Summer 8-15-2017

Is There a Higher-Order Mechanism that Explains Performance Across Prediction Tasks?

Michelle Lisa Eisenberg
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Clinical Psychology Commons, and the Cognitive Psychology Commons

Recommended Citation

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Psychological and Brain Sciences

Dissertation Examination Committee:
Thomas Rodebaugh, Co-Chair
Jeffrey Zacks, Co-Chair
David Balota
Deanna Barch
Todd Braver
Larry Snyder

Is There a Higher-Order Mechanism that Explains Performance Across Prediction Tasks?
by
Michelle Eisenberg

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August, 2017
St. Louis, Missouri

Table of Contents

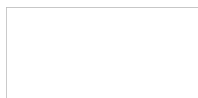# List of Figures

iv

List of Tables

v

<u>Acknowledgments</u>

ABSTRACT OF THE DISSERTATION

Is There a Higher-Order Mechanism that Explains Performance Across Prediction Tasks?

by

Michelle L. Eisenberg

Doctor of Philosophy in Psychological and Brain Sciences

Washington University in St. Louis

Professor Jeffrey Zacks, Chair

Professor Thomas Rodebaugh, Co-Chair

People constantly make predictions about what will happen in the near future. People anticipate how other people around them will act, what other people will say, and what actions will help them achieve the greatest rewards. Because all of these behaviors are typically called prediction, it is easy to make the assumption that performance across all of these types of tasks is driven by the same underlying mechanism. However, there has been little investigation into whether the mechanisms underlying prediction are the same across multiple task modalities. Therefore, in the current study, 226 participants completed four types of tasks that putatively involve prediction to determine whether there is a common factor that can account for performance on these tasks. Fluid and crystallized intelligence were also assessed to ensure that general intelligence did not drive correlations among the tasks. Preliminary evidence from a recent study suggested that people with Posttraumatic Stress Disorder (PTSD) have difficulty with predicting future activity; therefore, participants also completed a questionnaire screening for symptoms of PTSD. Performance across the four prediction tasks was not correlated, and PTSD severity was not significantly correlated with any of the tasks in the study. These results suggest that there is not an integrative prediction mechanism in the brain, but rather that there are multiple prediction

systems operating in parallel within the brain. In addition, these results suggest that PTSD may

only be associated with a subset, if any, of prediction tasks. Future researchers studying

prediction must be careful to investigate performance on various prediction tasks separately,

rather than assuming that prediction performance is stable across tasks.

Chapter 1: Introduction

The ability to anticipate what is going to happen in the near future is necessary for the survival of all animals. Prey animals must make predictions about the locations of their predators in order to avoid being eaten. Predators must anticipate the location of their prey so as not to starve. Humans, too, must constantly make predictions on timescales ranging from a fraction of a second to minutes, hours, and even days.

But, are all forms of prediction the same? In other words, do the same mechanisms that allow people to anticipate the next word in a sentence also allow people to predict the next action of the person in front of them, or whether it is likely to rain later in the day? For the past few decades, researchers have studied the neural basis of prediction to investigate how people and animals anticipate future input (e.g., Hikosaka, Sakamoto, & Usui, 1989; Rao & Ballard, 1999; Tanaka et al., 2004; Dikker & Pylkkänen, 2013). However, there has been little work examining whether the mechanisms underlying prediction are the same across multiple task modalities (see Adams, Friston, and Bastos (2015) for some evidence of a ubiquitous prediction system). The current study therefore examined four types of tasks that putatively involve prediction to determine whether there is a common factor of prediction ability that accounts for performance on these tasks. If a common mechanism for prediction exists, clinical populations with deficits in one domain of prediction would be expected to display similar deficits across other prediction tasks. Previous research on Posttraumatic Stress Disorder (PTSD) has found that people with higher levels of PTSD symptoms display difficulty with prediction of future human activity (Eisenberg, Zacks, Rodebaugh, & Flores, in prep). Therefore, this study also included a screening measure for symptoms of PTSD to determine whether this prediction deficit would generalize across all of the prediction tasks.

There is evidence that many areas of the brain are involved in making predictions, and the predictive coding theory, an influential model of prediction mechanisms in the brain, suggests that prediction occurs in a hierarchical process that is similar across the brain (Friston, 2005). In this model, higher-order cortical areas use past experience to make predictions about future inputs and send these predictions to lower-order areas, which compare actual input from the environment to these predictions. When there is a mismatch between the actual input and the predictions, the lower-order areas send prediction error signals back to the higher-order areas (Friston, 2005). Much research provides evidence of hierarchical signaling from higher-order brain areas to lower-order brain areas and vice versa, including findings on prediction mechanisms in sensory areas (see Bendixen, SanMiguel, & Schröger, 2012, for a review), reward processing areas (e.g., Schultz, Tremblay, & Hollerman, 1998; Tanaka, et al., 2004), and language processing areas (Dikker & Pylkkänen, 2013; Fruchter, Linzen, Westerlund, & Marantz, 2015). Although Friston's (2005) model does not require a single system for making predictions, Adams, Friston, and Bastos (2015) provide evidence that the neural circuitry for driving predictions is very similar across systems. They suggest that the individual systems should not be considered separate, but "instead as a single active inference machine that tries to predict its sensory input in all domains" (p. 100).

Other theories of information processing also posit the existence of a single integrative system that combines sensory information across modalities to generate predictions. For example, Event Segmentation Theory (EST; Zacks, Speer, Swallow, Braver, & Reynolds, 2007) suggests that people create representations of the current environment, called event models, based on incoming sensory information from various sensory modalities and semantic knowledge. Predictions about the near future are then formed on the basis of these event models,

and when errors in these predictions arise, the event model is reset to better represent the actual state of the world. Research using narrative texts (e.g., Zacks, Speer, & Reynolds, 2009), simple moving stimuli (e.g., Zacks, 2004), and complex activities (e.g., Zacks, Kurby, Eisenberg, & Haroutunian, 2011) has provided support for this model, suggesting that an event model may be one way of describing a higher-order integrative system that allows for predictions across modalities.

On the other hand, it is possible that rather than relying on an integrative prediction mechanism, each brain system uses a separate prediction mechanism that operates using only the information present in each brain system. For example, the visual system might have a prediction mechanism that operates using only visual information from the world and previous visual experience. As mentioned above, the predictive coding theory does not require an integrative prediction system, as Friston (2005) primarily argues that predictions operate in a hierarchical fashion within each brain system. Therefore, it is possible that multimodal stimuli (e.g., visual scenes that also involves music and speaking) activate multiple brain systems in concert, resulting in the illusion of integration without involving an integratory prediction mechanism. Because there is little research attempting to differentiate between an integratory prediction mechanism versus separate prediction mechanisms within each brain system, the current study represents an initial step toward determining whether such an integratory prediction mechanism actually exists.

Previous research and theorizing on prediction, including research on the predictive coding theory and EST, are based, in part, in a large literature on the neural mechanisms involved in creating and maintaining predictions in the brain. As might be expected, an examination of this literature provides evidence both for and against a higher-order prediction mechanism that

3

integrates information across modalities. In particular, research on the neural mechanisms involved in creating and maintaining predictions in the sensory systems, reward processing systems, and language processing systems provides evidence both for and against such an integrative mechanism. In addition, error signaling from lower-order areas to higher-order areas is essential in models of a predictive brain, and research findings on error signaling also suggest evidence supporting and opposing a common prediction mechanism. Therefore, each of these topics is discussed in detail in this introduction, followed by a description of the tasks used in the current study.

1.1 Prediction Formation

According to the predictive coding model (e.g., Friston, 2005), prediction formation should occur within both lower-order and higher-order systems, with the higher-order systems capable of integrating information across modalities to form adaptive predictions about future input and then communicating these predictions to lower-order areas. Therefore, in this section, specific attention is given to evidence that higher-order areas are involved in the prediction process across modalities.

1.1.1   Sensory Systems

Within the sensory systems, the main focus of research has been on auditory and visual prediction mechanisms. One method of studying prediction formation in the auditory systems is to investigate the effects of repeating an auditory tone or a pattern of auditory tones a varying number of times, as predictions would be expected to increase in strength as the number of repetitions increases. Haenschel, Vernon, Dwivedi, Gruzelier, & Baldewag (2005), for example, investigated event related potentials (ERPs) while participants listened to series of tones that were presented 2, 6, or 36 times. They found a positivity that began 50 to 250 ms after a tone

4

was presented and that increased in strength as the number of repetitions increased, and they suggested that this repetition positivity is associated with the formation of a sensory memory representation of the repeated tone. Bendixen, SanMiguel, and Schröger (2012) took this finding a step farther and suggested that the sensory memory representation is used to predict future tones. They interpreted the repetition positivity as the signal that occurs when a predicted stimulus matches the actual stimulus.

To further determine how the brain represents expectations of future stimuli, Raij, McEvoy, Mäkelä, and Hari (1997) investigated the brain response to omitted tones. They used magnetoencephalography (MEG) while participants listened to repeated tones. Seven percent of these tones were randomly omitted, and the authors found bilateral activation in the auditory cortex, particularly the supratemporal cortex, when the tones were omitted. They argued that this activation represents the buildup of an expectation of the tone and a signal indicating that this expectation was not fulfilled. Mustovic et al. (2003) conducted a similar study using functional magnetic resonance imaging (fMRI). They had subjects listen to repeated patterns of sounds and interspersed a short period of deviant louder sound or a short period of silence into the pattern. Increased activity in the bilateral posterior secondary and association auditory cortices, right Heschl's gyrus, and right planum temporale occurred during both types of deviant periods. In addition, greater activity during the silent period than the louder period was seen in the right planum temporale and part of the right temporoparietal junction, suggesting that these regions are involved in retrieving auditory memory traces for predicted (but absent) stimuli.

Research on the visual system has found similar prediction mechanisms in the visual cortex. For example, Luft, Meeson, Welchman, & Kourtzi (2015) used fMRI and multi-voxel pattern analysis to examine predictions in primary visual cortex. They had participants view a

sequence of gratings with different orientations. After participants learned the sequence, they were able to detect patterns of activation in the primary visual cortex representing the participants' predictions of the orientation of the next grating, providing evidence that the primary visual cortex maintains predictions about future visual stimuli. In another study using multi-voxel pattern analysis, participants viewed gratings with different orientations and heard auditory cues that provided information about the orientation of the next grating stimulus (Kok, Jehee, & Lange, 2012). The authors found that top down expectations driven by the auditory cue sharpened the representation of the predicted orientation in early visual cortex. Specifically, they found that expectation of a particular orientation dampened the overall response of the early visual cortex to the stimulus while simultaneously making it easier for the classifier to predict the behavioral response of the participant. These results suggest that higher-order cortical regions send top-down predictive signals to early visual cortex that bias the response in that area and facilitate performance on the task.

