/BF03206230

Eviatar, Z., Zaidel, E., & Wickens T. (1994). Nominal and physical decision criteria in same-different judgments. *Perception & Psychophysics, 56*(1), 62-72. Doi: 10.3758/BF03211691

Farell, B. (1977). *Encoding and comparisons in "Same"-"Different" judgments*. McGill University: unpublished doctoral dissertation.

Farell, B. (1984). Attention in the processing of complex visual displays: detecting features and their combinations. Journal of Experimental Psychology. *Human Perception and Performance, 10*(1), 40–64. Doi: 10.10370096- 1523.10.1.40

Farell, B. (1985). "Same"-"Different" judgements: A review of current controversies in perceptual comparisons. *Psychological Bulletin, 98*(3), 419–456. Doi: 10.1037/0033-2909.98.3.419

Feintuch, U., & Cohen, A. (2002). Visual attention and co-activation of response decisions for features from different dimensions. *Psychological Science, 13*, 361-369. Doi: 10.1111/1467-9280.00465

Fortsmann, B. U., Ratcliff, R., Wagenmakers E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review in Psychology, 67,* 641-666. Doi: 10.1146/annurev-psych-122414-033645

Gondan, M., Götze, C., & Greenlee, M. (2010). Redundancy gains in simple responses and go/no-go tasks. *Attention, Perception, & Psychophysics, 72*(6), 1692–1709. Doi: 10.3758/APP

Gottlob, L. R. (2007). Aging and capacity in the same-different judgment. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 14*(1), 55–69. Doi: 10.1080/138255890969528

Gomez, P., Ratcliff, R., Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*(3). 389-413. Doi: 10.1037/0096-3445.136.3.389

Grice, R. G., Gwynne, J. W. (1987). Dependence of target redundancy effects on noise conditions and number of targets. *Perception & Psychophysics, 42*(1), 29-36. Doi: 10.3758/BF03211511

Harding, B., Goulet, M.A., Cousineau, D., & Chartier, S. (2017). Are Recurrent Associative Memories good models of decision making? Modelling discrimination decisions from different perspectives. In B. Howell (Eds.), *International Joint Conference on Neural Networks* (pp. 2621-2628), Anchorage, AK: Institute of Electrical and Electronics Engineers. Doi: 10.1109/IJCNN.2017.7966176

Harding, B., Goulet, M. A., Jolin, S., Tremblay, C. Villeneuve, S.-P., & Durand, G. (2016). Systems factorial technology explained to humans. *The Quantitative Methods for Psychology, 12*(1), 39-59. Doi: 10.20982/tqmp.12.1.p039

Harding, B., Tremblay, C., Cousineau, D. (2014). Standard errors: A review and evaluation of standard error estimators using Monte Carlo simulations. *The Quantitative Methods for Psychology, 10*(2), 107-123. Doi: 10.20982/tqmp.10.2.p107

Hawkins, H. (1969). Parallel processing in complex visual discrimination. *Perception & Psychophysics, 5*, 56-64. Doi: 10.3758/BF03210482

Hawkins, H. L., & Shigley, R. H. (1972). Irrelevant information and processing mode in speeded discrimination. *Journal of Experimental Psychology, 96*(2), 389–395. Doi: 10.1037/h0033642

Heathcote, A., & Hayes, B. (2012). Diffusion versus linear ballistic accumulation: Different models for response time with different conclusions about psychological mechanisms? *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 66*(2), 125–136. Doi: 10.1037/a0028189.

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*(150), 1-19. Doi: 10.3389/fnins.2014.00150

Henmon V. A. C. (1911). The relation of the time of a judgment to its accuracy. *Psychological Review Review, 18*, 186–201. Doi: 10.1037/h0074579

Howell, P. & Stockdale, J. (1975). Memory and display search in binary classification reaction time. *Attention Perception & Psychophysics, 18*(6), 379-388. Doi: 10.3758/BF03204109

Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology: General, 137*(2), 324–347. Doi: 10.1037/0096-3445.137.2.324

Huber, D. E., & O'Reilly, R. C. (2003). Persistence and accommodation in short-term priming and other perceptual paradigms: temporal segregation through synaptic depression. *Cognitive Science, 27*, 403-430. Doi: 10.1016/S0364-0213(03)00012-0

Hock, H. S. (1973). The effects of stimulus structure and familiarity on same-different comparison. *Perception & Psychophysics, 14*(3), 413–420. Doi: 10.3758/BF03211176

Holmgren, J. E., Juola, J. F., & Atkinson, R. C. (1974). Response latency in visual search with redundancy in the visual display. *Perception and Psychophysics*, *16*(1), 123–128. Doi: 10.3758/BF03203264

Irwin, R. J., Hautus, M. J., & Francis, M. a. (2001). Indices of response bias in the same-different experiment. *Perception & Psychophysics, 63*(6), 1091–1100. Doi: 10.3758/BF03194527

Jacob, J., Breitmeyer, B. G., & Trevino, M. (2013). Tracking the first two seconds: three stages of visual information processing? *Psychonomic Bulletin & Review, 20*(6), 1114-1119. Doi: 10.3758/s13423-013-0482-4

Jeffreys, H. (1931). *Scientifc Inference*. London, UK: Cambridge University Press.

Kinchla R. A., Collyer C. E. (1974). Detecting a target letter in briefly presented arrays: A confidence rating analysis in terms of a weighted additive effects model. *Perception & Psychophysics. 16*, 117-122. Doi: 10.3758/BF03203263

Kornblum, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review, 76*(2), 113-131. Doi: 10.1037/h0027245

Krueger, L. E. (1973). Effect of stimulus frequency on speed of “Same”-”Different” judgments. In S. Kornblum (Ed.), *Attention and performance V*. New York, NY: Academic Press.

Krueger, L. E. (1978). A Theory of Perceptual Matching. *Psychological Review, 85*(4), 278–304. Doi: 10.1037/0033-295X.85.4.278

Krueger, L. E. (1979). A model of unidimensional perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance, 5*(2), 277-288. Doi: 10.1037/0096-1523.5.2.277

Krueger, L. E. (1987). Effect of backward masking on same-different judgments. *Perception & Psychophysics, 41*(4), 375–81. Doi: 10.3758/BF03208239

Krueger, L. E., & Shapiro, R. G. (1981). A Reformulation of Proctor's Unified Theory for Matching-Task Phenomena. *Psychological Review, 88*(6), 573–581. Doi: 10.1037/0033-295X.88.6.573

Link, S. W., & Tindall, A. D. (1971). Speed and accuracy in comparative judgments of line length. *Perception and Psychophysics, 9*(3), 284–288. Doi: 10.3758/BF03212649

Luce, R. D. (1986). *Response Times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.

Lupker, S. J., Nakayama, M., & Perea, M. (2015). Is there phonologically based priming in the same−different task? Evidence from Japanese−English bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 41*(5), 1281-1299. Doi: 10.1037/xhp0000087

Megreya, A.M. & Burton, A.M. (2006). Unfamiliar faces are not faces: Evidence from a matching task. *Memory & Cognition*, *34*(4) 865-876. Doi: 10.3758/BF03193433

Miller, J. (1978). Multidimensional same-different judgments: Evidence against independent comparisons of dimensions. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(3), 411–422. Doi: 10.1037/0096- 1523.4.3.411

Miller, J. (1982). Divided attention: evidence for coactivation with redundant signals. *Cognitive Psychology, 14*(2), 247–279. Doi: 10.1016/0010-0285(82)90010-X

Miller, J. & Bauer, D. W. (1981). Irrelevant differences in the "Same"-"Different" task. *Journal of Experimental Psychology: Human Perception and Performa*nce, *7*(1), 196-207. Doi: 10.1037/0096-1523.7.1.196.

Morey, R. D. (2008). Confidence intervals from normalised data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology, 4*(2), 61-64. Doi: 10.20982/tqmp.04.2.p061

Mordkoff, J. T., & Yantis, S. (1991). An interactive race model of divided attention. *Journal of Experimental Psychology: Human Perception and Performance, 17*(2), 520–538. Doi: 10.1037/0096-1523.17.2.520

Mordkoff, J. T., & Yantis, S. (1993). Dividing attention between color and shape: evidence of coactivation. *Perception & Psychophysics, 53*(4), 357–366. Doi: 10.3758/BF03206778

Nickerson, R. S. (1965). Response times for “Same” “Different” judgements. *Perceptual & Motor Skills, 20,* 15–18. Doi: 10.2466/pms.1965.20.1.15

Nickerson, R. S. (1967a). Categorization time with categories defined by disjunctions and conjunctions of stimulus attributes. *Journal of Experimental Psychology, 73*(2), 211–219. Doi: 10.1037/h0021270

Nickerson, R. S. (1967b). “Same”-“Different” response times with multi-attribute stimulus differences. *Perceptual and Motor Skills, 24,* 543–554. Doi: 10.2466/pms.1967.24.2.543

Nickerson, R. S. (1968). Note on “Same”-“Different” Response Times. *Perceptual & Motor Skills, 27,* 565–566. Doi: 10.3758/BF03330991

Nickerson, R. S. (1969). “Same”-’different’ response times: A model and a preliminary test. In W. G. Koster (Ed.*), Attention and Performance II* (pp. 257–275). Amsterdam, NL: North-Holland Publishing Company.

Nickerson, R.S. (1978). On the time it takes to tell things apart. In Requin, J. (Ed) *Attention and Performance VII* (pp. 77-88). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nickerson, R. S., & Pew, R. W. (1973). Visual pattern matching: An investigation of some effects of decision task, auditory codability, and spatial correspondence. *Journal of Experimental Psychology, 98*(1), 36–43. Doi: 10.1037/h0034300

Pachella, R. G., & Miller, J. O. (1976). Stimulus probability and same-different classification. *Perception & Psychophysics, 19*(1), 29–34. Doi: 10.3758/BF03199382

Pan, K., & Eriksen, C. W. (1993). Attentional distribution in the visual field during same-different judgments as assessed by response competition. *Perception & Psychophysics, 53*(2), 134–144. Doi: 10.3758/BF03211723

Petrov, A. a. (2009). Symmetry-based methodology for decision-rule identification in same--different experiments. *Psychonomic Bulletin & Review, 16*(6), 1011–1025. Doi: 10.3758/PBR.16.6.1011

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review, 78*(5), 391-408. Doi: 10.1037/h0031333

Posner, M. I. & Mitchell, R. F. (1967). Chronometric analysis of classification. *Psychological Review, 74,* 392–409. Doi: 10.1037/h0024913

Proctor, R. W. (1981). A unified theory for matching-task phenomena. *Psychological Review, 88*(4), 291–326. Doi: 10.1037//0033-295X.88.4.291

Proctor, R. W. (1986). Response bias, criteria settings, and the fast-same phenomenon: A reply to Ratcliff. *Psychological Review, 93*(4), 473-477. Doi: 10.1037/0033-295X.93.4.473

Proctor, R. W., & Healy, A. F. (1985). Order-relevant and order-irrelevant decision rules in multiletter matching. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 11*(3), 519–537. Doi: 10.1037/0278-7393.11.3.519

Proctor, R. W. & Rao, K. V. (1982). On the “misguided” use of reaction-time differences: A discussion of Ratcliff and Hacker (1981). *Perception & Psychophysics, 31*(6), 601-602. Doi: 10.3758/BF03204200

Proctor, R. W. & Rao, K. V. (1983). Evidence that the same-different disparity in letter matching is not attributable to response bias. *Perception and Psychophysics, 34*(1), 72-76. Doi: 10.3758/BF03205898

Proctor, R. W., Rao, K. V., & Hurst, P. W. (1984). An examination of response bias in multiletter matching. *Perception & Psychophysics, 35*(5). 464-476. Doi: 10.3758/BF03203923

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. Doi: 10.1037/0033-295X.85.2.59

Ratcliff, R. (1981). A theory of order relations in perceptual matching. *Psychological Review, 88*(6), 552-572. Doi: 10.1037/0033-295X.88.6.552

Ratcliff, R. (1985). Theoretical interpretations of the speed and accuracy of positive and negative responses. *Psychological Review, 92*(2), 212-225. Doi: 10.1037/0033-295X.92.2.212

Ratcliff, R. & Hacker, M. J. (1981). Speed and accuracy of same and different responses in perceptual matching. *Perception & Psychophysics, 30*(3), 303–307. Doi: 10.3758/BF03214286

Ratcliff, R. & Hacker, M. J. (1983). On the misguided use of reaction-time differences: A reply to Proctor & Rao (1982). *Perception & Psychophysics, 31*(6), 603-604. Doi: 10.3758/BF03204201

Ratcliff, R., McKoon, G., & Verwoerd, M. (1989). A bias interpretation of facilitation in perceptual identification. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 15*(3), 378–387. Doi: 10.1037/0278-7393.15.3.378

Schwarz, W. (1989). A new model to explain the redundant-signals effect. *Perception & Psychophysics, 46*(5), 498-500. Doi: 10.3758/BF03210867

Sekuler, R. W., & Abrams, M. (1968). Visual sameness: A choice time analysis of pattern recognition processes. *Journal of Experimental Psychology, 77*(2), 232–238.

