The time course recovery of confidence judgments using interruptions

Kevin Zish (kzish@gmu.edu)
George Mason University
4400 University Dr, Fairfax, VA 22030

Nathan Aguiar (naguiar@masonlive.gmu.edu)
George Mason University
4400 University Dr, Fairfax, VA 22030

Malcolm McCurry (malcolm.mccurry.ctr@nrl.navy.mil)
Peraton
12975 Worldgate Dr, Herndon, VA 20170

Erik M. Altmann (ema@msu.edu)
Michigan State University
316 Physics Rd, East Lansing, MI 48824

J. Gregory Trafton (greg.trafton@nrl.navy.mil)
U.S. Naval Research Laboratory
4555 Overlook Ave SW, Washington, DC 20375

Abstract

Participants were interrupted during a procedural ask. Choice response time and confidence were measured up to seven trials after the interruption. Empirical data suggests a curvilinear pattern of recovery for choice response time and confidence. Two models of recovery for choice response time and confidence judgments were built. A comparison of the models provided support for post-decisional theories of confidence which suggest that confidence judgments are formed after a choice is made.

Keywords: ACT-R, confidence, activation, modeling, decision-making

Introduction

When are confidence judgments formed? Current theories of confidence debate whether confidence judgments are formed at the time a decision is made or after. If someone asks: "How confident are you that the person you identified committed the crime?", decisional theories of confidence say that confidence is available at the same time as the decision. Post-decisional theories of confidence say that confidence is only available after the decision is made.

Decisional theories are modeled in the context of signal detection theory (SDT: Green & Swets, 1966) and the role of strength in recognition memory (Egan, Schulman, & Greenberg, 1959; Hart, 1967; Norman & Wickelgren, 1969; Wickelgren, 1968). In these theories, confidence judgments are directly related to the strength of a retrieved memory. Stronger memories elicit higher confidence responses than weaker memories. Confidence is the strength of a memory or the memory's distance to a decision criterion (Donaldson, 1996; Wixted & Mickes, 2010). The confidence judgment is tied to the decision process and confidence is available at the same time a decision is made.

Post-decisional theories state that forming a confidence judgment begins after a decision is made (Pleskac & Busemeyer, 2010; Vickers, 2001, 2014). These theories typically use sequential sampling models (Juslin & Olsson, 1997; Vickers, 1970), specifically drift diffusion models (Heath, 1984; Laming, 1968; Link & Heath, 1975; Ratcliff, 1978). In drift diffusion models, choice begins at some point z and evidence accrues on a series of counters (usually one) towards a criterion for decision A or decision B. Confidence is calculated as the speed with which evidence accumulates towards a criterion (Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2009, 2013) or as the difference between the counters (Merkle & Van Zandt, 2006; Van Zandt & Maldonado-Molina, 2004; Vickers, 2014).

There is evidence for both decisional and post-decisional theories of confidence. In a recent study, Dotan, Meyniel, & Dehaene (2018) provided support for decisional theories of confidence. Participants were presented with arrows that pointed to the left or right side of the screen. Arrows were presented one at a time. Participants were asked to move their finger from the bottom middle to targets at either the top left or top right of the screen. Participants moved to the side they thought the majority of the arrows was pointing towards. The authors found that y-speed, the speed that a participant was heading towards one choice or the other, was a better predictor of confidence than choice response time (RT). Because instantaneous speed was a better predictor of confidence than final choice RT, the authors argued that confidence is calculated online and repeatedly throughout the process of a judgment.

Post-decisional theories suggest that evidence continues to be collected after a choice is made as shown by response reversals (Resulaj, Kiani, Wolpert, & Shadlen, 2009; Van Zandt & Maldonado-Molina, 2004). Therefore, confidence judgments are based on the evidence collected for the choice plus some additional evidence collection.

One influential post-decisional theory is the two stage dynamic signal detection theory (2DSD: Pleskac & Busemeyer, 2010) which models the relationship between choice and confidence in a drift diffusion model. Participants collect evidence for a choice using a standard drift diffusion process. Evidence retrieved from memory or the environment accrues on a counter towards a response threshold for given alternatives. When the counter reaches the choice criteria a choice is made. The evidence collected for the choice continues to be sampled to make a confidence judgment.