Trapp and Bar (2015) posited a model of top-down and bottom-up predictive processing in the visual system, suggesting that the orbitofrontal cortex (OFC) is involved in creating predictions that bias processing of visual stimuli based on context. Specifically, their model suggests that early visual areas send information at a low spatial frequency to the OFC, which uses that information and prior knowledge of the context to make predictions about the identity of the most likely input. These predictions bias the analyses performed by the visual areas toward the relevant options, which is consistent with predictive coding theory.

1.1.2   *Reward Processing Systems*

Predictions related to reward processing appear to be generated in the striatum, which includes the caudate and putamen (see Schultz, Tremblay, & Hollerman, 1998, for a review). For

example, populations of neurons in the caudate and putamen of behaving monkeys increase and maintain their firing rate during a delay before an expected target appears and during a delay before a reward is dispensed, suggesting that these neurons signal the expectation of a reward (Hikosaka, Sakamoto, & Usui, 1989; Apicella, Scarnati, Ljungberg, & Schultz, 1992).

fMRI studies on reward prediction in humans have supported these animal findings. Tanaka, et al. (2004), for example, found that when participants learned a task involving immediate rewards, activity increased in the striatum, insula, and the lateral OFC, among other areas. In addition, when participants needed to maintain a representation of the reward structure in order to obtain future rewards, activity increased in the striatum, insula, ventrolateral PFC, the dorsolateral PFC, and other areas. In another fMRI study, activity in the striatum increased when participants saw cues that predicted rewards compared to when they saw cues that did not predict rewards (Ramnani, Elliott, Athwal, & Passingham, 2004). Providing further support for the involvement of the striatum in prediction, Ernst et al. (2004) found activation in the ventral striatum during the period right before participants received a reward. They also found activity in the left lateral and medial OFC and left insula (among other areas) during this period of reward anticipation. On the other hand, a meta-analysis of 142 studies of reward processing found that the bilateral insula, anterior cingulate cortex, inferior parietal lobule, and brain stem displayed activation related to anticipation of rewards, whereas the ventral striatum, medial OFC, and amygdala displayed increased activation during reward outcome stages (Liu, Hairston, Schrier, & Fan, 2011).

Although the lower-order brain areas found to be involved in reward prediction tasks are different than those involved in visual and auditory prediction tasks, the tasks are similar in that higher-order brain areas are recruited during the performance of all of the tasks. These findings

again provide support for Friston's predictive coding theory and suggest that higher-order areas may be necessary for integration of predictive information across modalities.

1.1.3   *Language Processing System*

The neural basis for prediction formation in the language processing systems has also received much attention. For example, Dikker & Pylkkänen (2013) used MEG while participants viewed pictures that were either strongly predictive or weakly predictive. After viewing each picture, participants saw a word that either matched or did not match the prediction generated by the picture and indicated whether the word was a match or a mismatch for the preceding picture. For example, in a predictive trial, participants might see a picture of an apple followed by the word "apple," whereas in a weakly predictive trial, participants might see a picture of a grocery bag (which could represent any type of edible object) followed by the word "apple." During the predictive trials compared to the weakly predictive trials, they found increased activity in the mid-temporal cortex and the ventromedial PFC around 350 ms before the onset of the noun and increased activity in the occipital lobe right before the noun was presented. The authors suggested that these results represent a predictive feedback process from higher- to lower-order cortical regions. Specifically, they posited that the activity in the visual cortex right before the noun was presented represented the preactivation of features associated with the predicted noun, that the activity in the mid-temporal cortex represented the preactivation of the predicted lexical representation of the noun, and that the activity in the ventromedial PFC represented the combination of lexical and semantic representations into a prediction of the future input. They further argued that the activation in the visual cortex corresponded to top-down activation of relevant features and the suppression of irrelevant features in response to the previously presented image.

Fruchter, Linzen, Westerlund, and Marantz (2015) found similar results when participants read adjective-noun phrases. They found increased activity in the left middle temproral gyrus during the time after the presentation of predictive compared to unpredictive adjectives but before the noun was presented. In addition, they found decreased activity in the left middle temporal gyrus when predictable nouns were presented, suggesting that activity in this area decreased once predictions were fulfilled. The results of these studies again suggest a hierarchical prediction system in which higher-order areas form predictions and then communicate these predictions to other areas of the brain.

1.2 Human Electroencephalographic Studies of Prediction Error

At the same time as the higher-order brain areas form predictions and communicate these predictions to lower-order areas, the lower-order areas must send signals to the higher-order areas when these predictions are incorrect. Most studies on prediction error in humans have used electroencephalography (EEG) to study event related potentials in the brain, as prediction error signals in the brain emerge very quickly and EEG is capable of measuring these signals as they occur. The most commonly reported error signals are the error related negativity (ERN), mismatch negativity (MMN), P300, N400, and P600. Although most of these error signals are elicited by stimuli in multiple sensory modalities, suggesting a common eliciting mechanism, the existence of multiple error signals suggests that different mechanisms may drive the detection of error. Findings related to these error signals are therefore informative for determining whether there is a higher order prediction mechanism that integrates information across modalities.

1.2.1 Error Related Negativity (ERN)

The ERN was first reported in two studies that presented a series of stimuli either visually or auditorily. When participants made an error in their response, a negativity with a fronto-central

maximum was observed 0 to 100 ms after the error (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993). Since then, the ERN has been the subject of numerous studies and has been found in tasks of various difficulty levels and response modalities (for a review, see Weinberg, Dieterich, & Riesel, 2015). In particular, the ERN has been implicated in studies of reward processing (e.g., Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003), and there is strong evidence that the ERN is generated by the anterior cingulate cortex (ACC; e.g., Gehring, Goss, Coles, Meyer, & Donchin, 1993; Dehaene, Posner, & Tucker, 1994; Ullsperger & von Cramon, 2001), which is thought to integrate pain/punishment and gain/reward to drive behavior (Weinberg, Dieterich, & Riesel, 2015). The ACC is thought to integrate signals from multiple sensory systems, and it has dense connections to the prefrontal cortex and the midbrain dopamine system, suggesting that error signals from the dopamine system in response to a reward or loss may be sent to the ACC and then to the prefrontal cortex for further processing (for a review, see Weinberg, Dieterich, & Riesel, 2015).

1.2.2   *Mismatch Negativity (MMN)*

The mismatch negativity (MMN) signal has repeatedly been found in auditory tasks in which rare sounds are inserted into a sequence of repeated sounds. In these tasks, the MMN signal begins around 100-250 ms after the rare deviant sound and is localized in the bilateral auditory cortices and in the right frontal cortex (see Kujala, Tervaniemi, and Schröger, 2007, for a review). The MMN has also been found in tasks involving much more complex patterns of which participants are not consciously aware, such as a rule that short tones must be followed by low tones and long tones by high tones (Paavilainen, Arajärvi, & Takegata, 2007). There is some evidence that the MMN begins in the bilateral auditory cortices and later is generated by the right

frontal cortex (Rinne, Alho, Ilmoniemi, Virtanen, & Näätänen, 2000), suggesting that the error signal may be sent from lower- to higher-order brain areas.

The MMN also has been elicited in tasks involving other modalities, including visual processing. For example, Tales, Newton, Troscianko, and Butler (1999) presented a series of visual stimuli in the peripheral visual field and interspersed rare deviant stimuli. They found that the deviant stimuli elicited a negativity that began 250-400 ms after the stimulus and that appeared to be generated by supplemental visual areas in the occipital lobe and posterior temporal cortex. They suggested that this negativity is similar to the MMN found in response to deviant tones in studies of auditory processing. Czigler, Balázs, and Pató (2004) found a similar negativity that began 140-200 ms after the presentation of a deviant visual stimulus. Providing further support that this signal is analogous to the auditory MMN, Wei, Chan, and Luo (2002) used tasks that required both visual and auditory processing. Participants viewed a series of repeated and rare deviant stimuli while at the same time listening to a series of repeated and deviant auditory stimuli. In the first block of the task, the participants were instructed to attend to the visual stimuli, and in the second block, they were instructed to attend to the auditory stimuli. They found a negativity beginning 100-200 ms after the deviant stimuli regardless of modality. When participants attended to the auditory stimuli, the negativity was greatest in the temporal lobe, whereas when participants attended to the visual stimuli, the negativity was greatest in the occipital lobe. Furthermore, 200-250 ms after deviant visual and auditory stimuli, a negativity was observed in frontal regions, suggesting the presence of a feed-forward error signal that integrated the error signals from both modalities.

1.2.3   P300

The P300 is another signal that is often elicited in tasks involving unpredictable or surprising auditory or visual stimuli (for a review, see Polich, 2007). For example, Pollich and Margala (1997) presented participants with a series of repeated and deviant auditory stimuli and found a positivity approximately 300 ms after the presentation of the deviant stimuli. They found the same positivity when rare target tones were presented within periods of silence. A very similar response has been observed in tasks involving repeated and rare deviant visual stimuli (e.g., Bledowski et al., 2004; Bledowski, Prvulovic, Goebel, Zanella, & Linden, 2004). Although there is not a strong consensus about the neural origins of the P300, lesion studies suggest that frontal lobe and temporal-parietal junction integrity are necessary for the generation of the P300, suggesting that these areas are likely candidates (Pollich, 2007).

The P300 is actually thought to consist of two distinct signals: the P3a and the P3b. The P3a is typically observed when participants passively listen to infrequent tones or view visual stimuli that are embedded within a series of repeated stimuli, whereas the P3b is elicited when participants must overtly respond to infrequent stimuli that are presented within a series of repeated stimuli (for a review, see Pollich, 2007). Pollich (2012) suggests that the P3a is driven by attentional processing of novel or unexpected stimuli in the frontal lobe and that the P3b is generated when memory storage in temporal-parietal areas is accessed in the service of performing a discrimination task. This suggests that the P3a may signal the presence of an error in prediction and that, in response to that error, memory stores are accessed to allow for better predictions on future trials, generating the P3b.