Silverman, W. P. & Goldberg, S. L. (1975). Further confirmation of same vs. different processing differences. *Perception & Psychophysics*, *17*(2), 189-193. Doi: 10.3758/BF03203884

Snodgrass, J. G. (1972). Matching patterns vs matching digits: The effect of memory dependence and complexity on "Same"-"Different" reaction times. *Perception & Psychophysics, 11*(5), 341–349. Doi: 10.3758/BF03206264

Snodgrass, J. G., & Townsend, J. T. (1980). Comparing parallel and serial models: Theory and implementation. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 330–354. Doi: 10.1037/0096-1523.6.2.330

St. James, J. D. & Eriksen, C. W. (1991). Response competition produces a "fast same effect" in same-different judgments. In Lockhead, G. R.; Pomerantz, J. R. (Eds). *The perception of structure: Essays in honor of Wendell R. Garner*, (pp. 157-168). Washington, DC: American Psychological Association xv. Doi: 10.1037/10101-009

Sternberg, S. (1998). Inferring mental operations from reaction-time data: how we compare objects. In D. N. Osherson, D. Scarborough, & S. Sternberg (Eds.), *An invitation to* *cognitive science, volume 4: methods, models, and conceptual issues* (pp. 365–454). Cambridge, MA: MIT Press

Taylor, D. A. (1976a). Effect of identity in the multiletter matching task. *Journal of Experimental Psychology: Human Perception Performance; 2*(3), 417-28. Doi: 10.1037/0096-1523.2.3.417

Taylor, D. A. (1976b). Holistic and analytic processes in the comparison of letters. *Perception & Psychophysics, 20*(3), 187-190. Doi: 10.3758/BF03198599

Taylor, D. A. (1977). Time Course of Context Effects. *Journal of Experimental Psychology-General, 106*(4), 404–426. Doi: 10.1037/0096-3445.106.4.404

Taylor, R. L. (1969). Comparison of short-term memory and visual sensory analysis as sources of information. *Journal of Experimental Psychology, 81*(3), 515–522. Doi: 10.1037/h0027894

Townsend, J. T. & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes*. Cambridge, MA: Cambridge University Press

Townsend, J. T. & Nozawa, G. (1995). Spatio-temporal properties of elementary perception an investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39*(4), 321–359. Doi: 10.1006 / jmps.1995.1033

Treisman A. M. (1985). Pre-attentive processing in vision. *Computer Vision, Graphics, and Image Processing. 31*(2) 156–177. Doi: 10.1016/S0734-189X(85)80004-9

Tversky, B. (1969). Pictorial and verbal encoding in a short-term memory task. *Perception & Psychophysics, 6*(4), 225–233. Doi: 10.3758/BF03207022

Van der Heijden, A. H. C. (1975). Some evidence for a limited capacity parallel self terminating process in simple visual search tasks. *Acta Psychologica, 39*(1), 21-41. Doi: 10.1016/0001-6918(75)90019-0

Van der Heijden, A. H. C., La Heij, W., & Boer, J. P. A. (1983). Parallel processing of redundant Targets in simple visual search tasks. *Psychological Research, 45*, 235-254. Doi: 10.1007/BF00308704

Van der Heijden, A. H. C., Schreuder, R., Maris, L., & Neerincx, M. (1984). Some evidence for correlated separate activation in a simple letter-detection task. *Perception and Psychophysics, 36*(6), 577–585. Doi: 10.3758/BF03207519

Van Opstal, F. & Verguts, T. (2011). The origins of the numerical distance effect: The same–different task. *Journal of Cognitive Psychology, 23*, 112–20. Doi: 10.1080/20445911.2011.466796

Well, A. D., & Green, J. (1972). Effects of color differences in a letter matching task. *Psychonomic Science, 29*(2), 109–110. Doi: 10.3758/BF03336585

Well, A. D., Pollatsek, A., & Schindler, R. M. (1975). Facilitation of both same and different judgments of letter strings by familiarity of letter sequence. *Perception & Psychophysics, 17*(5), 511–520. Doi: 10.3758/BF03203303

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202–238. Doi: 10.3758/BF03200774.

Appendix A: Controlling the fast-same phenomenon

In this appendix, I explore other parameter settings to show that the model's successes as a predictor of Same-Different performance is not bound to the selected parameters of Chapter 4.

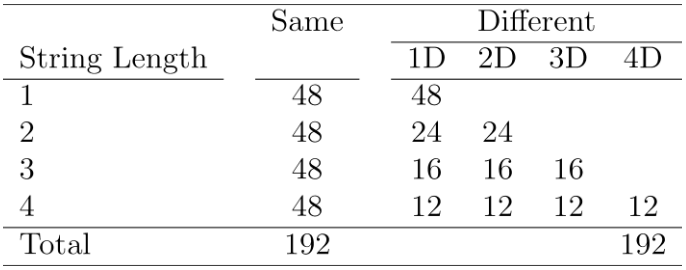
Because the relationship between the parameters are more important than the values per se, all simulations hereafter will target, and modify, a single (or group when the parameters are different for both decisions) of the accumulator's parameters. In the first simulation, I will remove all priming benefits from the accumulator to show how "Same" decisions perform in the complete absence of priming benefits (there was still a little priming benefit implemented for the attenuated-"Same" predictions of Chapter 4). Afterwards in the second, third, and fourth simulations I will adjust the accumulation rate, the rate's variability, and the threshold value respectively. During these simulations, only the targeted parameter will be modified, and all other details will be identical to those of the first simulation of this appendix (including the full cancelation of all priming benefits). The parameter settings for this appendix are presented in Table A.1, and the results are presented in Figure A.1.

INSERT TABLE A.1 AND FIGURE A.1 ABOUT HERE

As is seen, "Same" decisions are always as fast as or faster than those of "Different", a staple of this task, regardless of the parameter values. Furthermore, whether one was to affect accumulation rate, the variability of the accumulation, or the threshold, predictions resemble the expected trend; only the scale is affected. However, in the case where rate variability was increased, we can see that "Same" decisions are flat when they should be positively sloped. This is of no real issue as the variability is quite high leading to the possibility of very high accumulation rates or rates near 0 (1 SD from 40 is 20 or 60, and 2 SD is 0 and 80).

Tables and Figures

Table 2.1. Number of trials based on string length for both “Same” and “Different” trials. With “Different” trials there could be 1 difference (1D) or more depending on the string’s length. This design has been used for all experimental variants presented here except where specified. This table represents half of the total number of trials, equivalent to one of the experimental manipulations. Both experimental manipulations were presented to the participant within one test session. All conditions are presented in a random order and both “Same” and “Different” responses for all letter string lengths occur with the same frequency. For “Different” strings, the number of trials with a given number of differences are equal.



	Same	Different			
String Length		1D	2D	3D	4D
1	48	48			
2	48	24	24		
3	48	16	16	16	
4	48	12	12	12	12
Total	192				192

Table 2.2. Summary of mean slope (in ms/L) and intercepts (in ms) for both experimental manipulations of all four experiments for all participants. Measures were only recorded for the "Same" and 1D conditions. In parenthesis are the mean standard error of the individual estimates. In the leftmost column are the experiment names as well as the targeted experimental manipulation

		Condition 1		Condition 2	
Experiment Name	Measure	Same	1D	Same	1D
Case manipulation	Slope	16.61 (0.30)	37.94 (0.63)	30.09 (0.28)	46.24 (0.90)
(Match vs. Mismatch)	Intercept	463.63 (0.82)	493.40 (1.33)	461.20 (0.76)	489.75 (1.9)
Font manipulation	Slope	25.11 (0.31)	50.46 (0.90)	36.39 (0.30)	42.55 (0.76)
(Match vs. Mismatch)	Intercept	481.66 (0.85)	527.82 (1.85)	473.75 (0.82)	532.32 (1.62)
Stimuli Presentation	Slope	19.81 (0.35)	19.63 (0.73)	145.77 (0.68)	142.17 (1.33)
(Visual vs. Auditory)	Intercept	482.96 (0.95)	594.40 (1.52)	370.19 (1.85)	490.34 (2.82)
LTM association	Slope	21.80 (0.79)	29.01 (0.43)	34.17 (0.81)	27.50 (0.67)
(Present vs. Absent)	Intercept	499.81 (1.71)	573.52 (1.17)	590.92 (1.74)	663.46 (1.82)

Table 3.1. Experimental design for Experiment 1, the Same-Different task with blocked conditions. Number of trials based on string length for both “Same” and “Different” trials. With “Different” trials there could be 1 difference (1D) or more depending on the string’s length. String length was blocked for each quarter session (192 trials). However, each condition within each *l* block are presented in a random order and both “Same” and “Different” responses occur with the same frequency. For “Different” strings, the number of trials with a given amount of differences are equal.

	Same	Different			
String Length		1D	2D	3D	4D
1	96	96			
2	96	48	48		
3	96	32	32	32	
4	96	24	24	24	24
Total	384	384			

Table 3.2. Experimental design for Experiment 2, the Same-Different task with wildcard-"Same" decisions. As is seen, some conditions have been omitted including 3L stimuli altogether (see Table 1 for the full, classic, design) in order to introduce a second response type for "Same" stimuli that needs no threshold variation (the wildcard trials were inspired by Mordkoff & Yantis's, 1991, experimental design of a similar nature). In the wildcard-"Same" trials, participants had to record a "Same" decision as soon as a string composed entirely of "X" (with an identical length to S_1) was presented on screen during S_2.

	Same		Different			
String Length	Standard	Wildcard	1D	2D	3D	4D
1	32	64	96			
2	32	64	64	32		
3						
4	32	64		64		32
Total	288		288			

Table 4.1. Implementation parameters for the Same-Different model presented in this Chapter. Values for all three variants of the task are presented in the rightmost columns (the name of the columns refers to how “Same” decisions were affected during the experiment. In the left column are the values for the standard task with no changes to the experimental design, in the center column are the parameter values when priming for “Same” responses was hindered, and in the right column are the parameter values when “Same” decisions were replaced with wildcard stimuli. The number of matching letters within the string is reflected by n_{Same} and the number of mismatching letters is reflected by n_{Diff}; the total number of letters composing the string is reflected by n_{Total}

(CONTINUED ON NEXT PAGE)

List of Parameters

		Values for Various Experiments		
Parameters	Descriptions	Standard-”Same”	Attenuated-”Same”	Wildcard-”Same”
A	Starting point upper bound	500	500	500
b	Threshold for ”Different” responses	10000	10000	10000
	Threshold for ”Same” responses	$10000\ \times \mathrm{n}_{\mathrm{Total}}$	$10000\ \times \mathrm{n}_{\mathrm{Total}}$	10000
ν	Mean accumulation rate of all evidence detectors	40	40	40
	Mean accumulation rate of all noise detectors	$\sqrt{40}$	$\sqrt{40}$	$\sqrt{40}$
s	Standard deviation for ”Same” ν	$10\ \times \sqrt{\mathrm{n}_{\mathrm{Same}}}$	$10\ \times \sqrt{\mathrm{n}_{\mathrm{Same}}}$	$10\ \times \sqrt{\mathrm{n}_{\mathrm{Same}}}$
	Standard deviation for ”Different” ν	$10\ \times \sqrt{\mathrm{n}_{\mathrm{Diff}}}$	$10\ \times \sqrt{\mathrm{n}_{\mathrm{Diff}}}$	$10\ \times \sqrt{\mathrm{n}_{\mathrm{Diff}}}$
λ	Priming benefits for ”Same”	$\nu \times 4.5$	$\nu \times 1.5$	$\nu \times 1.5$
T_0	Motor Time	200	200	200

Implementation Equations

Architecture	Equation
General Equations	$v_i \sim N(\lambda \times \nu_i; s)$ $a_i \sim U(0; A_i)$ $\mathrm{RT} = \frac{\mathrm{b} - \sum a_i}{\frac{\sum v_i}{\sqrt{\mathrm{n}_{\mathrm{Total}}}}} + \mathrm{T}_0$

Table A.1. Variations of parameter implementation to further illustrate the model’s predictions. In the first column are the parameters for the simulation in which all priming benefits were removed, in the second column are the parameters for when accumulation variability is increased, in the third column are the parameters for when accumulation rate is increased, and in the fourth column are the parameters for when the threshold is increased. All other parameters are always identical to those of Table 4.1 except for the modified parameter (noted by the dash). The implementation equation is not modified in any of the simulations.