There is also support for both decisional and post-decional explanations in a single experiment (Baranski & Petrusic, 1998). Baranski & Petrusic (1998) asked participants to determine the longer or shorter of two horizontal lines under accuracy stress or speed stress. Under speed stress, the authors found a negatively linear relationship for confidence response time and confidence, where faster responses were more confident than slower responses. Under accuracy stress there was no relationship. The authors argued that speed stress caused participants to spend less time calculating confidence during the primary decision and rendered confidence post-decisonally. In contrast, participants had no change in confidence response time across levels of confidence for accuracy stress because confidence was processed during the decision.

The evidence from Dotan et al. (2018), 2DSD, and Baranski & Petrusic (1998) suggest that rendering confidence judgments is a cognitively complex process that can originate at different times during decision making.

One way to investigate cognitively complex processes is to interrupt them and record performance data at different points in time as the process recovers. Altmann & Trafton (2007) used interruptions to investigate time course recovery of memory and attention after an interruption.

Here, we incorporated confidence judgments in a procedural task that generates rich error data. Error data allows us to investigate process recovery of confidence judgments. Confidence judgments are relevant to study in this context because they could help inform both models of information processing in procedural tasks and performance in applied contexts where procedural errors have high potential cost.

Altmann & Trafton (2007) had participants complete a dynamic decision making task. Participants were interrupted after completing an action. The time to resume the task after the interruption (i.e. resumption lag: Trafton, Altmann, Brock, & Mintz, 2003) was measured up to seven actions after the interruption was complete. The authors found that interruptions increased resumption time after the interruption. However, participants did not immediately recover after the first action following an interruption. Instead, the data showed that resumption lag followed a curvilinear pattern of recovery wherein the response time for each action after the interruption was faster than the preceding action.

Altmann & Trafton (2007) developed a mathematical model that fit the pattern of recovery for resumption time. The model suggests that in procedural tasks spreading activation plays a critical role in facilitating the selection of the next action in a task. Activation spreads through associative links that form between actions in a task. An associative link means that when an action is completed, priming from the completed action is added to the activation of all following actions. The activation for an element at position p is represented by

$$A(p) = -1 + \sum_{i=1}^{p} assoc^{i-1}, assoc < 1$$

where *assoc* is the amount of activation received by preceding elements of the task that have already been retrieved. The action directly after the current action receives the most priming. Subsequent actions receive lower and lower amounts of priming.

When participants are allowed to complete an uninterrupted task, an element receives small amounts of priming from each associatively linked action before it. Interruptions effectively cut off that priming making it so that activation for the action to be resumed is lower than if it had not been interrupted. The curvilinear pattern of choice RT is evidence of cumulative priming building for the task as it recovers.

The recovery process of choice RT is based on the ACT-R theory (Anderson et al., 2004) and is represented by

$$RT(p) = F * exp[1 - \sum_{i=1}^{p} assoc^{i-1}]$$

where F is a scaling parameter representing non-decisional processes.

Using a similar approach, we can investigate the time course of confidence judgments by interrupting participants during a decision and building a model of choice, taken from Altmann & Trafton (2007), and a model of confidence as those processes recover. We assume, as does SDT theory, that confidence is a scaled measure of strength. We represent strength using activation which is the *assoc* parameter in the Altmann & Trafton (2007) model.

Activation, has been explored extensively in the ACT-R cognitive architecture (Anderson et al., 2004) of which the findings of Altmann & Trafton (2007) were based. We used the activation-based properties of ACT-R to make predictions about the relationship between choice RT and confidence.

ACT-R suggests that activation of a memory element m is defined by the relationship between the number of times a goal has been rehearsed n and the time that has passed T. The following is a simplified equation for activation adapted from equation 2.2 in Anderson, Bothell, Lebiere, & Matessa (1998).

$$m = ln(n/\sqrt{T})$$

Goals that have been rehearsed many times in the recent past have more activation than goals that have been rehearsed fewer times or rehearsed in the distant past.

Goals also undergo decay. Decay is an important part of forgetting and is indexed by time. The more time that has passed since a goal has received activation (from being retrieved or from associative priming) the lower the activation for the goal. Interruptions decrease activation by decreasing the probability that a goal can be rehearsed, increasing the amount of time between rehearsal, or some combination of the two.

Because we assume that confidence is based on activation, it follows that confidence should behave in several systematic ways according to ACT-R. First, confidence should decrease after an interruption because interruptions decrease activation. This finding has already been demonstrated by Aguiar, Zish, McCurry, & Trafton (2016) and Zish, Hassanzadeh, McCurry, & Trafton (2015).