1.2.4   N400

The N400 has been observed mainly in tasks involving language processing, particularly tasks in which stimuli do not match the preceding context. It was first reported in a study in which participants read sentences with either semantically congruent or incongruent final words (Kutas & Hillyard, 1980). The authors found a negative deflection between 300-600 ms after participants read the incongruent words, suggesting that the N400 represents a response to semantic errors. Since then, the N400 has been observed in response to many types of semantically incongruent stimuli, including to sentences that do not match the context of the preceding paragraph (e.g., van Berkum, Brown, & Hagoort, 1999). In addition, van Berkum, Brown, Zwisterlood, Kooijman, and Hagoort (2005) found an N400 in response to words that were unlikely given the grammatical structure of the sentence. They had participants listen to sentences in Dutch, which uses gendered suffixes on adjectives based on the gender of upcoming nouns in noun phrases. The authors recorded event related potentials (ERPs) after a gendered adjective was presented but before the associated noun was presented. Participants heard sentences such as "The burglar had no trouble locating the secret family safe. Of course it was situated behind a big, but unobtrusive painting." In this sentence, the adjective "big" had a neuter gender suffix, which was consistent with the gender of the word "painting." On the other hand, the sentence, "Of course, it was situated behind a big, but unobtrusive bookcase," was inconsistent because in this sentence the adjective "big" had an inconsistent gender when in the same sentence as the word "bookcase." The authors found a very large N400 response at the time that the prediction-inconsistent nouns were presented. The authors suggested that this ERP response meant that people predicted upcoming words based on the structure of current and preceding sentences.

The N400 has been localized to a source in the anterior medial temporal lobe, in middle and superior temporal areas, inferior temporal areas, and prefrontal areas, including the dorsolateral frontal cortex (for a review, see Kutas & Federmeier, 2011). Consistent with the predictive coding theory, Kutas and Federmeier (2011) suggested that the N400 actually consists of a wave of feed-back activity in response to an unexpected word: The activity begins in the left posterior temporal gyrus at 250 ms after word presentation, then spreads to more forward and ventral areas in the temporal lobe by 350 ms, and finally spreads to the right anterior temporal lobe and to the bilateral frontal lobes by 370-500 ms.

1.2.5   P600

The P600 is similar to the N400, in that both have been studied mainly during language processing; however, the P600 is observed when syntactic structure is violated, whereas the N400 is observed when semantic context is unexpected (for a review, see Swaab, Ledoux, Camblin, & Boudewyn, 2012). The P600 was first reported in a study in which participants read sentences that either conformed to expected syntactic structure (e.g., "The broker planned to conceal the transaction.") or violated expected syntactic structure (e.g., "The broker persuaded to buy the stock."; Osterhout & Holcomb, 1993). The authors found a slow positive signal around 600 ms following the word "to" in the sentences that violated syntactic structure. The P600 has also been observed for other types of syntactic violations, including gender and case marking violations (e.g., Coulson, King, & Cutas, 1998) and verb tense violations (e.g., Osterhout & Nicol, 1999). However, some recent research has led to questions about whether the P600 is elicited only in response to syntactic violations, as it has also been observed in response to certain types of semantic violations (for a review, see Swaab, Ledoux, Camblin, & Boudewyn,

2012). Overall, the evidence suggests that the P600 represents a prediction error response to stimuli that violate predictions about syntactic, and possibly semantic, structure.

There has not been much research attempting to localize the P600 to specific brain areas. In one of the few studies discussing the neural generator of the P600, Service, Helenius, Maury, & Salmelin, (2007) used MEG while participants read a series of sentences, some of which violated syntactic rules and some of which violated semantic rules. They found evidence for both the P600 and the N400, and they localized the P600 to the superior temporal cortex, posterior to the generator of the N400 response. Brouwer and Hoeks (2013) disagreed with this finding, and used existing evidence from neuroimaging studies of language processing to suggest that the P600 originates from the left inferior frontal gyrus. However, although they provided recommendations for future studies meant to support their hypothesis, they did not collect any data localizing the P600 to the inferior frontal gyrus. It is therefore difficult to determine whether the neural mechanisms underlying the P600 conform to the predictions of predictive coding theory.

1.3 Interim Summary

Evidence on prediction formation and prediction error signaling strongly suggests that predictive processes occur throughout the brain, from lower-order sensory areas to higher-order cortical areas that include the prefrontal cortex. However, the literature provides evidence both for and against the hypothesis that there is a higher-order prediction mechanism that drives performance across tasks. On one hand, there seem to be similar higher-order brain areas, primarily in the frontal cortex, that are activated across tasks requiring prediction formation. In addition, many of the error signals are elicited by stimuli in multiple modalities. For example, the P300 is sensitive to both simple tone sequences and to complex semantics and the N400

responds to sentence-level incongruity as well as situation-level incongruity. On the other hand, the fact that different lower-order brain regions are activated during prediction formation depending on the task and that there are so many different prediction error responses suggests that there may not be a single integrative prediction mechanism. These two possibilities have very different implications for individual differences in performance across prediction tasks. If there is a higher-order integrative prediction mechanism that drives performance across tasks, prediction performance across tasks should be highly correlated. Conversely, if predictions are generated separately within each neural system, prediction performance across tasks might not be highly correlated due to individual strengths and weaknesses within specific modalities.

Discriminating between these alternate possibilities requires the use of tasks that are likely to require multiple modalities, as tasks that fall only within a specific modality might not require a higher-order integrator. Fortunately, although prediction formation and error signaling are often studied separately in different modalities, most real-world tasks do not involve single modalities, and even the laboratory tasks discussed above rarely require only a single sensory system. For example, the reward prediction literature often uses visual cues to signal upcoming rewards (e.g., Ramnani, Elliott, Athwal, & Passingham, 2004; Ernst et al., 2004), which means that prediction mechanisms in the visual system and the reward processing system must be active at the same time. In fact, in the real world, it is common for people to employ the visual, auditory, language, and reward processing systems simultaneously. This strongly suggests the presence of a higher-order system that integrates information across modalities in order to make adaptive predictions, and given the evidence discussed in the previous sections, this higher-order system likely resides in the frontal cortex. However, there is not enough evidence to support the delineation of specific areas within the frontal cortex as multimodal integrators and predictors.

The tasks used in the current study were chosen because they require prediction across multiple modalities and were therefore likely to involve a higher-order cortical prediction system. Specifically, the current study used two types of predictive looking tasks, which use eye tracking to determine whether participants are making predictions about future input, a probabilistic classification task, in which participants use cues to predict which of two outcomes will occur, and a gambling task, in which participants predict which choices will lead to the highest rewards. I hypothesized that if a higher-order integrative prediction mechanism existed, there would be high correlations across these tasks. On the other hand, I hypothesized that if predictions were generated separately by modality specific brain systems, correlations across these tasks would be low due to individual strengths and weaknesses in different modalities.

While there have not been any previous studies that have investigated whether performance is correlated across the four tasks included in this study, it is informative to examine previous research on the similarities and differences in the brain regions and systems activated during performance of these tasks. The research literature on these tasks provides evidence both for and against the hypothesized integrative prediction mechanism, and the following sections therefore discuss prior research on each of these tasks, with an emphasis on the neural mechanisms involved in their performance (where this literature exists).

1.4 Predictive Looking Tasks

There are two main types of predictive looking tasks: non-verbal predictive looking tasks and language predictive looking tasks. In non-verbal predictive looking tasks, participants complete short sequences of actions or watch short movies while their eyes are tracked using an eye tracker. Researchers are typically interested in whether participants look at objects before they are acted upon (e.g., how early participants look at a bowl before the actor picks it up), and

these tasks require integration of visual and motor information to make predictions. In language predictive looking tasks, participants typically view static images of objects while listening to sentences, and researchers are interested in whether participants look at objects before they are mentioned in the sentences. These tasks require auditory, visual, and language processing to achieve accurate predictions.

1.4.1   *Non-Verbal Predictive Looking Tasks*

Non-verbal predictive looking has been studied in ages ranging from infants to adults. In one study, six-, eight-, twelve-, fourteen-, and sixteen-month-old infants were shown short movies of a person interacting with a common object (Hunnius & Bekkering, 2010). In these movies, the objects were either brought to a correct or incorrect location. For example, in one movie a cup was brought to a person's mouth, while in another movie, a cup was brought to a person's ear. The authors found that infants were more likely to display anticipatory looking to the target location when the object and target locations were congruent than when they were incongruent. This study suggests that infants as young as six-months-old are capable of making predictions about objects and object-related goals. Cannon and Woodward (2012) also studied predictive looking in 11-month-old infants. They showed infants movies of a hand making repeated reaching movements toward one of two objects. Then, the locations of the objects were switched. Infants were more likely to predictively look at the original object rather than at the original location, suggesting that they were predicting that the actor would continue interacting with the same object rather than simply reacting to the actor's motion. Falck-Ytter, Gredebäck, and von Hofsten (2006) found similar results when they had 12-month-old infants watch a movie of an actor placing three toys in a bucket. Infants displayed reliable predictive eye movements to the

bucket before the toys contacted the bucket. These studies suggest that infants are capable of making goal-directed predictions about future action.

Studies in adults have also examined goal-directed predictive looking while participants carried out an action themselves. For example, Land, Mennie, and Rusted (1999) used eye tracking while participants made tea. A head mounted video camera and a second video camera located across the room were used to obtain fixation location. The authors found that participants first fixated on an object an average of .56 seconds before touching the object and that participants fixated the next object an average of .61 seconds before finishing their use of the previous object. In a very similar study, participants made a sandwich while their gaze location was tracked using an eye tracker. The authors found that 30% of the reaches to objects were preceded by a fixation to that object within the previous eight seconds (Hayhoe, Shrivastava, Mruczek, & Pelz, 2003). In another study, participants grasped a bar and moved it around an obstacle toward a target. Participants looked at the grasp site on the bar and at the target before making contact with the bar and the target. In addition, participants stopped fixating these objects after contact was made (Johansson, Westling, Bäckström, & Flanagan, 2001).

Not only do adults engage in predictive looking when performing a task themselves, but they also perform goal-directed predictive looking when watching someone else complete a task. Flanagan and Johansson (2003) had participants both stack blocks themselves and watch an actor stack the blocks in the same manner. The authors found that in both the passive and the active trials, almost all fixations were directed to the sites of contact on the blocks and the locations where the blocks were to be set down. In addition, in both types of trials, participants' fixations occurred an average of 150 ms before contact actually occurred. In another study, Elsner, Falck-Ytter, & Gredebäck (2012) created 12 s movies of point-light displays by attaching markers to a

hand moving laterally in space. They created two motion conditions: a biological motion condition in which the hand moved naturally and a non-biological motion condition in which the hand moved at a constant velocity. In both movies the hand moved toward and contacted a target object that was partially occluded by a barrier. In a between-subjects design, participants watched either the biological or the non-biological motion movies ten times while their eyes were tracked using an eye tracker. The authors found that participants in the biological motion condition looked at the target object an average of 124 ms before contact occurred. On the other hand, participants in the non-biological motion condition looked at the target object an average of 21.5 ms after contact occurred, which the authors stated constitutes reactive, rather than predictive, looking. This study suggests that people are equipped to be able to predict future biological motion, whereas people are less able to predict non-biological motion. From an evolutionary perspective, this would make sense, as living organisms are much more likely to move around on a regular basis that non-living objects.