List of Parameters

		Values for Various Experiments			
Parameters	Descriptions	Adjusted Priming	Adjusted Accumulation	Adjusted Variability	Adjusted Threshold
A	Starting point upper bound	500	-	-	-
b	Threshold for ”Different” responses	10000	-	-	50000
	Threshold for ”Same” responses	$10000 \times n_{\text{Total}}$	-	-	$50000 \times n_{\text{Total}}$
ν	Mean accumulation rate of all evidence detectors	40	100	-	-
	Mean accumulation rate of all noise detectors	$\sqrt{40}$	$\sqrt{100}$	-	-
s	Standard deviation for ”Same” ν	$10 \times \sqrt{n_{\text{Same}}}$	-	$20 \times \sqrt{n_{\text{Same}}}$	-
	Standard deviation for ”Different” ν	$10 \times \sqrt{n_{\text{Diff}}}$	-	$20 \times \sqrt{n_{\text{Diff}}}$	-
λ	Priming benefits for ”Same”	$\nu \times 4$	-	-	-
T_0	Motor Time	200	-	-	-

Figure I.1. Typical response time (RT; left panel) and accuracy (right panel) trends found in a Same-Different task inspired by Bamber's (1969) seminal paradigm. Data stem from Same-Different task variants found in Chapter 2 (control conditions for Experiment 1 and 2). Dependent variable is located on the y-axis and letter string length is located on the x-axis. The number of differences within the string is denoted by each individual colored line. Error bars denote the correlation and difference adjusted 95% confidence interval of the mean (Baguley, 2012)

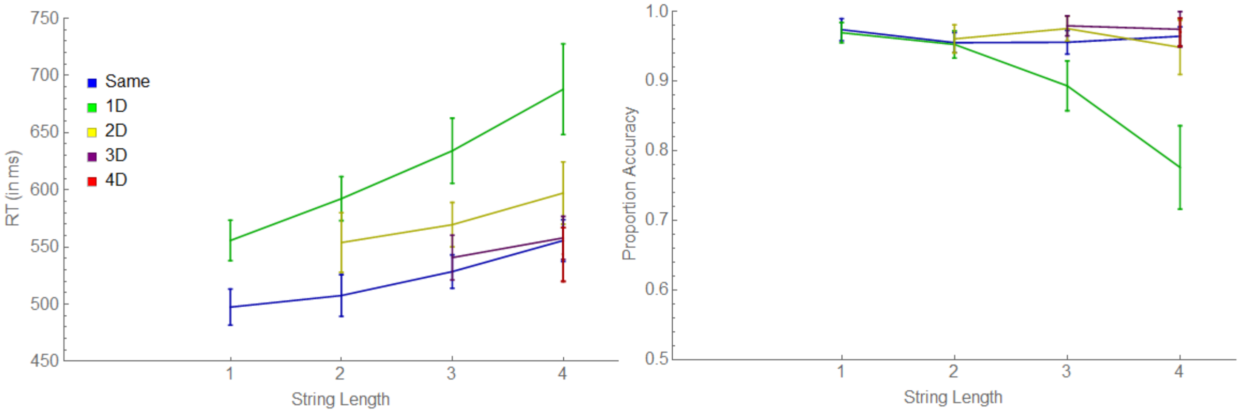


Figure 2.1. Timeline of a trial in Experiment 1, 2, and 4. S_1 denotes the first stimulus that is presented to the participant (the criterion stimulus) and S_2 denotes the second stimulus to be presented to the participant (the test stimulus). Feedback was only presented on error and non-response trials.

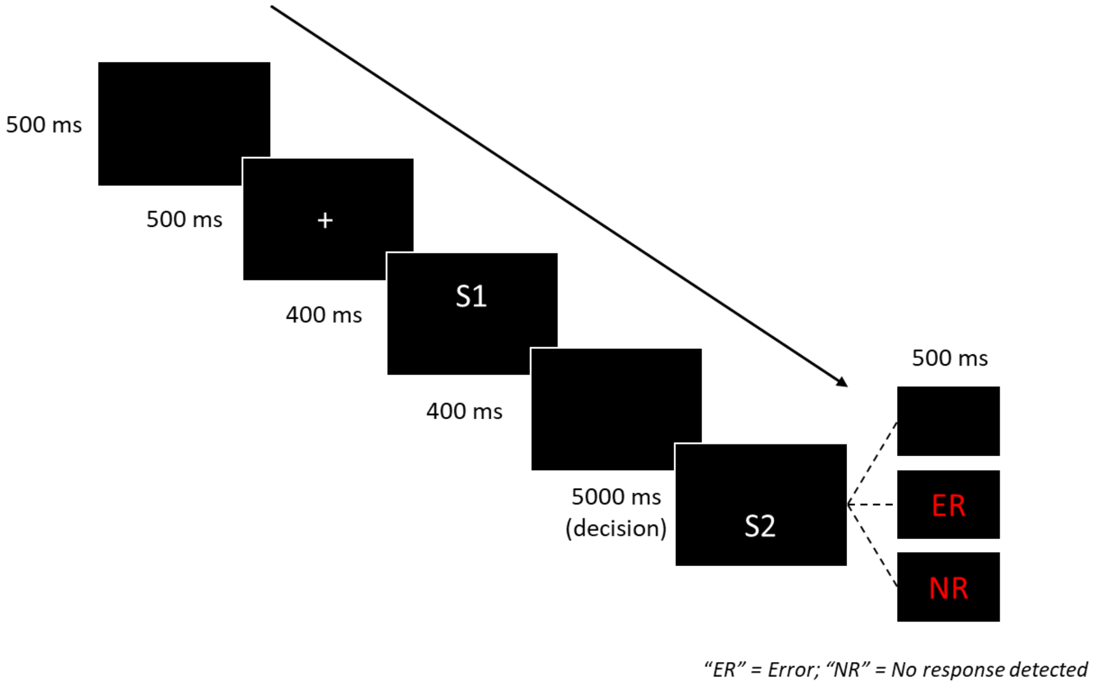


Figure 2.2. RT (top panels) and accuracy (bottom panels) results for Experiment 1. In the left columns are results for physically matching trials, and in the right columns are results for physically mismatching trials. Error bars denote the de-correlated and difference adjusted 95% confidence interval of the mean (Baguley, 2012).

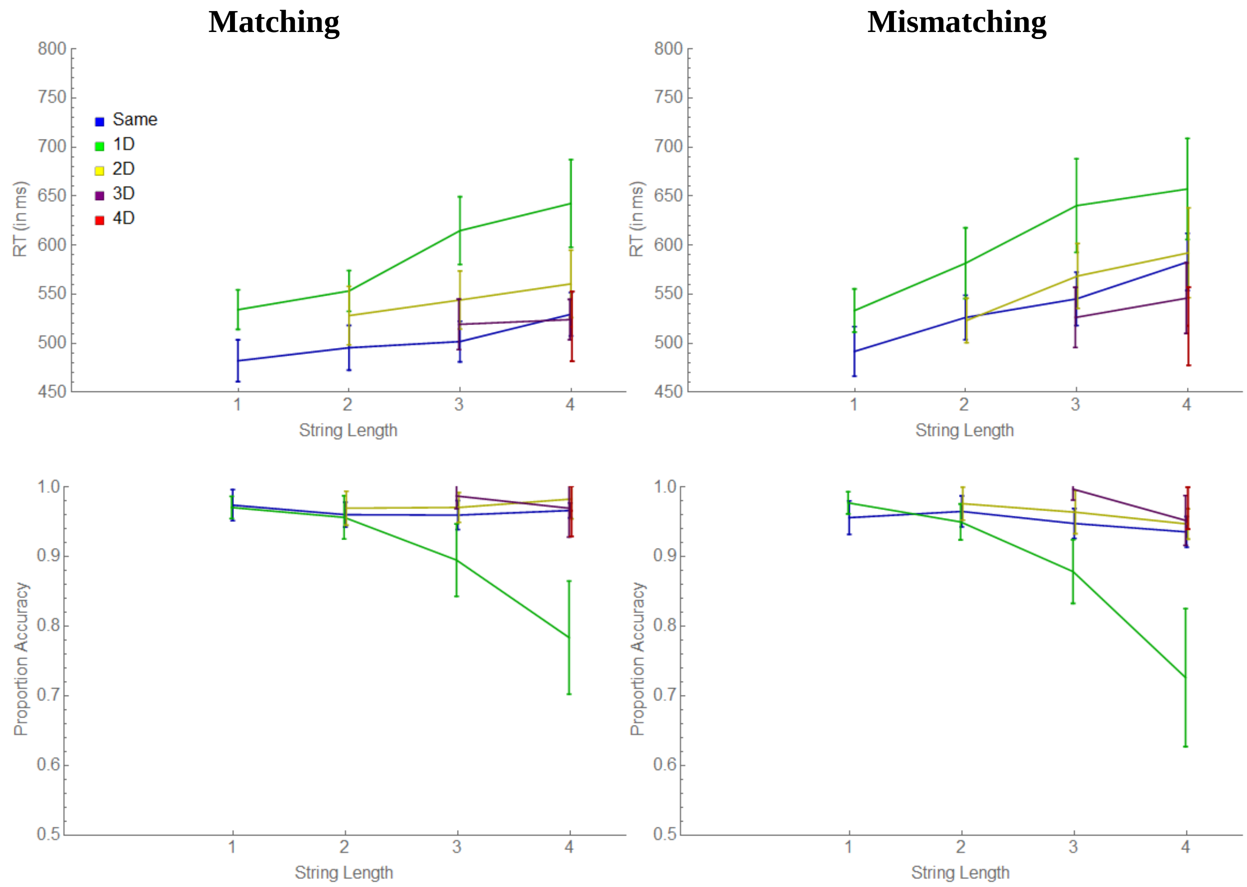
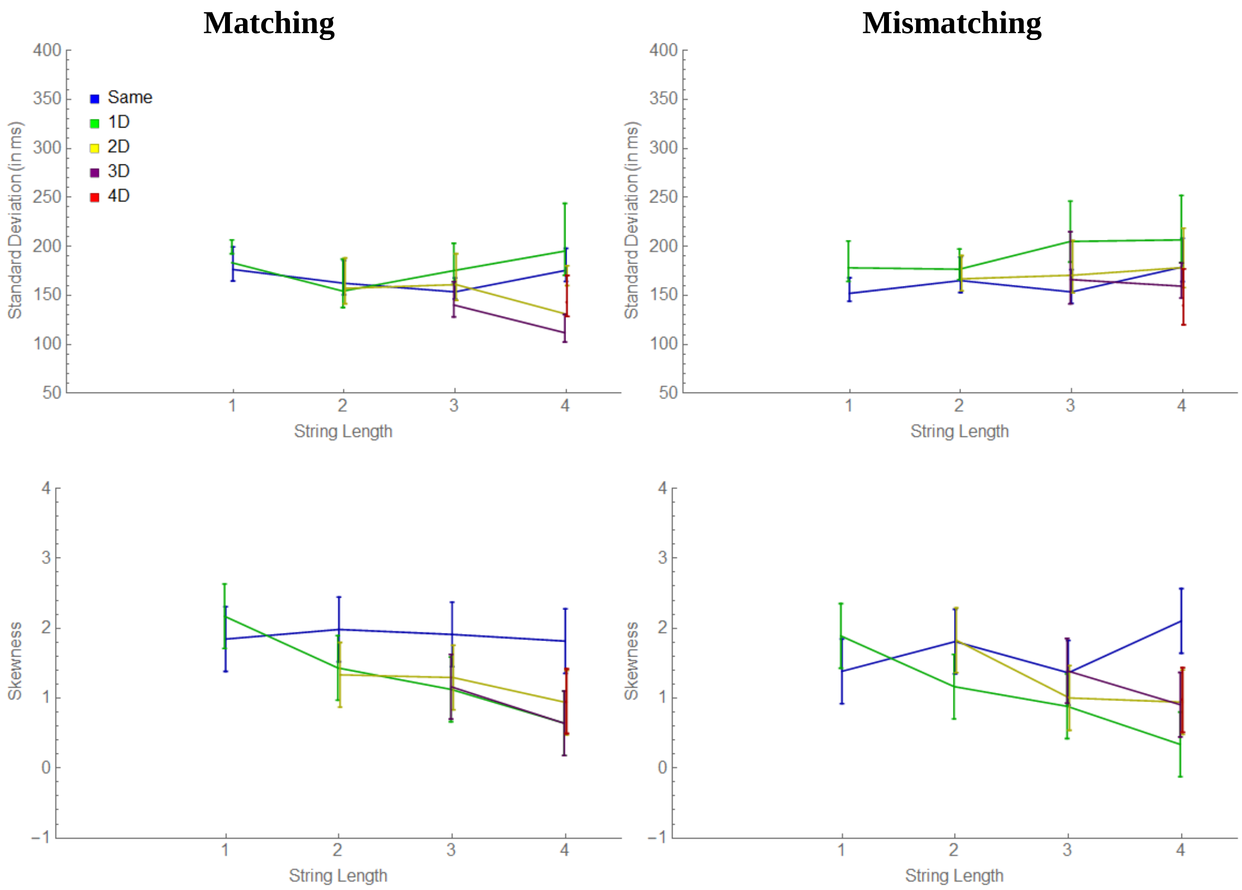