Second, confidence should increase in a curvilinear pattern after an interruption similar to Altmann & Trafton (2007). Cumulative priming after an interruption should result in a decrease in confidence after an interruption followed by a gradual increase in confidence for later actions as the decision process recovers.

Third, a mathematical model of the time course of recovery for confidence C can be built that will match empirical data. The model we propose is

$$C(p) = S * exp[-1 + \sum_{i=1}^{p} assoc^{i-1}]$$

Different parameters were used for the RT model and the confidence model as RT and confidence are not on the same scale. We change the F parameter to S and the activation parameter from $-A(p)$ for RT to $A(p)$ given that RT decreases after an interruption and confidence increases.

Fourth, we can use the two models to provide evidence for decisional or post-decisional theories. In particular we can compare the *assoc* parameter in each model (RT and confidence). Decisional theories would predict that the *assoc* parameter for the RT model and the *assoc* parameter for the confidence model should be equivalent because the theories claim that both choice and confidence emerge at the same time. In contrast, post-decisional theories would predict that choice and confidence emerge at different times and in particular that confidence judgments occur after decision. Therefore, post-decisional theories would predict that the *assoc* parameter would be significantly smaller for the confidence model than the *assoc* parameter for the RT model.

Methods

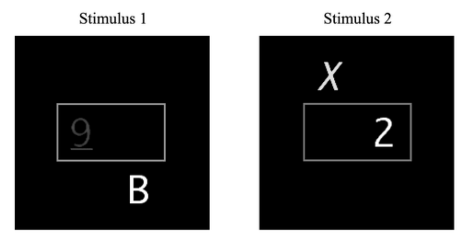
Participants

One hundred and fifty-five George Mason University undergraduates participated for course credit.

Task

Primary Task The UNRAVEL task was adapted from Altmann, Trafton, & Hambrick (2014). The UNRAVEL task has seven rules each represented by a letter (i.e. U, N, R, A, V, E, L). Participants are presented with one number and one letter at the same time. Each letter and number has certain characteristics that change from trial to trial such as color, font, position, etc. Participants are instructed to keep the UNRAVEL rule in memory, interpret what characteristic of the stimuli they are asked to identify, analyze the stimuli, and using the keyboard to submit what characteristic they identified (Figure 1). The UNRAVEL rules are available to the participant at anytime by holding the Shift + ? keys.

(a) Sample stimuli for UNRAVEL task:



(b) Choice rules and candidate responses for UNRAVEL task, and responses to stimuli in (a):

Step	Candidate responses		Choice rules	Responses to sample stimuli Stimulus 1	Stimulus 2
U	u	i	character is Underlined or in Italics	u	i
N	n	f	letter is Near to or Far from start of alphabet	n	f
R	r	y	character is Red or Yellow	r	y
A	a	b	character is Above or Below the box	b	a
V	v	c	letter is Vowel or Consonant	c	c
E	e	o	digit is Even or Odd	o	e
L	l	m	digit is Less than or More than 5	m	l

Figure 1. Example of the UNRAVEL task from Altmann et al. (2014).

For example, the U action in UNRAVEL prompts participants to identify if a number or letter is underlined or italicized. If a letter or number is underlined they press the “u” key on the board. If the letter is italicized they are instructed to press the “i”. After they submit their response participants will be presented with a brand new stimulus. They will search the stimulus for a characteristic prompted by the N action. The N action prompts participants to determine if the letter is near (“n”) or far (“f”) from the beginning of the alphabet. Participants continue to proceed through the UNRAVEL rules. Once completed, participants wrap around to the U action. The goal is to complete the rules in order and correctly identify the prompted characteristic for each stimulus.

Each action in UNRAVEL has a different set of keys associated with a response. As a result, the keystrokes reveal what action participants think they are on allowing for an analysis of sequence errors.

Interruption Task We used an equation to determine when participants were given a secondary task that served as an interruption. The equation can be found in Altmann et al., (2014) and resulted in an interruption 11.85% of the time.

After an UNRAVEL response was submitted, the UNRAVEL task was occluded and participants were asked to type in a series of letters into a box. Once the letters were typed in correctly the UNRAVEL task was revealed again. Participants were asked to return to the UNRAVEL task in the correct order.