Building on this literature, Eisenberg, Zacks, and Flores (in prep) developed a novel paradigm, called the Predictive Looking at Action Task (PLAT) to assay predictive looking while participants watch videos of actors performing everyday activities. Unlike the studies discussed above that used short videos of an actor interacting with only one object, the PLAT uses movies during which an actor interacts with many objects sequentially. In preliminary work, twenty-five participants passively watched three five-to-six minute long videos of an actor completing an everyday activity (e.g., making breakfast, preparing for a party). For each movie, the points at which the actor came into contact with a new object were identified and 500 ms bins were created for the three seconds before each point of contact. Both the proportion of participants who looked at the target object and participants' fixation time on the target object

increased as time to contact approached (See Figure 1). These results suggest that the PLAT can provide an online measure of prediction ability. This same pattern of results was replicated in another study of twenty-eight participants. Additional analyses using these two data sets also found that predictive looking decreased around event boundaries, suggesting that participants formed event models while viewing these movies, and that they updated their event models when predictions became more difficult (Eisenberg & Zacks, in preparation).
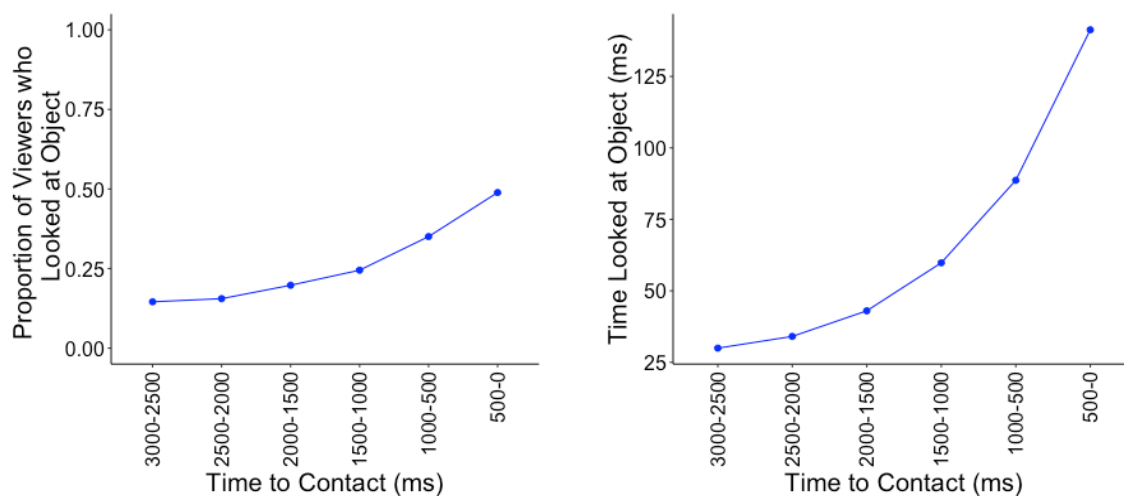
Figure 1. Results from Eisenberg, Zacks, and Flores (in prep) for the PLAT. The figure on the left displays the proportion of participants who looked at the target object during each of the six 500 ms bins. The figure on the right displays the amount of time participants looked at the target object during each of the six bins. For both figures, the time bins progress in time from left to right from 3000-2500 ms before the actor contacted the target object to 500-0 ms before the actor contacted the target object

1.4.2   Language Predictive Looking Tasks

Language-related predictive looking tasks typically use the visual world paradigm to investigate anticipatory language processing in adults. In the *visual world* paradigm, participants view an array of objects on a computer screen while listening to a sentence. An eye tracker is used to determine the point at which participants begin looking at the next object to be mentioned in the sentence. In one of the earliest studies investigating anticipatory eye

movements while participants looked at an array of images, Altmann and Kamide (1999) had participants listen to sentences such as "The boy will move the cake" or "The boy will eat the cake" while viewing a collection of images that, in this example, included depictions of a boy, a cake, and various toys. In one version of each sentence, the verb could only apply to one of the images (in this example, only the cake could be eaten), and in the other version, the verb could apply to all of the objects (in this example, all of the objects could be moved). The authors found that participants began looking at the cake much earlier when the sentence included the word "eat" than when the sentence included the word "move." In a similar study, Kamide, Altmann, and Haywood (2003) had participants listen to sentences such as "The woman will spread the butter on the bread." The authors found that participants looked at the goal object (the bread, in this case), immediately after they heard the referring expression (spread the butter). In another study, Altmann and Kamide (2007) used a similar paradigm but varied the tense of the verbs in the sentences. For example, participants heard sentences such as, "the man will drink" or "the man has drunk" while viewing a screen with both a full glass of beer and an empty wine glass. Participants looked more often at the object that matched the tense of the verb in the sentence, even before the target object was mentioned. These studies suggest that people engage in predictive looking during language processing as well as when they view short movie clips of human actions.

1.5 Probabilistic Classification Tasks

In probabilistic classification tasks, participants are asked to classify stimuli into two or more categories. Participants usually receive feedback after each trial, allowing them to learn to predict the category of subsequent items. Although these tasks are typically used to measure executive control, rather than prediction ability, these tasks require participants to learn

information over a series of trials and use this information to make subsequent predictions. There are a variety of probabilistic classification tasks that require participants to use previously learned information to respond on subsequent trials, including the weather prediction task and the Mr. Potato Head task.

In the weather prediction task (Knowlton, Squire, & Gluck, 1994), participants see various combinations of four cards with simple geometric designs and are asked to predict, on the basis of these cards, whether there will be rain or sun. The four cards are associated with 75%, 57%, 43%, or 25% probability with one of the outcomes. Participants receive feedback after each trial. Using this task, Knowlton, Squire, and Gluck (1994) found that after 50 trials, healthy participants performed above chance, choosing the optimal answer on an average of 68.2% of the trials. After 350 trials, healthy participants chose the optimal answer on an average of 74% of the trials. Gluck, Shohamy, and Myers (2002) found similar results, finding that after 200 trials, participants chose the optimal answer on an average of over 70% of the trials.

In a very similar paradigm, Shohamy et al. (2004) created probabilistically predictive stimuli using a Mr. Potato Head doll. In this study, four features of the Mr. Potato Head doll could vary, and participants had to use this information to predict whether each Mr. Potato Head customer at an ice cream shop wanted vanilla or chocolate ice cream. They found that healthy adult participants made optimal predictions on approximately 80% of the trials, which is consistent with findings from the weather prediction task. Aron et al. (2004) found almost identical results in another sample of participants using the same Mr. Potato Head task, with participants making optimal predictions on an average of approximately 70-80% of trials.

In addition to this behavioral replication, Aron et al. (2004) used fMRI to determine the brain areas involved in making these predictions. Specifically, they broke the trials down into

three phases: stimulus, delay, and feedback. Of these phases, activation during the delay was most relevant to prediction, as it was during this time that participants most likely made their predictions about ice cream flavor. The authors found significant activation in the right inferior frontal cortex, caudate nucleus, parietal cortex, and cerebellum during this delay. In addition, they found significant deactivation in the medial prefrontal cortex, medial temporal cortex, and parietal cortex. The authors then correlated neural activity with the degree of uncertainty on each trial. They found a significant positive correlation between activity in a region of interest in the midbrain (centered on the substantia nigra) and increasing uncertainty during the delay period. The authors therefore suggested that this midbrain region codes for uncertainty when people make predictions. Finally, the authors examined the functional connectivity of this midbrain region with the rest of the brain, and found significant correlations between activity in this region and ventral striatum, orbitofrontal cortex, and dorsomedial frontal cortex, suggesting that feed-forward and feed-back connections drive performance on this task.

In another type of probabilistic classification task, participants viewed a series of eight rapidly presented circles and triangles. Participants then predicted whether the next stimulus was likely to be a circle or a triangle. The amount of uncertainty on each trial varied depending on how many circles and triangles were presented during the stimulus presentation phase. For example, if all of the stimuli were circles, there was an 80% probability that the next stimulus would be a circle as well. On the other hand, when there was an equal number of each type of stimulus presented during the stimulus presentation phase, there was a 50% probability that the next stimulus would be a circle. The authors found that even though participants were never told about these different probabilities, participants quickly began to use the probabilities to help

them make their predictions; as uncertainty decreased, participants made more confident and correct decisions (Huettel, Song, & McCarthy, 2005).

In the same study, Huettel, Song, and McCarthy (2005) used fMRI to examine the brain regions activated while participants performed this task. They found that bilateral insula, inferior frontal gyrus, and intraparietal sulcus, along with right thalamus, and right inferior parietal lobule displayed a significant increase in activation as uncertainty increased. In addition, the authors analyzed the data to determine whether the order in which stimuli were presented affected the neural response. They found that when a stimulus that was incongruent with the preceding stimuli was presented late in a trial, the posterior parietal cortex, specifically, the intraparietal sulcus, displayed significantly greater activation compared to trials in which an incongruent stimulus was presented early in the trial. The authors suggested that the activation in this region represents the neural correlates of the attempted resolution of uncertainty and, therefore, the formation of a prediction about the target stimulus. In addition, based on previous research, the authors suggested that while activity in the posterior parietal cortex reflects uncertainty about which behavior to choose and the ultimate resolution of this uncertainty, the activity they observed in the anterior insula is likely related to uncertainty about future reward outcomes.

1.6 Gambling Tasks

In the Iowa gambling task, participants are shown four decks of cards, are given a starting amount of money, and are told to choose cards such that they make the most money and lose the least money. When participants choose to turn over a card from the A or B decks, they usually earn $100. When they choose to turn over a card from the C or D decks, participants usually earn $50. However, every once in a while, turning over a card results in a penalty, with a larger penalty associated with the A and B decks. This penalty occurs randomly, meaning that

participants have no way of knowing when they will incur this penalty. Because of the high penalties associated with the A and B decks, choosing cards from decks C and D results in the higher scores on this task. In the first study using this task, Bechara, Damasio, Tranel, and Damasio (1997) measured participants' skin conductance (SCR) while they completed this task. The authors found that healthy participants began choosing more cards from decks C and D and began showing a higher SCR when choosing a card from decks A and B between the 10th-50th cards, despite being unable to verbalize whether one deck was better than the other. This pattern became very strong during trials 50-80, when participants began to express the possibility that decks C and D were better and continued to show a high SCR when choosing cards from decks C and D. During the last 20 trials, the pattern of choices remained relatively unchanged, but most participants were confident that decks C and D resulted in the most advantageous outcome. During this final phase, SCR remained high for decks A and B but became lower for decks C and D, suggesting that participants no longer experienced as much concern when choosing cards from decks C and D. The results of this study suggest that participants begin making predictions about which deck will result in less monetary loss, even before they become consciously aware of these predictions.