Figure 2.3. Mean standard deviation (top panels) and skewness (bottom panels) results for Experiment 1. In the left panels are results for physically matching stimuli, and in the right panels are when the stimuli are physically mismatching. Error bars denote the correlation and difference adjusted 95% confidence interval of the mean (Baguley, 2012) constructed with the proper estimator (Harding, Tremblay, & Cousineau, 2014).



Matching

Mismatching

Figure 2.4. RT (top panels) and accuracy (bottom panels) results for Experiment 2. In the left panels are results for physically matching trials, and in the right panels are results for physically mismatching trials. Error bars denote the de-correlated, difference adjusted 95% confidence interval of the mean (Baguley, 2012).

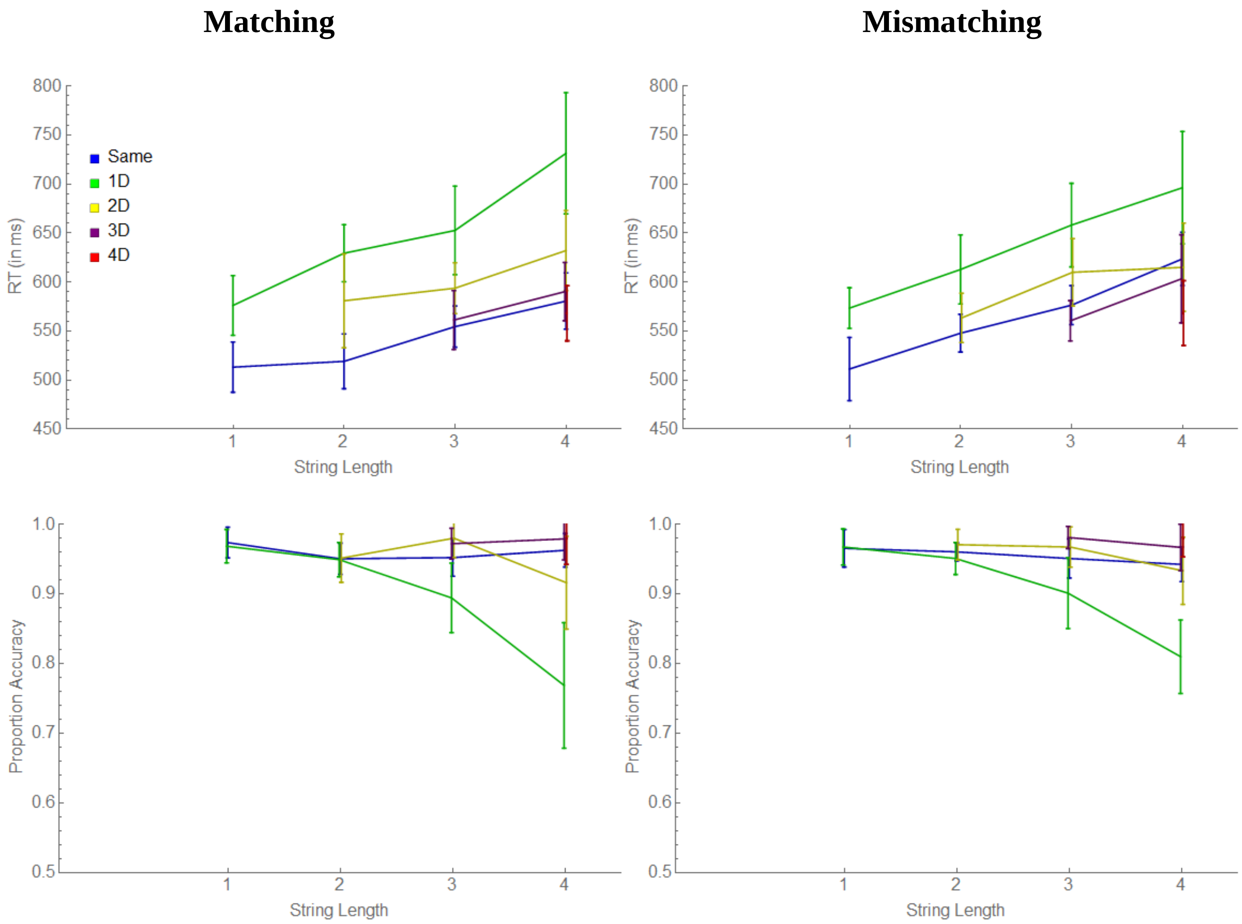


Figure 2.5. Mean standard deviation (top panels) and skewness (bottom panels) results for Experiment 2. In the left panels are results for physically matching stimuli, and in the right panels are results for physically mismatching stimuli. Error bars denote the de-correlated, difference-adjusted 95% confidence interval (Baguley, 2012) constructed with the proper estimator (Harding, et al., 2014).