Confidence Question Participants received a confidence question after completing an UNRAVEL action following half of the interruptions and an equal number of the control trials. The screen was replaced with a question that asked: "How confident were you that you just chose the correct step during the UNRAVEL task? Enter your choice on a scale from 1 to 6, with 1 being least confident and 6 being most confident." The participant typed in their response into a text field. After submitting their response the participant was returned to the UNRAVEL task

Procedure

Participants filled out an approved IRB consent form as well as biographical information. The task was first described using screenshots.

Participants were given a practice session where each rule of UNRAVEL was explained. They were exposed to all elements of the task including interruptions and confidence questions. Participants were shown that they could hit a certain key to access a list of the UNRAVEL rules at any time.

Results

One hundred and fifty-five participants completed 42442 UNRAVEL actions. We treated each action as a trial. There were 4925 confidence judgments. Only trials with confidence judgments were analyzed. We averaged RT and confidence for each participant for the first seven actions after an interruption.

Modeling the Time course of Recovery for Decisions RT

Not every participant had a confidence question at each step after the interruption. We used a linear mixed-effects model which can account for unbalanced repeated measures designs (Lindstrom & Bates, 1990) to look for differences in RT across step. There was a significant effect of step [$F(1, 723.61) = 69.16$, $p < .05$]. To investigate differences between steps, we compared Step 1 with Steps 2-7. Step1 after an interruption was significantly higher than Steps 2-7 [$F(1,154) = 171.9$, $MSE = 207.15$, $p < .05$]. This result replicates the disruptive effects of interruptions (Altmann et al., 2014; B. Edwards & Gronlund, 1998; Gillie & Broadbent, 1989).

Replicating Altmann & Trafton (2007) the response time for the primary judgment has a curvilinear pattern of recovery after an interruption [Overall: $F(6, 848) = 17.64$, $p < .05$; Linear: $t = -6.89$, $p < .05$; Quadratic: $t = 4.65$, $p < .05$]. Following Altmann & Trafton (2007), we fit the model to the data by estimating *F* and *assoc* for each participant. We used the mean decision RT for the first trial after the interruption for each participant as the *F* parameter. An RMSE was calculated for the *F* parameter while varying the *assoc* parameter between .005 and 1 by .005. The lowest RMSE for each fixed *F* and varied *assoc* parameter was set for each participant. The *F* and *assoc* were then averaged across participants to give us an *F* of 4.79 and an *assoc* of .29 for our model. The mean RMSE was 1.11 and R^2 was .37. Figure 2 shows the empirical data for choice RT for each UNRAVEL step after an interruption and predicted choice RT from our model. This replicates Altmann & Trafton (2007).

To test goodness of fit we ran runs tests (Bradley, 1968) on the signs of the deviations from the model minus the data. The runs test showed that the model and data were not significantly different from each other [$t(154) = 1.47$, $p = .14$].

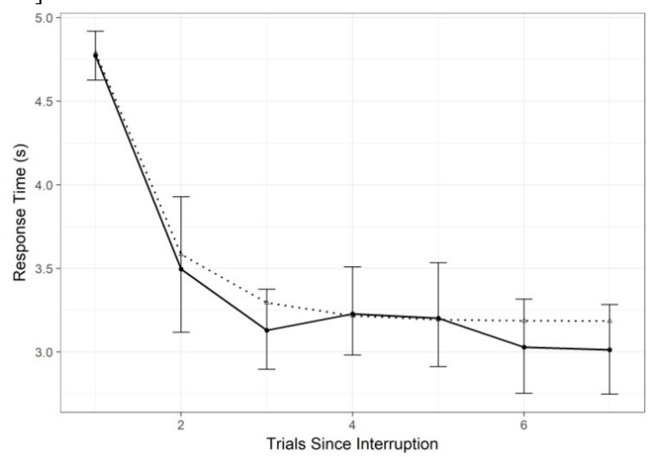


Figure 2. Data (solid) and model (dashed) for time course recovery of decision RT. Error bars are 95% confidence intervals data.