There have been multiple fMRI studies examining neural activity while participants perform the Iowa gambling task. In the first fMRI study of this task, Fukui, Murai, Fukuyama, Hayashi, and Hanakawa (2005) examined the selection period during which participants made their choice about which deck to choose. The behavioral results replicated those of Bechara et al. (1997), demonstrating that participants began preferentially choosing cards from the advantageous decks beginning around trial 40 of 100. In addition, they found significant activation in the medial prefrontal cortex when participants chose cards from the risky decks

compared to the safe decks. Furthermore, the more successful participants were on the task, the greater the activity in this same region. The authors suggest that when deciding which deck to choose, participants create an estimate of the probability of gain versus loss, and that it is this prediction that is represented in the activity in the medial prefrontal cortex.

Lawrence, Jollant, O'Daly, Zelaya, and Phillips (2009) found similar results in their fMRI study of seventeen men who completed a similar version of the Iowa gambling task. They again replicated the original Bechara et al. (1997) behavioral results. In addition, they compared brain activity during the selection period compared to activity in a control task in which participants were told which choices to make, and they found significant increased activation in the medial orbito-frontal cortex and the ventral anterior cingulate cortex. Furthermore, when they compared activation during the selection period on trials in which participants chose the risky decks compared to trials when participants chose the safe decks, they found increased activation during risky decisions in the medial frontal gyrus, lateral orbitofrontal cortex, insula, and the occipital cortex. Providing further support for these results, another fMRI study found a correlation between expected gain and activation in the hippocampus, superior frontal gyrus, right medial frontal gyrus, inferior frontal gyrus, inferior orbito-frontal cortex, right amygdala, insula, and orbito-frontal/ventromedial prefrontal cortex (Li, Lu, D'Argembeau, Ng, and Bechara, 2010). The authors suggest that the amygdala sends a signal to the orbito-frontal/ventromedial prefrontal cortex when the potential for risk is present, and that the orbito-frontal/ventromedial prefrontal cortex then allows for conscious processing of the risk and resulting decision. Although the authors do not explicitly mention prediction when discussing these results, it seems likely that the activity in the orbito-frontal/ventromedial prefrontal cortex reflects predictions about gains and losses occurring during the decision making process.

In sum, non-verbal predictive looking tasks, language predictive looking tasks, probabilistic classification tasks, and gambling tasks all require the use of prior information to make predictions about future input. In addition, although they depend primarily on different sensory modalities, they all overlap in their use of visual processing. On the other hand, there are some clear differences between the tasks. Both predictive looking tasks use information stored primarily in semantic memory, whereas the probabilistic classification and gambling tasks rely on working and short-term memory to make predictions. In addition, the measurement tools differ across the tasks, with performance on the predictive looking tasks measured using oculomotor data and performance on the probabilistic classification and gambling tasks measured using accuracy data. The tasks also have similarities and differences at the neural level. Although there are little EEG or imaging data available for the predictive looking tasks, the literature discussed earlier on predictions in the visual and language processing systems suggest that these tasks likely involve feed-back and feed-forward signals between lower-order areas and the frontal lobe. Similarly, imaging data for the probabilistic classification and gambling tasks suggest that information related to prediction is activated in the lower-order and higher-order brain areas, again suggesting the presence of feed-back and feed-forward connections. The behavioral and neural similarities between these tasks suggest that similar mechanisms may support performance on all of these tasks and that these tasks may load onto a single factor that represents overall prediction ability.

To test whether performance on these types of tasks involves a higher-order integrative prediction system, the current study included the PLAT, the visual world paradigm, the weather prediction task, and the Iowa gambling task. The visual world, weather prediction, and Iowa gambling tasks were chosen because all three have been the focus on extensive research. While

the PLAT is a relatively new task, it was chosen because preliminary research found that it reliably measured predictive looking during viewing of naturalistic movies and because eye tracking may allow for a more sensitive measure of prediction than an overt behavioral response.

Surprisingly, there are extremely few studies that have previously investigated whether performance is correlated across any combination of these types of tasks, and there have not been any previous studies that have combined three or more of these tasks into a single study. The few studies that have investigated performance on two of the tasks were both focused on performance in clinical populations (HIV/AIDS and schizophrenia) and used only the Iowa gambling task and the weather prediction task (Gonzalez, Wardle, Jacobus, Vassileva, & Martin-Thormeyer, 2010; Wasserman, Barry, Bradford, Delva, & Beninger, 2012). Both studies found that performance across these tasks was not as similar as expected, though neither study included a control group of healthy participants. It is therefore not possible to determine from these previous studies whether this lack of relationship between the two tasks would generalize to healthy populations.

In addition to completing these prediction tasks, participants also completed tasks involving crystallized and fluid intelligence to ensure that similarities among these tasks were not completely explained by other areas of cognitive functioning. The current study therefore tested whether there was a unique factor that explained shared variance across the prediction tasks, even when the fluid and crystallized intelligence tasks were included in the model.

In addition to furthering understanding of predictive processing, this study can inform applied research with clinical populations. For example, people with Posttraumatic Stress Disorder (PTSD) often experience hypervigilance and engage in constant surveillance of the environment to prevent themselves from experiencing a reoccurrence of their traumatic event (Ehlers & Clark, 2000); in other words, they make often erroneous predictions about potential

dangers in their environment. To illustrate, a military veteran might anticipate the presence of

enemy soldiers around every corner, even though these predictions are incorrect. In addition,

research has found that people with PTSD display increased arousal not only in response to

threat-related information, but also in response to novel, demanding, or unpredictable cues

(Stam, 2007) and that people with PTSD have deficits on neutral attention tasks, primarily with

inhibiting responses to distracters (Vasterling, Brailey, Constans, & Sutker, 1998). In fact, in a

recent study in our laboratory, we found that people with PTSD made slower and less accurate

predictions about everyday activity compared to controls (Eisenberg, Zacks, Rodebaugh, &

Flores, in prep). Furthermore, imaging studies have found reduced activity in the dopamine

system in combat veterans (van Wingen et al., 2012), increased activity in the striatum in people

with PTSD (e.g., Linnman, Zeffiro, Pitman & Milad, 2011; Falconer et al., 2008), and increased

activity in the dorsal anterior cingulate cortex in people with PTSD (e.g., Shin & Liberzon,

2010). These studies suggest that feed-back and feed-forward signaling necessary for successful

predictions may be different in people with PTSD compared to people without PTSD. Because

these studies suggest that people with PTSD make incorrect predictions about future stimuli and

have difficulty inhibiting responses to distracters, I hypothesized that people with PTSD also

have difficulty on other tasks that require correct predictions. Therefore, the present study

screened participants for the presence of PTSD symptomology to determine whether people with

higher PTSD severity experience difficulty with a variety of tasks involving prediction.

Chapter 2: Methods

2. 1 Participants

Two hundred seventy-six participants were recruited from the student participant pool at Washington University. Fifty participants were dropped from analyses because they did not complete both sessions of the study (26), the eye tracker could not track their eyes and no eye tracking data was collected (14), they were not fluent in English (3), they did not follow task instructions (4), or computer problems prevented them from completing the first session (3). An additional 15 participants were dropped from only the visual world task analyses because these participants were missing eye tracking data for more than 20% of the trials on this task but had adequate data for the other tasks. This left data from 226 participants for all but the visual world task and data from 211 participants for the visual world task. All analyses on single tasks and all simple pairwise correlations that did not involve the visual world task included data from all 226 participants. All modeling was conducted with data only from the 211 participants with full data sets. Participants ranged in age from 18 to 59 (mean = 19.73), and were 64% female. The study took place over two sessions: a group testing session that lasted 1.5 hours and an individual eye tracking session that lasted 1 hour.

2. 2 Eye-Tracking

An EyeLink 1000 eye tracker was used to collect oculometric measures. This eye tracker records data at 1000 Hz. Gaze location was the measure of particular interest; however, we also collected other oculometric measures including pupil size, fixation duration and saccade distance. Participants were required to keep their head in the headrest throughout all of the tasks requiring eye tracking. Nine-point calibration followed by validation of the resulting calibration was used; however when nine-point calibration did not allow for adequate calibration, thirteen-

point calibration was used. If both nine-point and 13-point calibration failed, five-point calibration was used. The infrared illuminator was initially set to 75% illumination; if calibration failed at this illumination it was adjusted to either 50% illumination or 100% illumination depending on which level of illumination provided the best calibration. After successful calibration and validation, the experimenter used a simulated pupil of known size printed on an index card to calibrate the pupil size measure, because the eye-tracker measures pupil size in pixels rather than in millimeters. All eye-tracking tasks were presented on a 19-inch (74 cm) monitor (1440x900 resolution, viewing distance of 58 cm from the forehead rest, viewing angle of 38.6°) using the Experiment Builder software designed by S-R Research (http://www.sr-research.com) to be used with this eye tracker.

2.3 Procedure

Participants first completed the group session. During the group session, participants reviewed the consent form and completed the demographics questionnaire. Participants were then told to begin the tasks on the computer. All computer tasks were presented on 23-inch (58.4 cm) monitors, with a viewing distance of 68.5 cm. Each task began once the previous task was complete, with no intervention from the experimenter. Participants began by completing the weather prediction task and then the Iowa gambling task. They then completed the letter sets task, followed by the synonym and antonym vocabulary tasks. Participants continued with the paper folding task and then the Information task. They finished the group session by completing the Raven's Progressive Matrices task. During the individual eye tracking session, participants completed the PLAT and then the visual world task. They then filled out the PTSD questionnaires. They were given a debriefing form explaining the purpose of the study before

they left this session. (See Table 1 for a list of the tasks participants completed in each session of the study.)

Table 1. List of tasks within each session of the study.

Group Session
Demographics Questionnaire
Weather Prediction
Iowa Gambling
Letter Sets
Synonym
Antonym
Paper Folding
Information
Raven's Progressive Matrices

Individual Eye Tracking Session
Predictive Looking at Action
Visual World
PTSD Questionnaires

2.4 Measures

2.4.1 Demographics

Participants completed a short demographic questionnaire that included age, gender, handedness, ethnicity, current employment, highest level of education, history of major medical problems, and hours of exercise.

2.4.2 PTSD Questionnaires

Participants first completed the Life Events Checklist for DSM-5 (LEC-5; Weathers, Blake, Schnurr, Kaloupek, Marx, & Keane, 2013), which included 17 questions about a variety of potentially traumatic events and eight additional questions about the severity of the most severe event. Participants then completed the PTSD Checklist for DSM-5 (PCL-5; Weathers, Litz, Keane, Palmieri, P.A., Marx, & Schnurr, 2013), which consisted of 20 questions assessing the

severity of all DSM-5 PTSD symptoms participants experienced over the past month. The National Center for PTSD has proposed a cut-point of 33, at or above which someone would be classified as having probable clinical PTSD, though they note that this could change after further research (National Center for Posttraumatic Stress Disorder). The PCL-5 has strong internal consistency ($\alpha = .94$), test-retest reliability ($r = .82$), and convergent and divergent validity (Blevins, Weathers, Davis, Witte, & Domino, 2015).