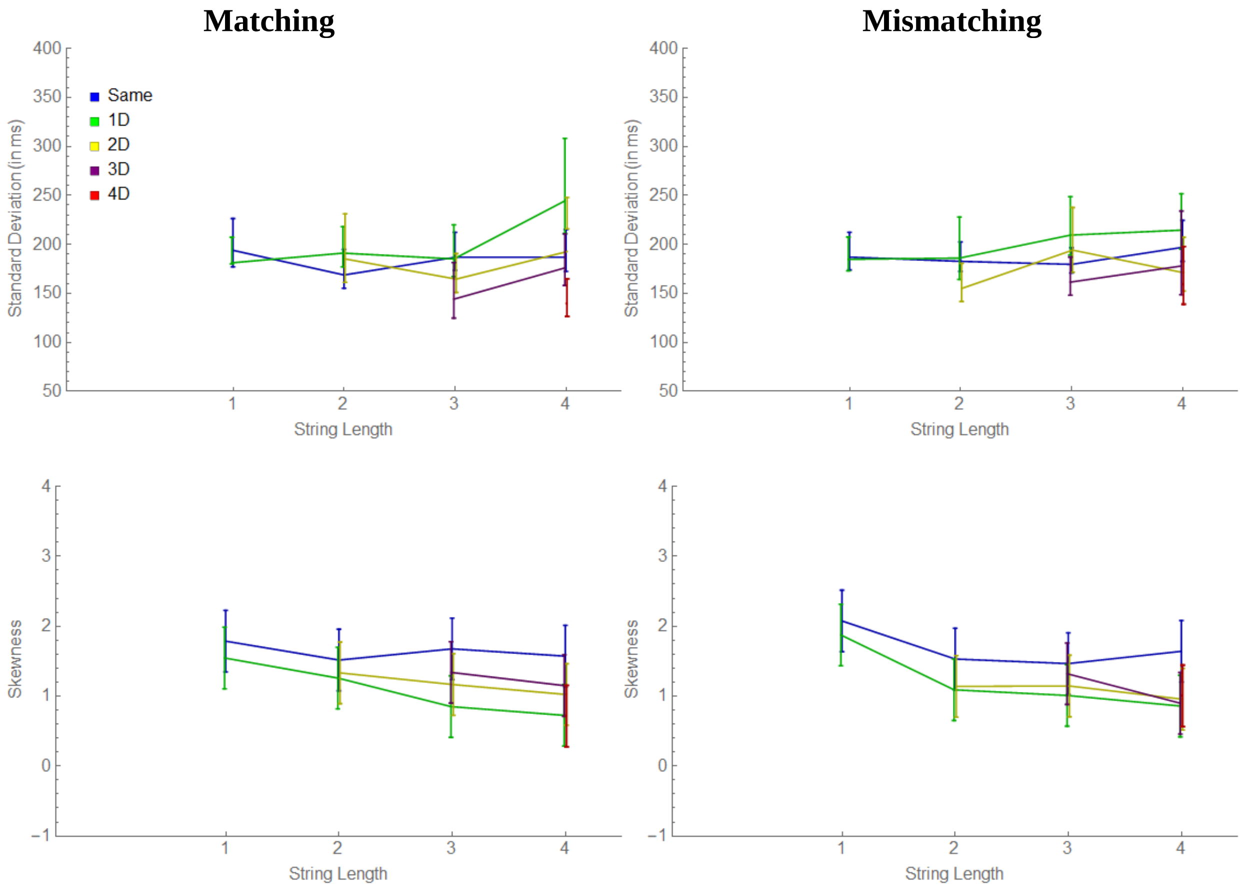
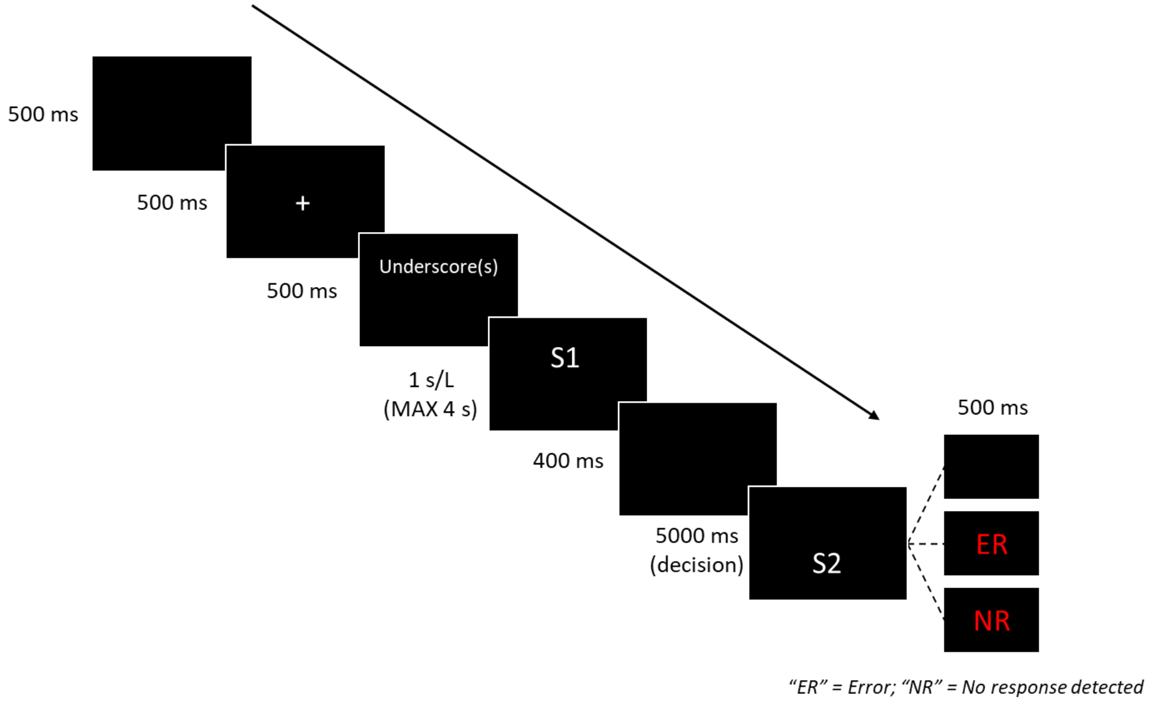
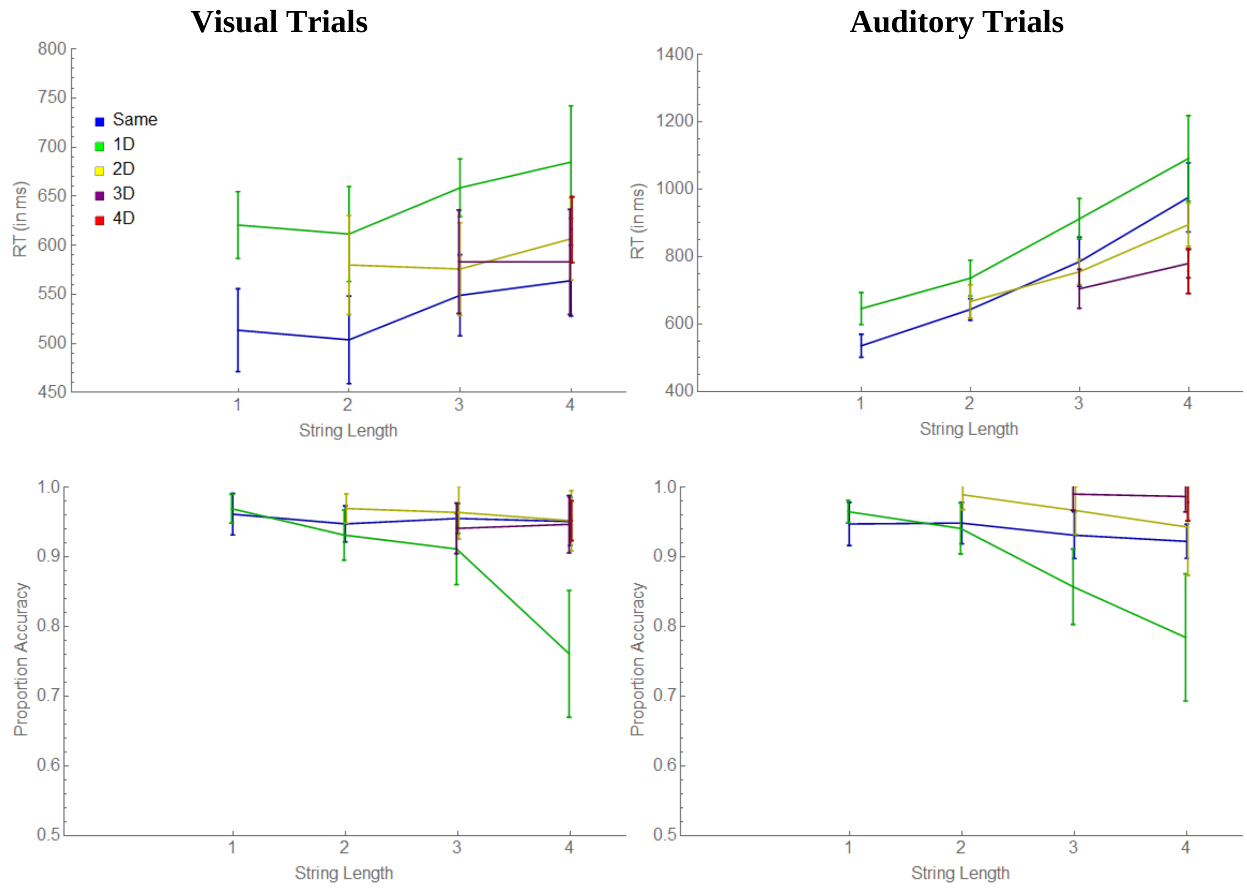


Figure 2.6. Timeline of the auditory condition for Experiment 3. S_1 denotes the auditory stimulus presented to the participant (letters are vocalized serially in 1 s increments) and S_2 denotes the second stimulus to be visually presented to the participant. The underscore(s) is/are presented immediately after the fixation slide and indicate the length of S_1. As seen, feedback is only given on errors and non-responses. In the visual condition, S_1 is presented visually for 400 ms.



"ER" = Error; "NR" = No response detected

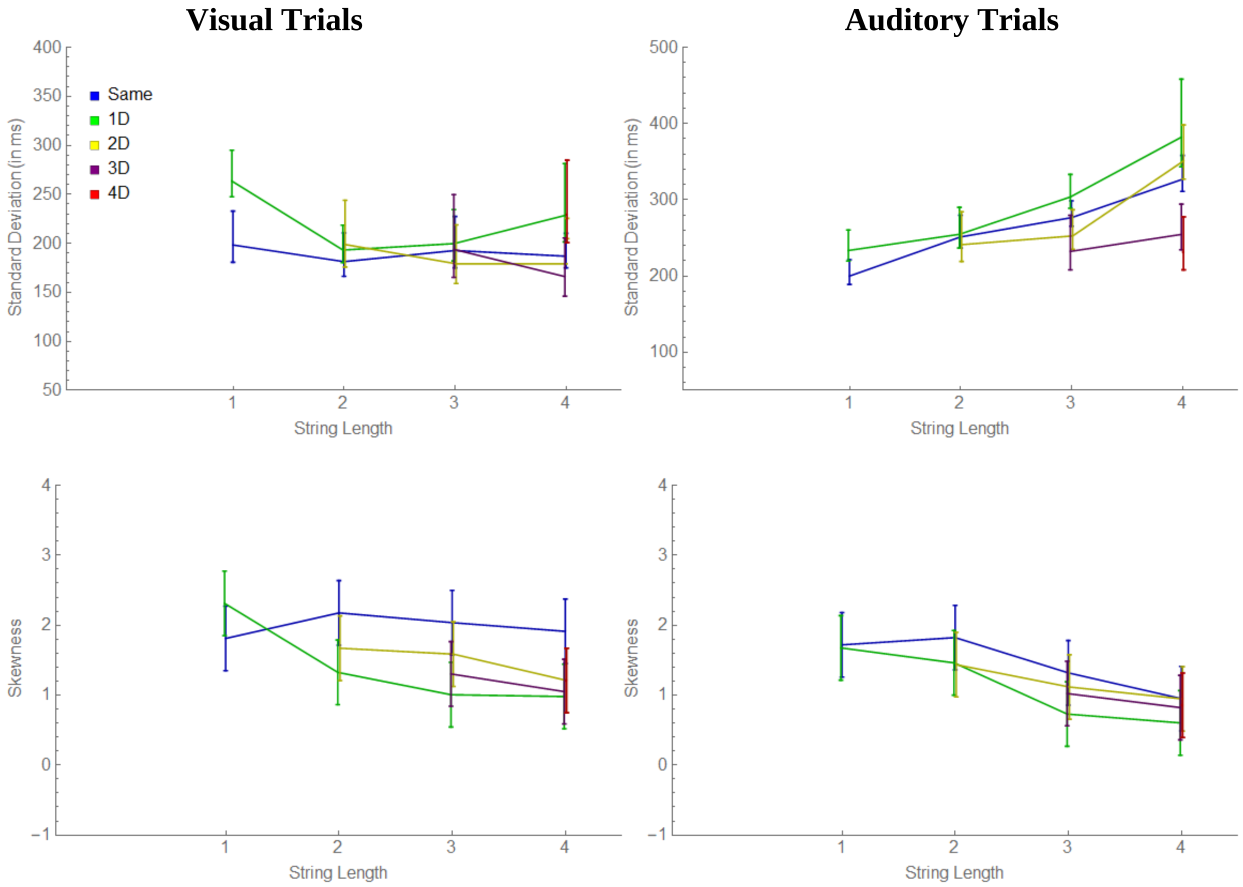
Figure 2.7. RT (top panels; note the different scales on the y-axis) and accuracy (bottom panels) results for Experiment 3. In the left panels are results for visually presented trials, and in the right panels are results for the auditory trials. Error bars denote the de-correlated, difference adjusted 95% confidence interval of the mean (Baguley, 2012).



Visual Trials

Auditory Trials

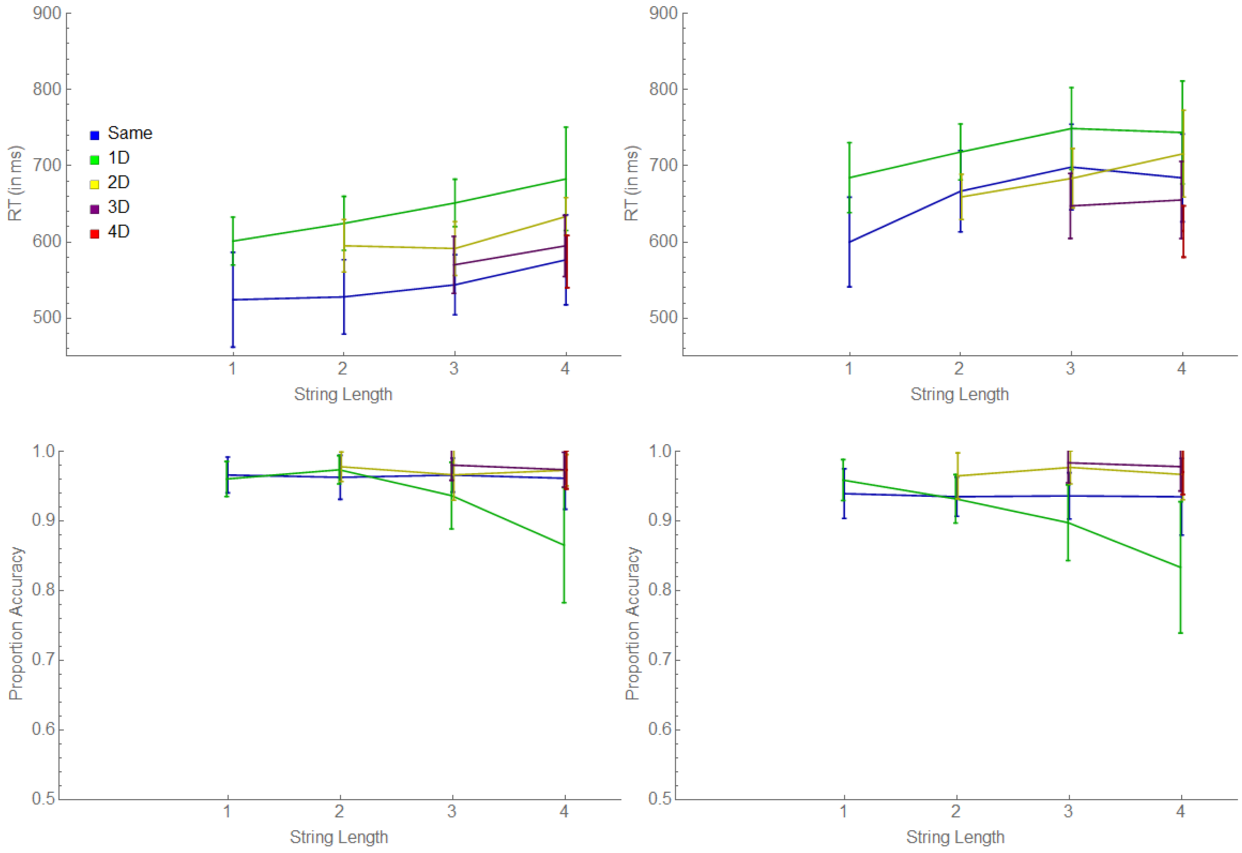
Figure 2.8. Mean standard deviation (top panels; note the differences on the y-axis) and skewness (bottom panels) results for Experiment 3. In the left panels are results for visually presented trials, and in the right panels are results for the auditory trials. Error bars denote the de-correlated, difference-adjusted 95% confidence interval (Baguley, 2012) constructed with the proper estimator (Harding, et al., 2014)