Modeling the Time course of Recovery for Confidence

There was a significant effect of step [$F(1, 701.85 = 35.80$, $p < .05$]. To investigate the differences between steps we compared Step 1 with Steps 2-7. Step 1 after an interruption was significantly lower than Steps 2-7 [$F(1,154) = 67.6$, $MSE = 11.32$, $p < .05$]. Similar to decision RT, confidence shows a pattern of recovery after an interruption [Overall: $F(6,848) = 3.89$, $p < .05$; Linear: $t = 3.05$, $p < .05$; Quadratic: $t = -2.37$, $p < .05$]. We used the same modeling process that Altmann & Trafton (2007) did and that was used above for the RT model. The *S* and *assoc* parameters were estimated for each participant and the parameters with the lowest RMSE were used for the final model. The *S* parameter was 4.95 and the *assoc* parameter was .08 for the final model. The mean RMSE was .41 and R^2 was .43. Figure 3 shows the empirical data for confidence for each UNRAVEL action after an interruption and predicted confidence from our model.

The goodness of fit for the runs test for the model and data showed no significant difference [t(154)= -.02, p = .98] suggesting that confidence is, indeed, tied to activation.

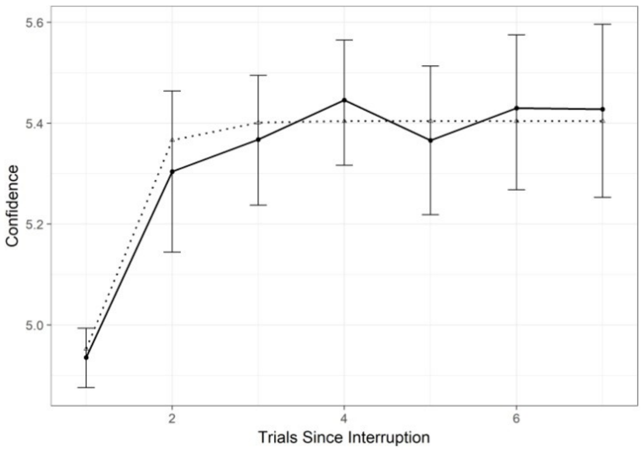


Figure 3. Data (solid) and model (dashed) for time course recovery of confidence. Error bars are 95% confidence intervals for data.

Comparing Activation for Choice RT and Confidence

To compare the timing of choice and confidence, we assume that the strength of a memory (activation) is the driver of both judgments.

Both of our models have an activation component that is represented by the *assoc* parameter. We took the *assoc* parameter from each participant's lowest RMSE model fit and compared *assoc* for choice RT and confidence using a within-subjects ANOVA. The *assoc* parameter was significantly lower for confidence (M=.08) than for choice RT (M=.29) [F(1,154) = 169.9, MSE = 3.56, p < .05, η2 =.31] .

Discussion

In this paper we built two models of complex cognitive processes: decision-making and confidence judgments. We instantiated decision-making using the model from Altmann & Trafton (2007) and modeling choice RT. We then built a model of confidence that also used an activation parameter so that we could compare the models. Our data and models produced five important findings.

First, we were able to replicate the empirical curvilinear pattern from Altmann & Trafton (2007) and show that RT recovers over time after an interruption.

Second, we built and replicated a model of RT for data on a new task.

Third, we showed that confidence is influenced by cumulative priming and that confidence recovers over time. This is a unique finding given that many models and experiments measuring confidence consider the contribution of memory only from the item just retrieved (DeSoto & Roediger, 2014; Merkle & Van Zandt, 2006; Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2009, 2013; Van Zandt & Maldonado-Molina, 2004; Douglas Vickers, 2014).

In this study we did not rely on the common list-learning or perceptual stimuli that have come to dominate the field. Instead we used a procedural task which allowed us to demonstrate that confidence judgments respond to priming from other elements of the task. Given that many of the current popular models of confidence consider confidence judgments a unitary process (Merkle & Van Zandt, 2006; Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2009, 2013; Van Zandt & Maldonado-Molina, 2004), our findings suggest that models of confidence should be able to account for carryover effects.

Fourth, we built a novel model for confidence judgments that explains the recovery of the confidence judgment process following an interruption. The model is driven by two parameters. The first is a scaling parameter which accounts for non-decisional processes. The second parameter is the amount of associative activation between elements. This second parameter is theoretically important because it suggests that in procedural tasks, confidence is sensitive to changes in activation.