2.4.3   *Crystallized and Fluid Intelligence Tasks*

Participants completed three tasks testing crystallized intelligence and three tasks testing fluid intelligence. The crystallized intelligence tasks included the Information Test (Wechsler, 2008) and the Synonym and Antonym Vocabulary tasks (Salthouse, 1993). The Information Test required participants to answer general knowledge questions in a variety of areas, and a meta-analysis suggests that this measure has a test-retest reliability of 0.92 (Calamia, Markon, & Tranel, 2013). The Synonym vocabulary (Chronbach's alpha = .67) and Antonym vocabulary (Chronbach's alpha = .79) tasks required participants to choose synonyms or antonyms, respectively, from among five possible choices (Salthouse, 2001).

The fluid intelligence tasks included a paper folding task (Ekstrom, French, Harman, & Dermen, 1976), a letter sets task (Ekstrom, French, Harman, & Dermen, 1976), and the odd numbered questions in the Raven's Advanced Progressive Matrices Set II (Raven, 1990). For the paper folding task, participants were shown a sequence in which a square piece of paper was folded. The final image in the sequence showed where a pencil was poked through one location on the folded paper. Participants had to choose which of five options correctly displayed the locations of the holes on the unfolded piece of paper. Participants completed two sets of ten questions each. They were given three minutes for each set of questions. Each question was

presented by itself on the computer screen and participants were allowed to click a button to skip questions if they chose. For the letter sets task, participants were shown five strings of four letters each. Participants were instructed to choose the string that did not match the pattern that the remainder of the letter strings followed. Participants were given seven minutes to complete fifteen questions. Reliability data is not available in the literature for the Letter Sets and Paper Folding tasks. Therefore split-half reliability for these tasks was calculated using data from the current study (see results section for the results of these analyses). The Raven's Advanced Progressive Matrices task required participants to choose which of eight items completed the pattern shown at the top of the screen. The correct item matched the pattern vertically and horizontally. Participants were given ten minutes to complete eighteen questions (Kane et al., 1990). The Raven's Advanced Progressive Matrices has high internal consistency (Chronbach's alpha = .83; Paul, 1985) and test-retest reliability (r = .83; Bors & Forrin, 1995)

2.4.4   *Prediction Battery*

The prediction battery consisted of four different tasks: the predictive looking at action task (PLAT), a visual world task (Altmann & Kamide, 1999), a weather prediction task (Knowlton, Squire, & Gluck, 1994), and the Iowa gambling task (Bechara, Damasio, Tranel, and Damasio, 1997).

As described earlier, the PLAT is a novel task that we have recently developed in the laboratory. For this task, participants passively watched 5-6 minute movies of an actor performing everyday activities while their eyes were tracked using an eye tracker. These movies included many goal-directed sequences of activity in which the actor orients toward an object, picks up the object, and then completes an action with that object. Predictive performance on this

task was measured by examining how early participants looked at objects before the actor came into contact with them. This task allowed for a relatively continuous measure of prediction.

To calculate performance on this task, an experimenter first identified all of the time points at which the actor came into contact with an object. Dynamic interest areas were then drawn around each contacted object. The dynamic interest areas were placed to capture fixations on the object of interest ranging from 3000 ms before contact to 1000 ms after contact. Interest areas were placed using the following rules: (1) All interest areas were rectangular in shape, (2) No interest areas were allowed to overlap in time and space, (3) If potential interest areas overlapped, only the first interest area was kept, (4) If the actor contacted an object by touching it with another object, the object in direct contact with the actor was considered the object of interest (e.g., if the actor put a bowl on the counter, the bowl was considered the object of interest), (5) Only objects that were fully onscreen when contacted were considered objects of interest, (6) If the longest dimension of an object was smaller than 105 pixels (visual angle of 2.9°), the interest area was created around the entire object, and if the longest dimension of an object was larger than 105 pixels, the interest area was created around the part of the object that the actor contacted, and (7) For objects smaller than 48 pixels (visual angle of 1.3°) on any side, interest areas were created with a minimum size of 48 pixels per side. (See Figure 2 for an example movie frame with an interest area highlighted).

Figure 2. An example frame taken from one of the three movies used in this study. The yellow box represents the interest area, which was drawn around the chandelier—the object the actor is about to contact in order to put up the streamer. The purple dot represents the gaze location of an example participant who watched this movie. Here, the participant looked at the chandelier before the actor contacted it. During the study, participants saw neither the yellow box nor their own gaze location.

Once these dynamic interest areas were created in Data Viewer, the amount of time participants spent fixating within each interest area during the 3000 ms before contact was calculated. The 3000 ms before contact was divided into thirty 100 ms bins, and bins with more than 20% of the eye tracking data missing were dropped from later analyses. Growth curve modeling was conducted using the lmer package in R (Bates, Maechler, Bolker, & Walker, 2015) to obtain growth estimates over the 30 time bins for the random effect of subjects. Growth curve modeling provides an estimate of how performance on a task changes over time and tests whether allowing each subject's growth over time to have different intercepts, linear slopes,

and/or quadratic slopes improves the fit of the model. All variables entered into the growth curve models were z-scored to ensure that the variances were at a similar scale. Three models were tested: 1) a model that only allowed the intercept to vary by subject, 2) a model that allowed the intercept and linear slope to vary by subject, and 3) a model that allowed the intercept and both the linear and quadratic slopes to vary by subject. For all of these models, movie was also included as a random effect, but only the intercept was allowed to vary by movie.

For the visual world task (Altmann & Kamide, 1999), the eye-tracker was used to obtain participants' gaze location while they viewed arrays of objects and listened to sentences that included some of the objects. The sentences all had the following structure: an article, a noun, a verb, a person's name, and a noun that was the subject of the verb (e.g., The boy kicked Tracy's ball). On some trials, the verb was predictive of only one object in the array (predictive trials; 12 trials of this type per subject), while on other trials, the verb was not predictive of any individual object in the array (unpredictive trials; 12 trials of this type per subject). For example, one array of objects included a picture of football, a tennis ball, a toy truck, and a piece of broccoli. The predictive sentence was, "The woman steamed Paul's broccoli," and the unpredictive sentence was, "The woman put away Paul's broccoli" (See Figure 3). Predictive looking was measured by determining how much earlier participants looked at the target picture while listening to predictive sentences than while listening to unpredictive sentences. Three types of control sentences were also included: 1) the verb predicted two of the objects (8 trials per subject), 2) the verb predicted three of the objects (8 trials per subject), and 3) the verb did not apply to any of the objects (20 trials per subject). Participants heard a beep immediately after the end of each sentence, and they were instructed to press one button if they thought the sentence applied to any

38

of the pictures on the screen and press a different button if they thought the sentence did not

apply to any of the pictures. They were told to wait until they heard the beep to respond.

Figure 3. Example stimulus from the visual world task. The predictive sentence for this trial was
"The woman steamed Paul's broccoli," and the unpredictive sentence was "The woman put away
Paul's broccoli."

To calculate performance on this task, an experimenter first determined the amount of time

between the verb and the subject in each predictive and unpredictive sentence, hereafter referred

to as verb-subject distance. The verb-subject distance ranged from 1406 ms to 2400 ms (mean =

1892.68 ms). Then, the proportion of time spent looking at the target object during the verb-

subject distance during the predictive sentences was calculated. To ensure that the amount of

time spent looking at the target object was not due to the salience of that particular object, a control measure of looking time was calculated using the same object in the matched unpredictive sentence that only differed in the verb. Because each participant only heard one version of each sentence, the mean looking time across participants who heard the unpredictive sentence was used as the control for each predictive sentence.

For the weather prediction task (Knowlton, Squire, & Gluck, 1994), participants were told to use cues to predict whether there would be rain or sun. On each trial, they saw one, two, or three cards, each with geometric symbols. Each combination of cards had a different probability of predicting each outcome. For example, if cards three and four were presented, there was a .1 probability of rain, whereas if cards one and two were presented, there was a .9 probability of rain (See Figure 4 for an example trial from this task). Each combination of cards was presented in a random order with the frequency displayed in the P(cue) column of Table 2. (See Table 2 for the cue card combinations, the frequency with which each combination was presented, and the probability of rain given each combination.) Participants completed 200 trials of this task. For each trial, prediction ability was measured by determining whether participants chose the optimal response based on the cues given. Trials for which the probability of rain was 0.50 were scored as correct regardless of the response given. The resulting accuracy scores for each trial were entered into a logistic growth curve model using the lmer package in R (Bates, Maechler, Bolker, & Walker, 2015), with trial number as the time variable and trial by trial accuracy as the dependent variable. The logistic slope was allowed to vary randomly by subject. Trial number was z-scored to improve numerical precision.

Sun   Rain

Figure 4. Example trial from the weather prediction task.

Table 2: Design of Weather Prediction Task (Modified from Knowlton, Squire, & Gluck, 1994)

Pattern	Cue				P (cue)	P (rain)
	1	2	3	4		
1	0	0	0	1	0.14	0.15
2	0	0	1	0	0.084	0.38
3	0	0	1	1	0.087	0.1
4	0	1	0	0	0.084	0.62
5	0	1	0	1	0.064	0.18
6	0	1	1	0	0.047	0.5
7	0	1	1	1	0.041	0.21
8	1	0	0	0	0.14	0.85
9	1	0	0	1	0.058	0.5
10	1	0	1	0	0.064	0.82
11	1	0	1	1	0.032	0.43
12	1	1	0	0	0.087	0.9
13	1	1	0	1	0.032	0.57
14	1	1	1	0	0.041	0.79