Visual Trials

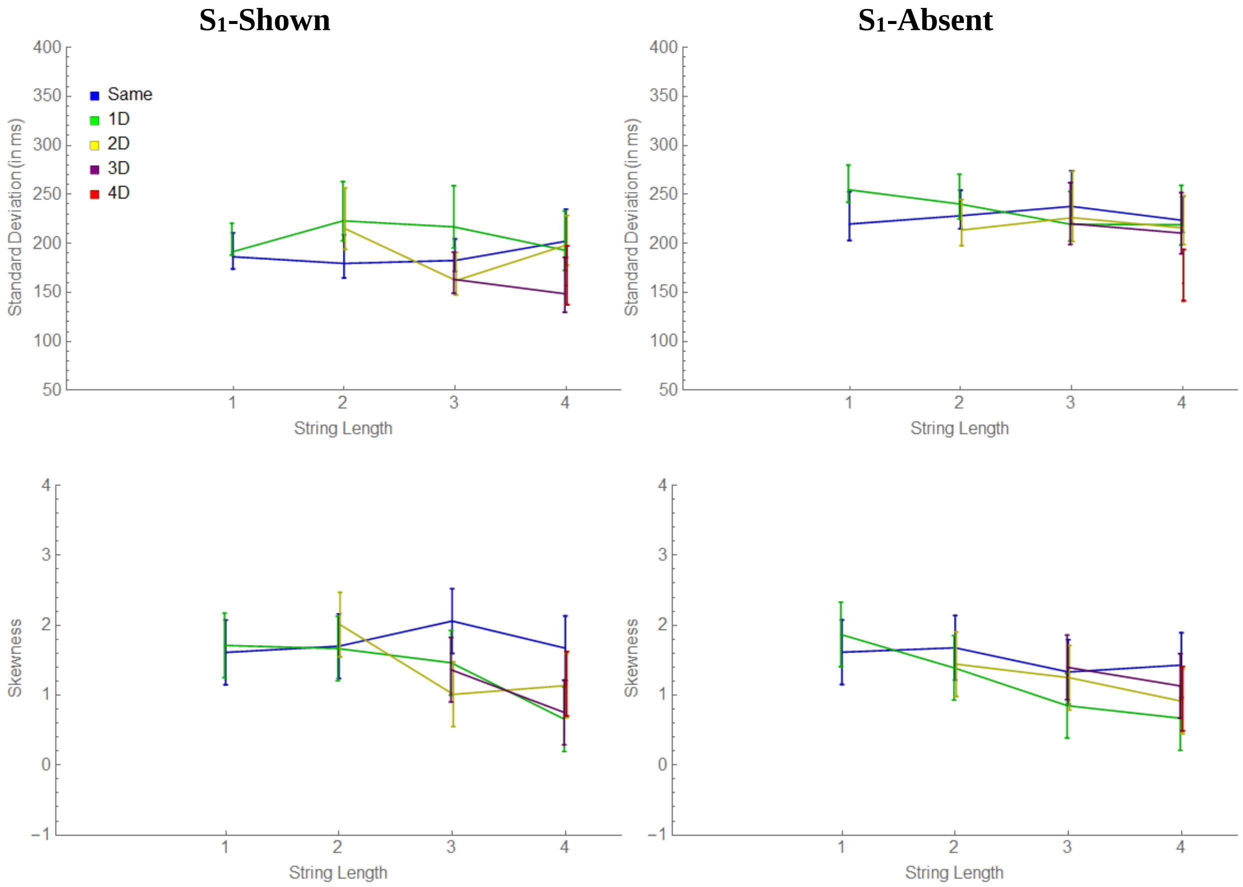
Auditory Trials

Figure 2.9. RT (top panels) and accuracy (bottom panels) results for Experiment 4. In the left panels are results for S_1-Shown trials, and in the right panels are results for S_1-Absent trials. Error bars denote the de-correlated, difference adjusted 95% confidence interval of the mean (Baguley, 2012).



S_1-Shown **S_1-Absent**

Figure 2.10. Mean standard deviation (top panels) and skewness (bottom panels) results for Experiment 4. In the left panels are results for S_1-Shown trials, and in the right panels are results for S_1-Absent trials. Error bars denote the de-correlated, difference-adjusted 95% confidence interval (Baguley, 2012) constructed with the proper estimator (Harding, et al., 2014).


S_1-Shown

S_1-Absent

Figure 3.1. RT (top-left panel), accuracy (top-right panel), standard deviation (bottom-left panel), and skewness (bottom-right panel) results for Experiment 1. Error bars denote the correlation and difference adjusted 95% confidence interval (Baguley, 2012) using the appropriate SE and CI estimator (Harding, Tremblay, & Cousineau, 2014)

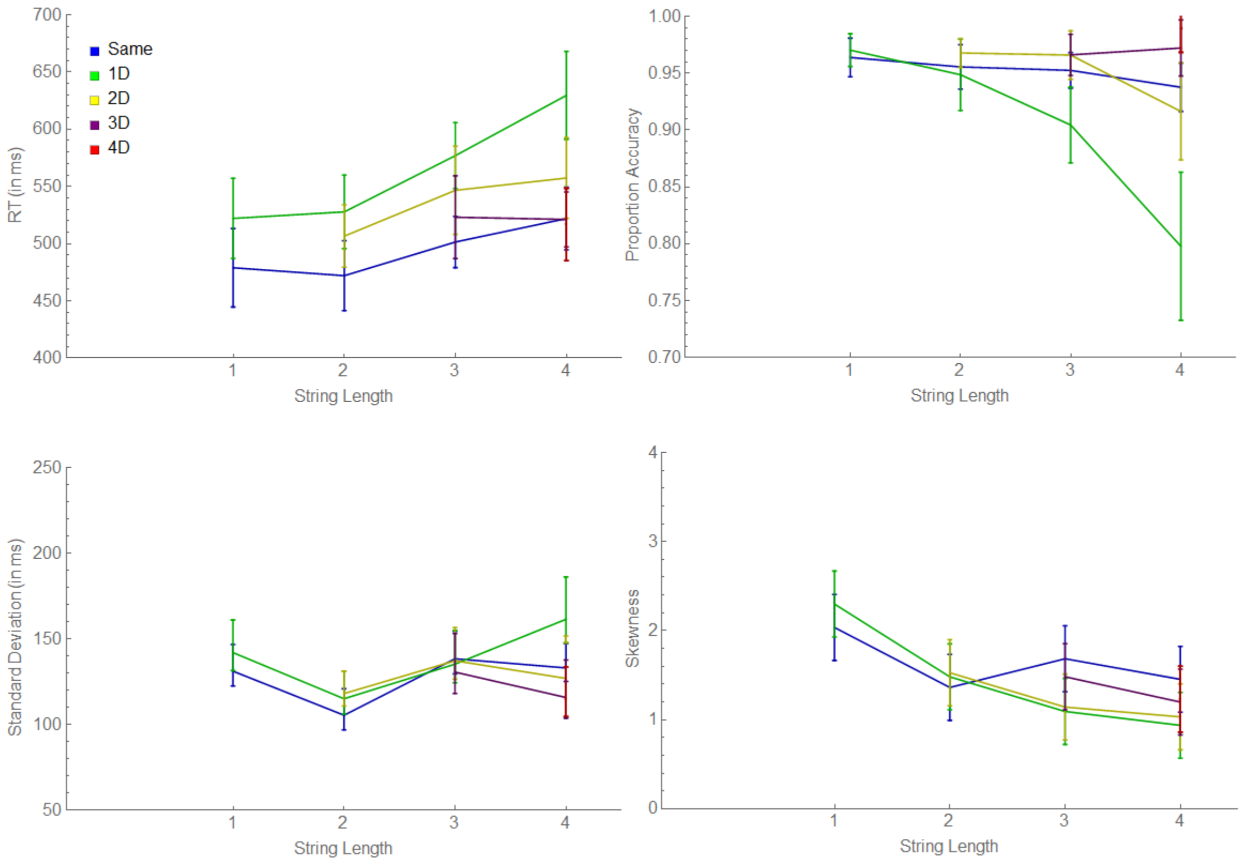


Figure 3.2. Results of the Race Model Inequality analysis for Experiment 1. In the left panel are the results for 2L stimuli and in the right panel are the results for the 3L stimuli. Each condition is denoted appropriately by solid coloured lines while the composite bounds are denoted by segmented lines. If the Both (2 letters) and All (three letters) mismatch condition crosses a composite bound at any point along its progression, there are redundancy gains and thus coactivity within the system.

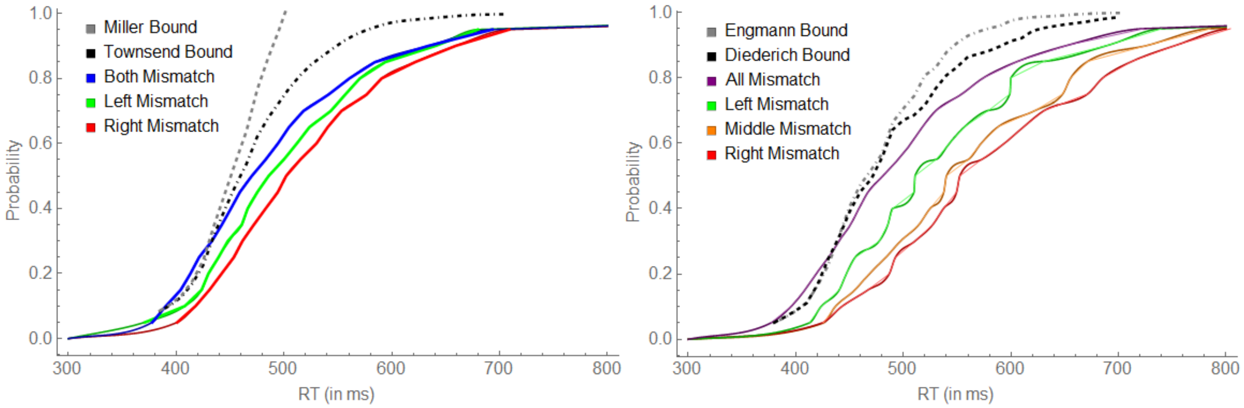


Figure 3.3. Mean RT (left panel) and accuracy (right panel) results for Experiment 2. Error bars denote the correlation and difference adjusted 95% confidence interval (Baguley, 2012)