Fifth, the comparison of the *assoc* parameter for choice RT and confidence strongly suggests that confidence happens after choice. Recall the predictions of the decisional and post-decisional theories of confidence. Decisional theories claim that confidence emerges as a result of the primary choice and that confidence is made available at the same time. Post-decisional models claim that additional information is collected about the primary choice and confidence is formed after a decision is made. Therefore, decisional models would predict that the *assoc* parameter used to model choice RT and confidence would be the same for both because they emerge at the same time. Post-decisional models predict that the activation in the *assoc* parameter would be lower for confidence than for choice RT because the confidence decision occurs after choice RT.

In our study we find that the *assoc* parameter is significantly less for confidence than it is for choice RT. This finding supports the post-decisional theories of confidence that say confidence is rendered after a decision is made. We believe that the *assoc* parameter is lower because the goal used to make the confidence judgment has undergone decay from when the same goal was measured earlier to make the primary choice.

Calculating activation through modeling has some unique benefits for measuring cognitive processes that are otherwise difficult to investigate empirically. For example, in Dotan et al., (2018) the authors offer an alternative explanation of their data. According to some views of confidence processing, their data could be interpreted as very fast and discrete post-processing. By this view, participants make decisions and confidence judgments several times before rendering a final judgment. However, it would be very difficult empirically to disentangle online confidence from rapid post-decision processing. As was the case in our study, modeling could be a helpful tool to help investigate such a hypothesis.

Acknowledgments

This work was supported by the Office of Naval Research to JGT. The views and conclusions contained in this document should not be interpreted as necessarily representing the official policies of the U. S. Navy.

References

Aguiar, N., Zish, K., McCurry, J. M., & Trafton, J. G. (2016). Interruptions Reduce Performance across All Levels of Signal Detection When Estimations of Confidence are Highest. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 60, pp. 254–258). SAGE Publications.

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review*, *14*(6), 1079–1084.

Altmann, E. M., Trafton, J. G., & Hambrick, D. Z. (2014). Momentary interruptions can derail the train of thought. *Journal of Experimental Psychology: General*, *143*(1), 215.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, *38*(4), 341–380.

Edwards M. B., & Gronlund, S. D. (1998). Task interruption and its effects on memory. *Memory*, *6*(6), 665–687.

Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(3), 929.

Bradley, J. V. (1968). *Distribution-free statistical tests*.

DeSoto, K. A., & Roediger, H. L. (2014). Positive and negative correlations between confidence and accuracy for the same events in recognition of categorized lists. *Psychological Science*.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition*, *24*(4), 523–533.

Dotan, D., Meyniel, F., & Dehaene, S. (2018). On-line confidence monitoring during decision making. *Cognition*, *171*, 112–121.

Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research*, *50*(4), 243–250.

Green, D. M., & Swets, J. A. (1966.). *Signal Detection Theory and Psychophysics*. New York City, New York: Wiley.

Heath, R. A. (1984). Random-walk and accumulator models of psychophysical discrimination: a critical evaluation. *Perception*, *13*(1), 57–65.

Juslin, P., & Olsson, H. (1997). Thurstonian and Brunswikian origins of uncertainty in judgment: a sampling model of confidence in sensory discrimination. *Psychological Review*, *104*(2), 344.

Laming, D. R. J. (1968). Information theory of choice-reaction times.

Lindstrom, M. J., & Bates, D. M. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, 673–687.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*(1), 77–105.

Merkle, E. C., & Van Zandt, T. (2006). An application of the poisson race model to confidence calibration. *Journal of Experimental Psychology: General*, *135*(3), 391.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864.

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

Ratcliff, R., & Starns, J. J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*(1), 59.

Ratcliff, R., & Starns, J. J. (2013). Modeling confidence judgments, response times, and multiple choices in decision making: recognition memory and motion discrimination. *Psychological Review*, *120*(3), 697.

Resulaj, A., Kiani, R., Wolpert, D. M., & Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature*, *461*(7261), 263.

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, *58*(5), 583–603.

Van Zandt, T., & Maldonado-Molina, M. M. (2004). Response reversals in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(6), 1147.

Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, *13*(1), 37–58.

Vickers, Douglas. (2001). *Where Does the Balance of Evidence Lie with Respect to Confidence?*

Vickers, Douglas. (2014). *Decision processes in visual perception*. Academic Press.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review*, *117*(4), 1025.

Zish, K., Hassanzadeh, S., McCurry, J. M., & Trafton, J. G. (2015). Interruptions can Change the Perceived Relationship between Accuracy and Confidence. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, pp. 230–234). SAGE Publications.