Finally, participants completed the Iowa gabling task on the computer. For each of 100

trials, participants were told to choose a card from one of four decks. Participants were told to

choose cards to earn as much money as possible over the course of the task. Choosing cards from

the A and B decks resulted in large gains but larger losses and choosing cards from the C and D

decks resulted in small gains and smaller losses. The optimal strategy on this task was to choose

cards from the C and D decks. Participants could choose their own strategy for picking cards

from the decks, but after choosing forty cards from a single deck, that deck disappeared, and

participants had to start picking from one of the other decks. (See Figure 5 for an example trial

from the Iowa Gambling task.) Responses were scored as correct if the participant chose from

one of the decks that resulted in a higher long-term payout (decks C or D). A final score for this

task was calculated by summing the total number of correct responses over the one hundred trials

of this task. See the Results section for split-half reliability measures for this and each of the
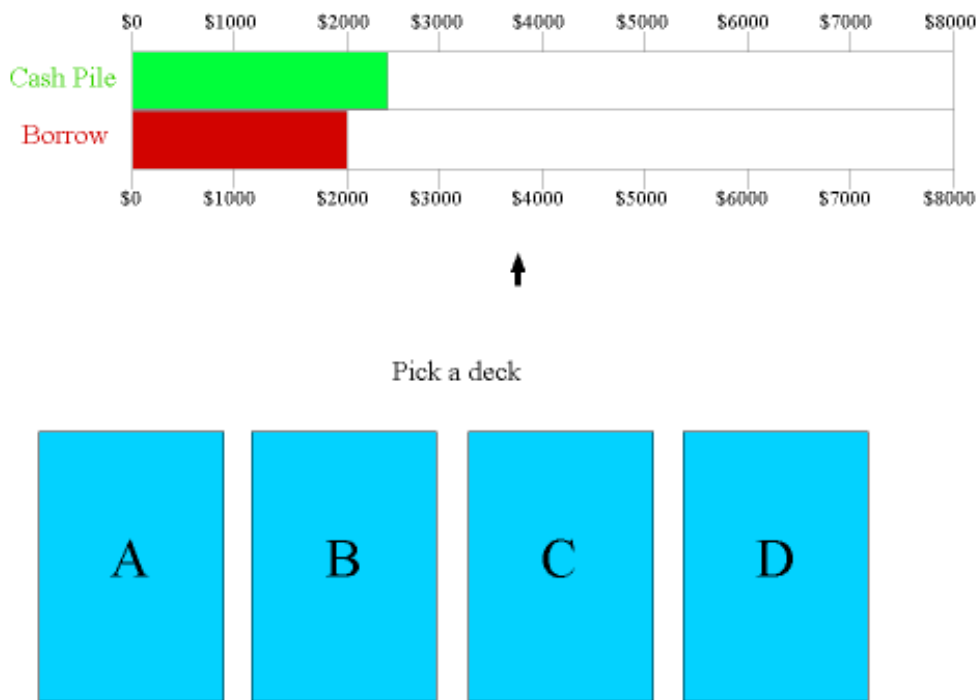
other prediction tasks.

Figure 5. Example trial of the Iowa Gambling task. Participants began with $2,000 in their cash
pile and earned or lost money by choosing from decks A through D.

2.4.5   Processing Speed

None of the tasks included in the battery described above directly tested processing speed; however, response times for all of the tasks were collected, allowing for an approximation of processing speed to be calculated. Originally, we planned to include response times for the visual world task, the weather prediction task, the synonym vocabulary task, and the antonym vocabulary task. However, as discussed in the results section below, preliminary analyses found that the response times for these tasks did not correlate, and we therefore decided to use only the response time for the visual world task, as this seemed the purest measure of processing speed available from the battery of tasks.

2.5 Modeling

Confirmatory factor analysis was used to determine whether the cognitive ability measures loaded onto the predicted latent variables, and whether the predictive processing tasks loaded onto the hypothesized latent variable. Model fit was calculated using the Tucker-Lewis Index (TLI), the Standardized Root Mean Square Residual (SRMR), the comparative fit index (CFI), and the root-mean squared error of approximation (RMSEA). The TLI compares the chi-square value of the model to the chi-square of the null-model (which specifies that there are no correlations among the measured variables) and takes into account model complexity; the SRMR is the square root of the difference between the residuals of the sample covariance matrix and the hypothesized covariance model; the CFI compares the fitted model with the null model, which specifies that the covariance of the variables is 0; and the RMSEA provides an estimate of how well the model fits the population covariance matrix. Hu and Bentler (1999) suggest that for sample sizes under 500, model fit is considered good when the TLI is greater than .95 and the SRMR is less than .09. They also suggest that model fit is good when the CFI is .95 or higher

and the RMSEA is .06 or lower. In addition, because of the relatively small sample size, the

Swain correction (Herzog, Boomsma, & Reinecke, 2007), which reduces the bias of model fit

estimators when the ratio of sample size to the number of estimated parameters is at least two to

one, was used to correct bias in the TLI, CFI, and RMSEA. It was planned a priori to use

structural equation modeling to test the model shown in Figure 6 if the confirmatory factor

analyses found that the hypothesized prediction and general cognitive functioning latent

variables provided a good fit for the data.
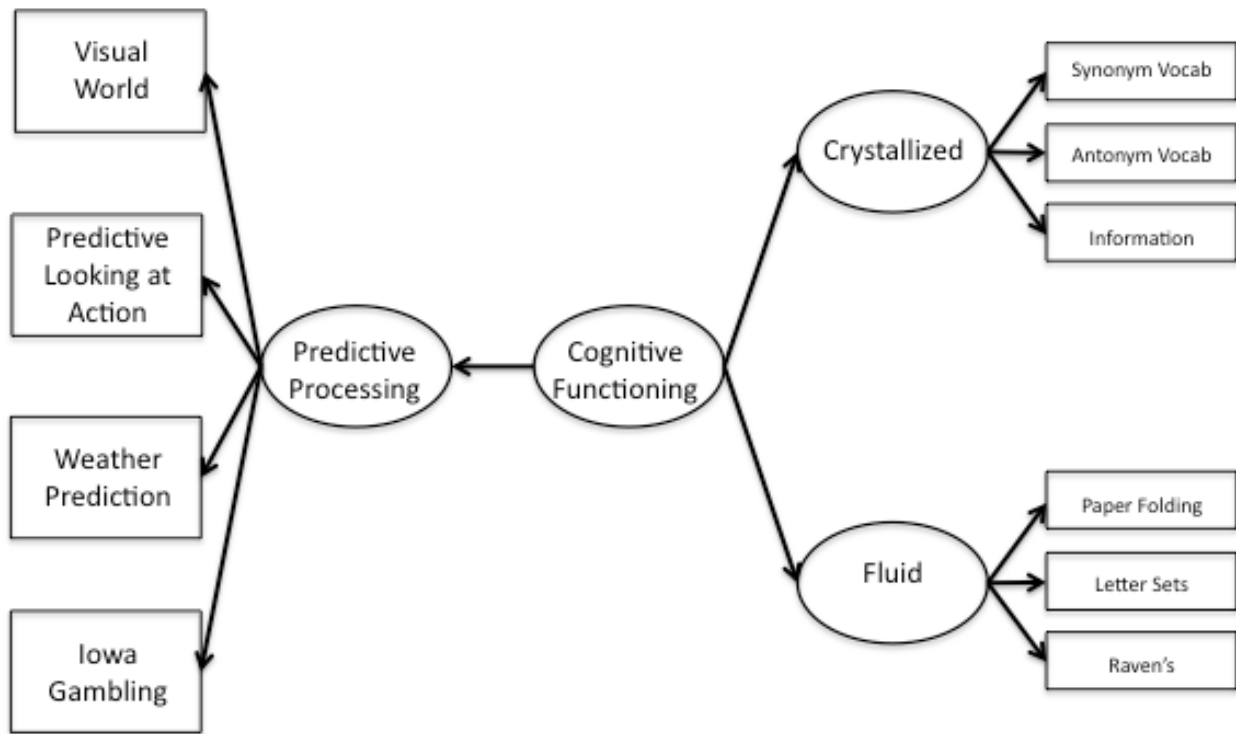
Figure 6. Proposed model structure with all four of the prediction tasks loading onto a predictive
processing latent variable. This predictive processing latent variable was hypothesized to load
onto the general cognitive functioning variable, and the crystallized and fluid intelligence
constructs were hypothesized to load onto the general cognitive functioning latent variable.

An additional model that included PTSD symptom severity was also tested to determine

whether higher PTSD symptom severity affects prediction ability and general cognitive

functioning. The hypothesized model that includes PTSD symptom severity is shown in Figure 7.
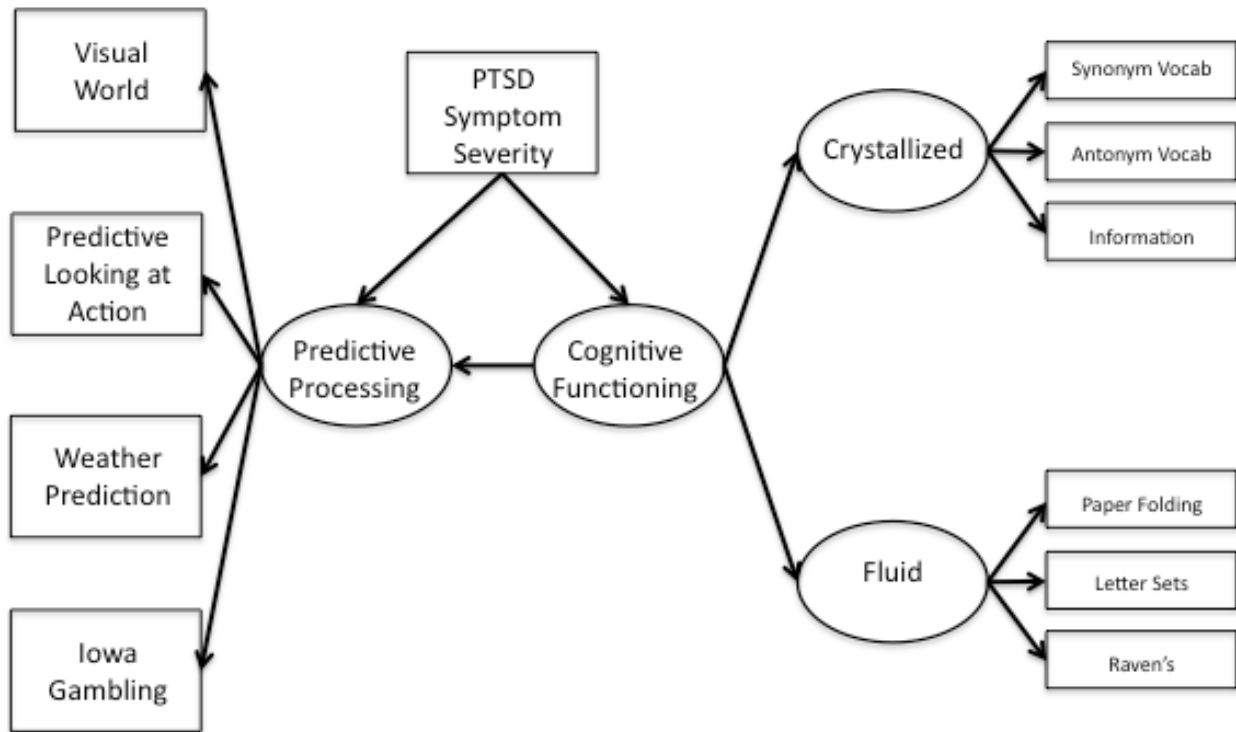
Figure 7. Model structure with PTSD as an additional predictor of general cognitive functioning and predictive processing.