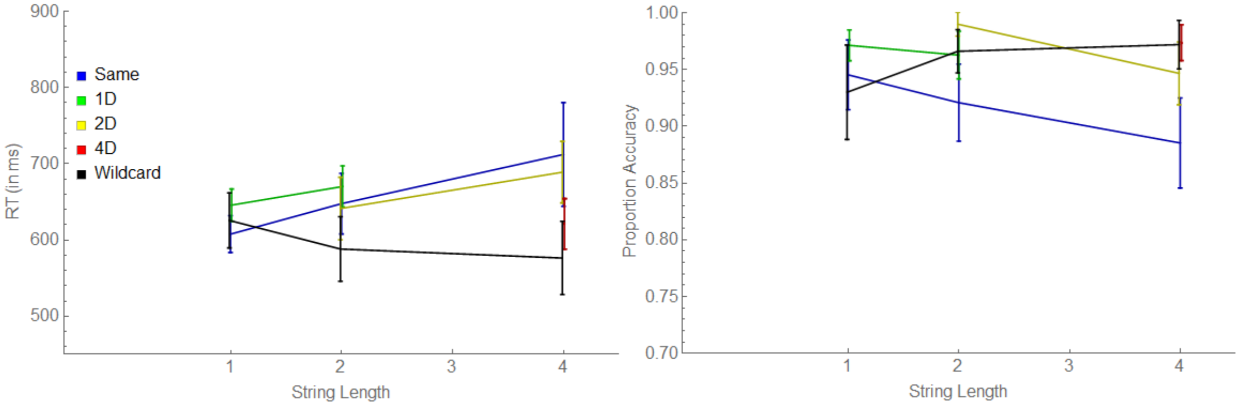


Figure 3.4. Response time distributions for X-0D1L, X-0D2L, and X-0D4L (left panels, in order, descending the page) as well as standard deviation (top-right panel) and skewness (bottom-right panel) results for Experiment 2's wildcard-"Same". Error bars denote the correlation and difference adjusted 95% confidence interval (Baguley, 2012) using the appropriate SE and CI estimator (Harding et al., 2014)

(CONTINUED ON NEXT PAGE)

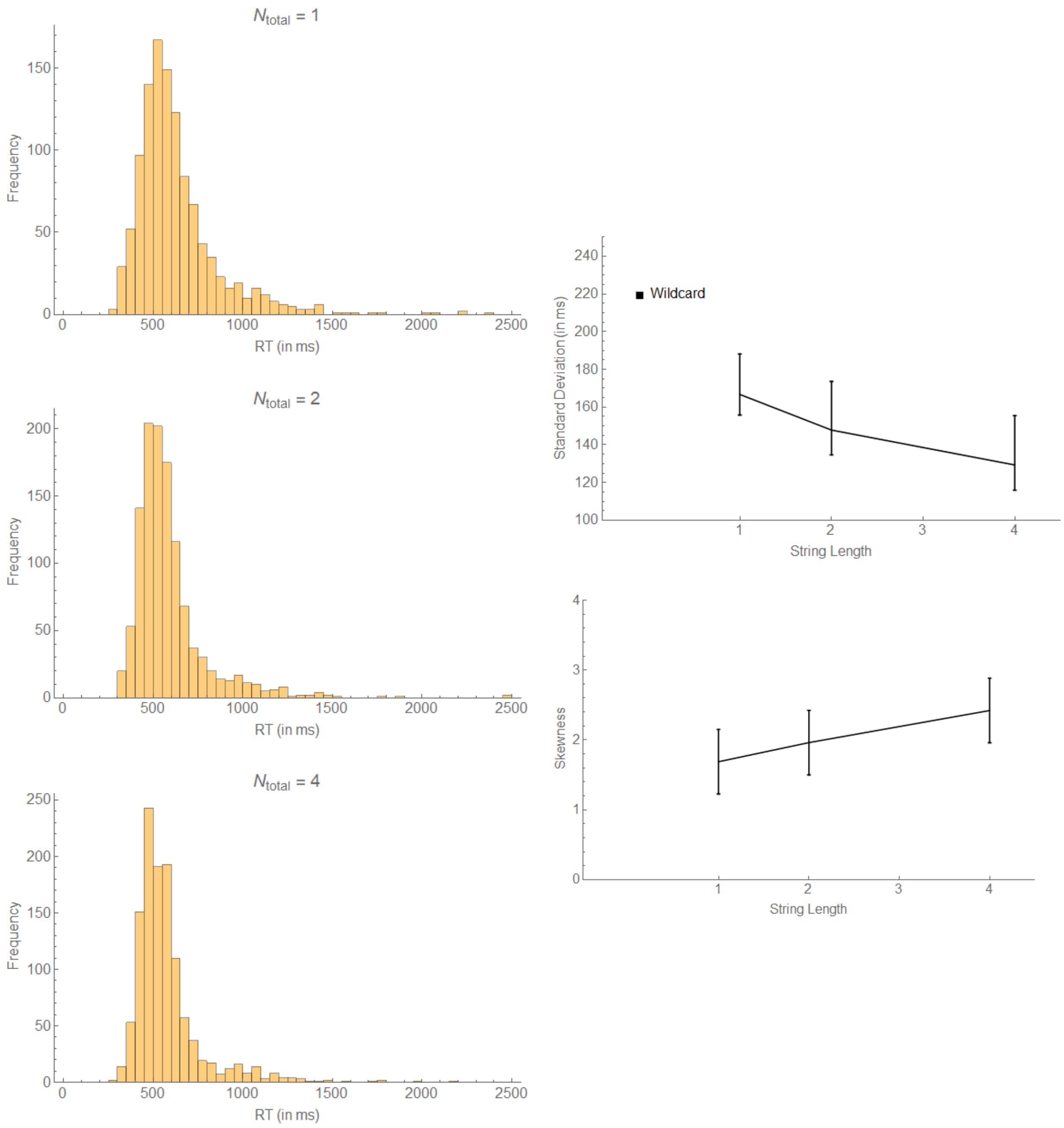

$N_{\text{total}} = 1$
Frequency
0
50
100
150
0
500
1000
1500
2000
2500
RT (in ms)
$N_{\text{total}} = 2$
Frequency
0
50
100
150
200
0
500
1000
1500
2000
2500
RT (in ms)
$N_{\text{total}} = 4$
Frequency
0
50
100
150
200
250
0
500
1000
1500
2000
2500
RT (in ms)
■ Wildcard
Standard Deviation (in ms)
100
120
140
160
180
200
220
240
1
2
3
4
String Length
Skewness
0
1
2
3
4
1
2
3
4
String Length

Figure 4.1. Schematic of competing "Same" and "Different" decisions when faced with a trial. In the top-left panel is a 0D1L stimulus, in the top-right panel is a 1D2L stimulus, in the bottom-left panel is a 1D4L stimulus, and in the bottom-right panel is a X-0D4L stimulus. Individual detectors are represented as arrows of which green arrows represent the response that triggered the decision whereas the red arrows represent the decision that was not made. When the arrow is full, it represents a standard accumulation of evidence whereas when the arrow is segmented, it represents an accumulation of noise (noise is sampled from a rate distribution with a smaller mean; the accumulation remains linear). The black horizontal line represents the threshold and the grey horizontal dotted line in the "Same" accumulator represents where a self-terminating threshold would be. The vertical dotted grey line in both accumulators represents when that decision was recorded.

(CONTINUED ON NEXT PAGE)

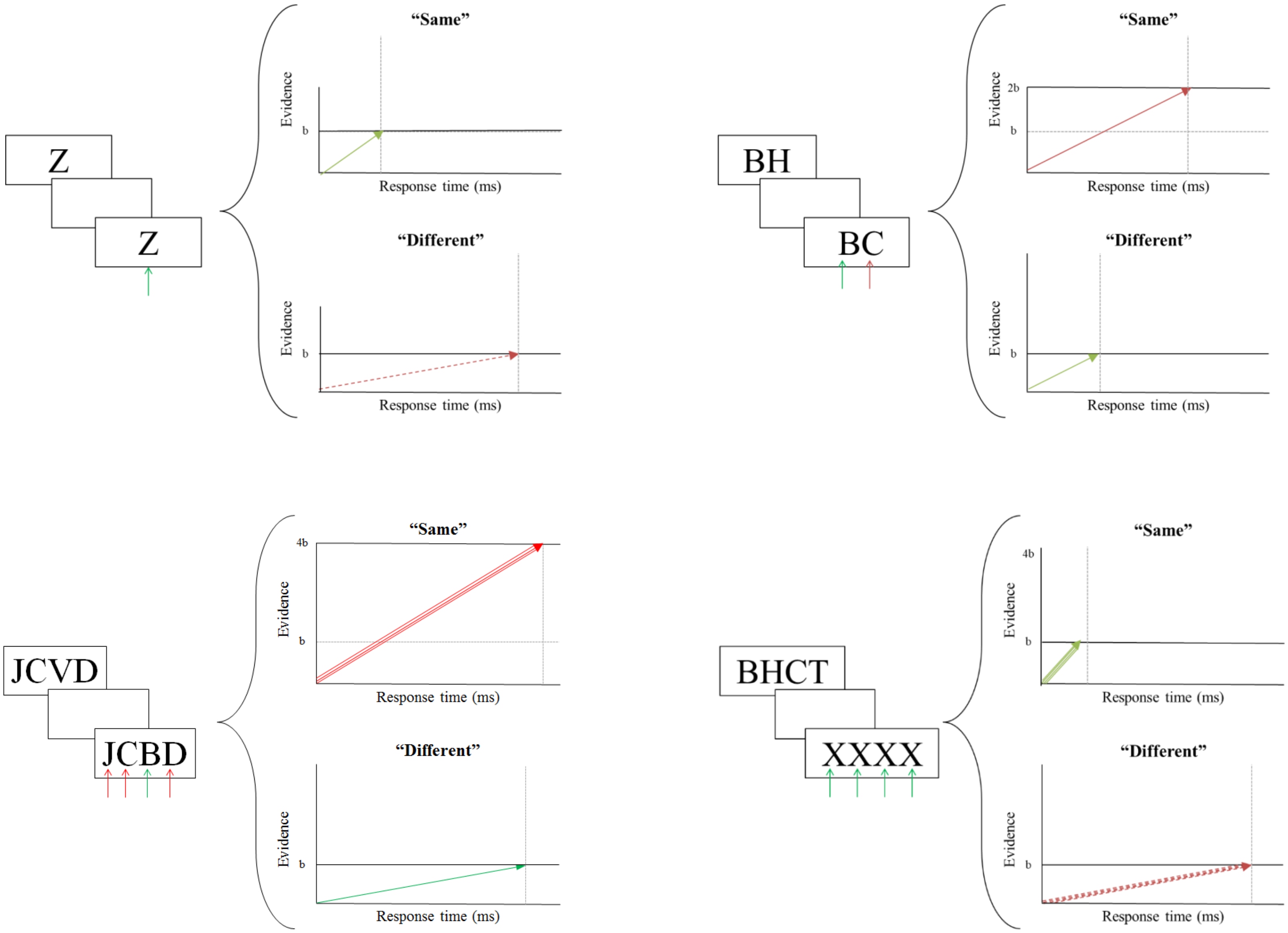

Z
Z
"Same"
Evidence
b
Response time (ms)
"Different"
Evidence
b
Response time (ms)
BH
BC
"Same"
Evidence
2b
b
Response time (ms)
"Different"
Evidence
b
Response time (ms)
JCVD
JCBD
"Same"
Evidence
4b
b
Response time (ms)
"Different"
Evidence
b
Response time (ms)
BHCT
XXXX
"Same"
Evidence
4b
b
Response time (ms)
"Different"
Evidence
b
Response time (ms)

Figure 4.2. The proposed model's predictions of RT (top panels) and accuracy (bottom panels) for Same-Different task variants discussed within this thesis. In the left panels are the predictions of a standard Same-Different task (explored in the control conditions of Chapter 2 and in Experiment 1 of Chapter 3), in the middle panels are the predictions when priming for “Same” is attenuated/cancelled (explored in Chapter 2), and in the right panels are the model’s predictions when “Same” responses are replaced with wildcard stimuli (explored in Chapter 3). Each individual condition represents the mean of simulated data from 1000 individual RTs. Error bars denote the appropriate 95% confidence interval of the mean.

(CONTINUED ON NEXT PAGE)

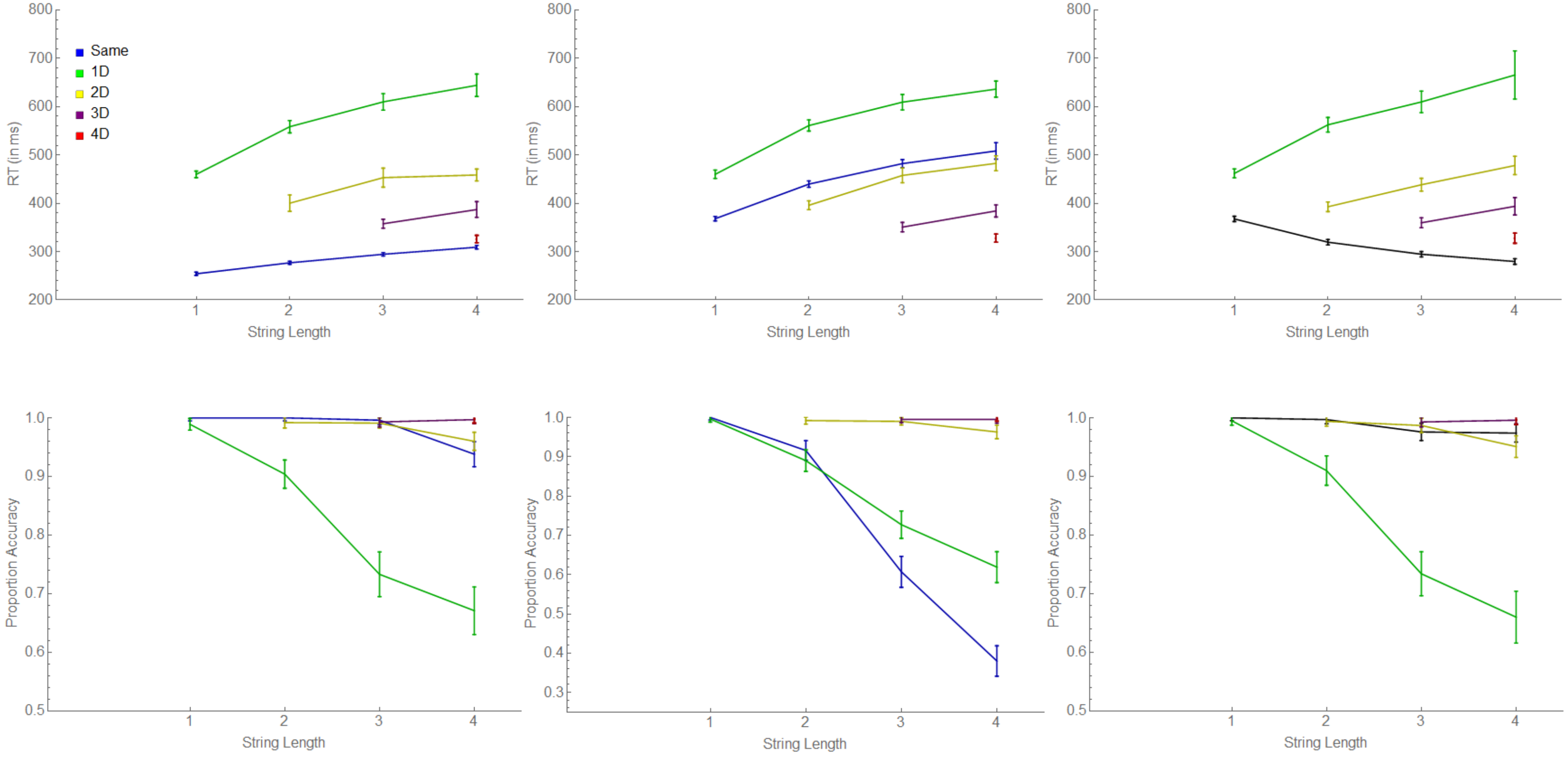

■ Same
■ 1D
■ 2D
■ 3D
■ 4D
RT (in ms)
800
700
600
500
400
300
200
String Length
1
2
3
4
Proportion Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3

Figure 4.3. The proposed model's predictions of standard deviation (top panels) and skewness (bottom panels) for Same-Different task variants discussed within this thesis. Columns follow the same order as Figure 4.2. Each individual condition represents the mean of simulated data from 1000 individual RTs. Error bars denote the appropriate 95% confidence interval of the mean (Harding, Tremblay, & Cousineau, 2014).

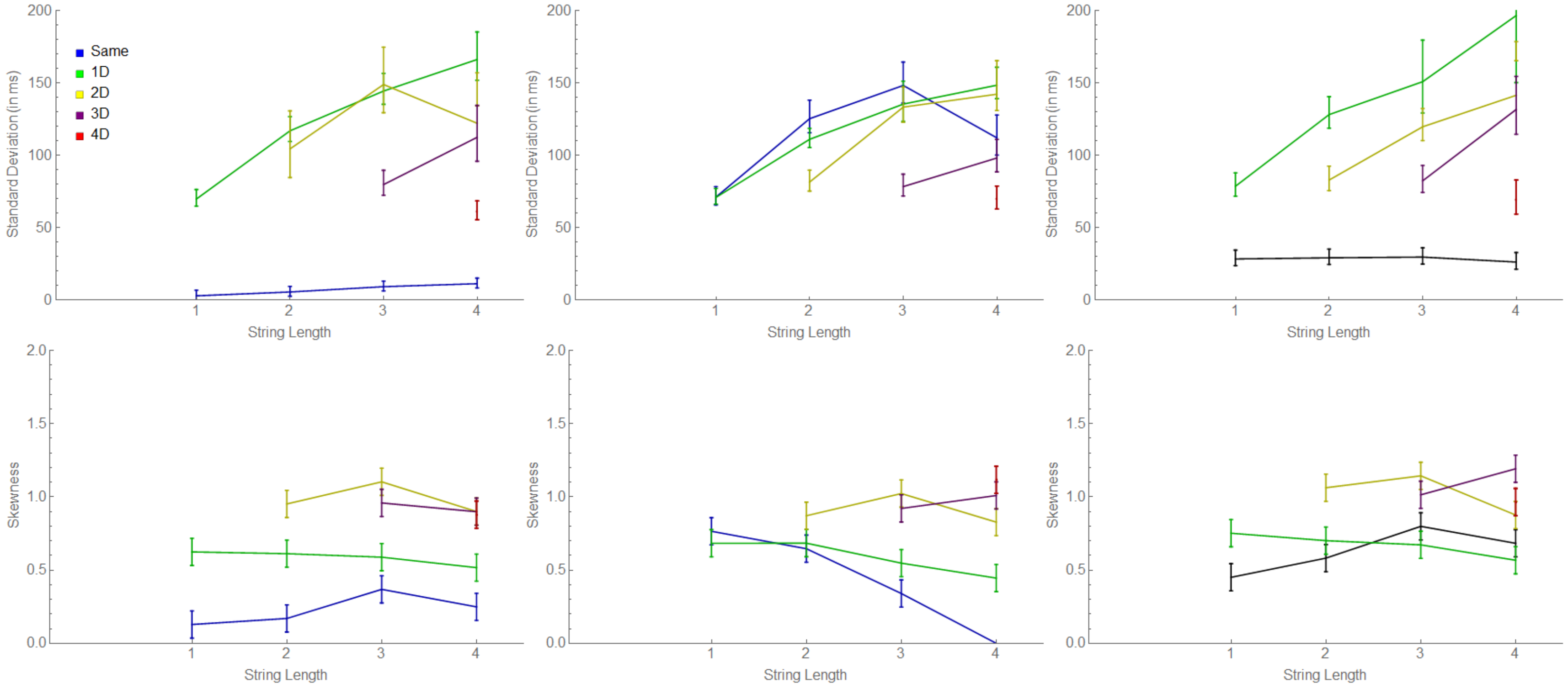
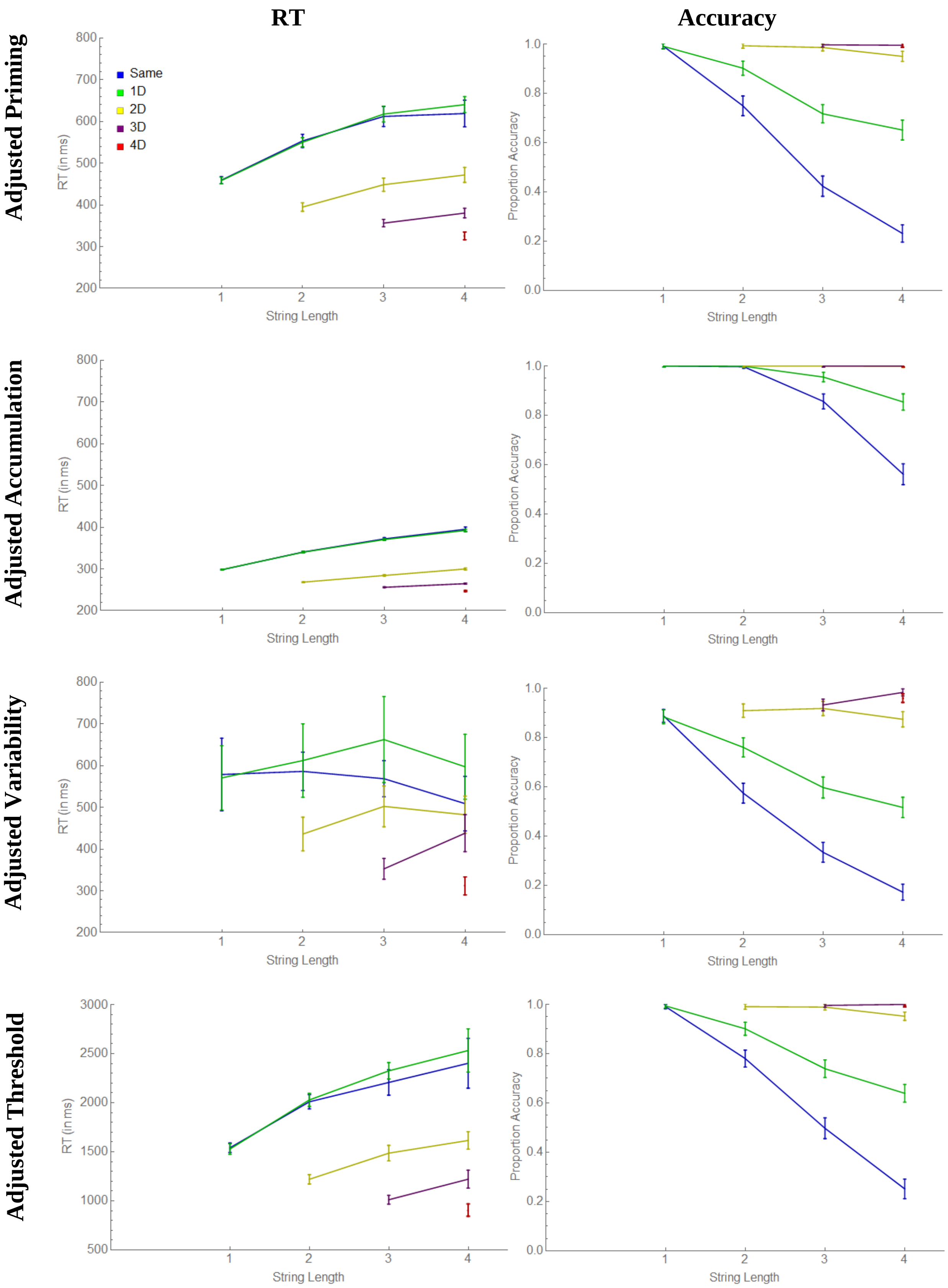


Figure A.1. The proposed model's predictions of RT (left panels) and accuracy (right panels) for Same-Different task variants discussed within Appendix A. In the first row are the model's predictions for adjusted priming, in the second row are the model's predictions for adjusted accumulation, in the third row are the model's predictions for adjusted variability, and in the fourth row are the model's predictions for adjusted threshold. All other details are identical to those presented in Chapter 4.

(CONTINUED ON NEXT PAGE)


RT
Accuracy
Adjusted Priming
Adjusted Accumulation
Adjusted Variability
Adjusted Threshold
Same
1D
2D
3D
4D
RT (in ms)
Proportion Accuracy
String Length
1
2
3
4
200
300
400
500
600
700
800
500
1000
1500
2000
2500
3000
0.0
0.2
0.4
0.6
0.8
1.0