Finally, an additional model, in which the fluid and crystallized intelligence constructs predict performance on each of the prediction tasks, was also tested to determine whether performance on the prediction tasks was simply a combination of other forms of cognitive functioning rather than a separate construct. This model is shown Figure 8.
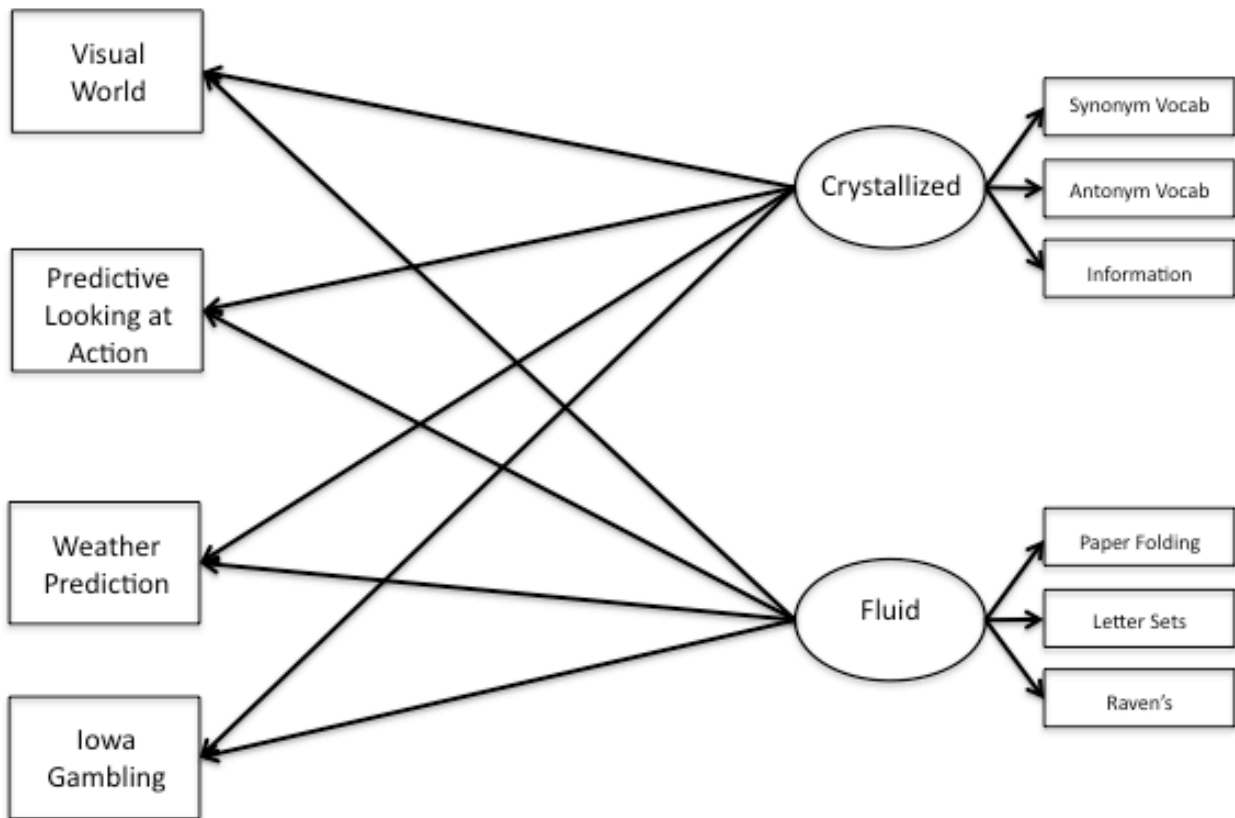
Figure 8. Model structure with crystallized and fluid intelligence constructs predicting performance on each of the prediction tasks.

Chapter 3: Results

3.1 Individual Task Performance

3.1.1   PTSD Questionnaires

Total PTSD severity scores and symptom cluster scores were calculated for each

participant (mean = 10.65, SD = 13.10, range: 0 to 60). The suggested cut-off score for probable

PTSD is 33 (National Center for Posttraumatic Stress Disorder, and only 19 participants had a

score of 33 or above. See Figure 9 for the distribution of PTSD severity scores. To determine

whether the scores on this measure followed the expected factor structure, scores for each

question were entered into a confirmatory factor analysis with each symptom loading on the

appropriate symptom cluster latent factor and all of the symptom cluster latent factors loading

onto a total PTSD score latent factor.

The model for the confirmatory factor analysis had a TLI of .834, a SRMR of .068, a CFI of

.855, and a RMSEA of .106. Although the SRMR suggests that the model is an adequate fit for

the data, the other measures suggest that this model should not be considered a good fit for the

data. This means that the factor structure of PTSD symptoms in this sample only loosely fits the

factor structure of PTSD in the DSM-5. Therefore, while scores on the PTSD questionnaires

were used in subsequent analyses, results from these analyses should be considered in the

context of a less than adequate model fit. See Table 3 for the standardized factor loadings for

each question on the PCL-5. As is evident from this table, the factor loadings for reexperiencing,

avoidance, and alterations in mood and cognition were very high. However, some of the factor

loadings for increased arousal were lower, which may have contributed to the failure of the

model to adequately fit the data in this study.

Distribution of PTSD Severity Scores
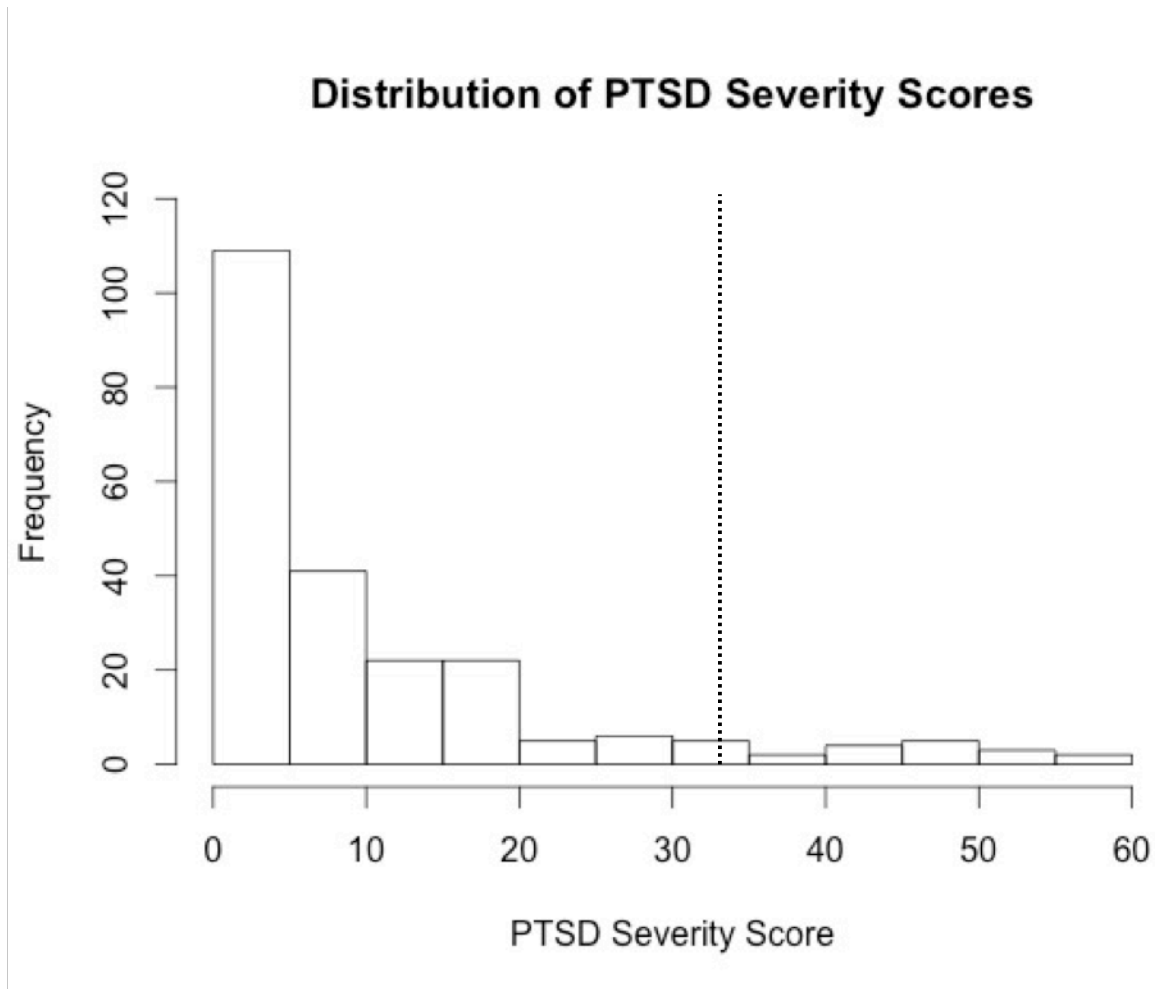
Figure 9. Distribution of PTSD Severity Scores. The suggested cut-off score for probable PTSD is 33, represented by the dotted line (National Center for Posttraumatic Stress Disorder).

Table 3. Factor loadings for the PCL-5

	p	Standardized Loading
Reexperiencing		
Q1		
Q2	<.001	0.65
Q3	<.001	0.66
Q4	<.001	0.82
Q5	<.001	0.82
Avoidance		
Q6		
Q7	<.001	0.84
Alterations in Cognition and Mood		
Q8		
Q9	<.001	0.86
Q10	<.001	0.80
Q11	<.001	0.86
Q12	<.001	0.75
Q13	<.001	0.79
Q14	<.001	0.73
Increased Arousal		
Q15		
Q16	<.001	0.53
Q17	<.001	0.57
Q18	<.001	0.61
Q19	<.001	0.87
Q20	<.001	0.73
PTSD		
Reexperiencing		
Avoidance	<.001	0.91
Alterations in Cognition and Mood	<.001	0.88
Increased Arousal	<.001	0.82

3.1.2   *Crystallized and Fluid Intelligence Measures*

Scores on the Synonym Vocabulary, Antonym Vocabulary, and Information tasks were summed to generate a total score for each task. See Table 4 for descriptive statistics for these tasks.

Table 4: Descriptive Statistics

	Mean	SD	Min	Max
Crystallized				
Synonym	6.51	2.21	1.00	10.00
Antonym	6.34	2.11	1.00	10.00
Information	17.35	3.46	5.00	23.00
Fluid				
Letter Sets	11.40	2.29	3.50	15.00
Paper Folding	12.95	3.57	1.00	
Raven's	11.58	2.62	3.00	18.00
Prediction Battery				
PLAT	1.15	0.40	0.15	2.15
VW	0.00	0.04	-0.11	0.11
Weather	0.00	0.00	-0.01	0.01
Iowa Gambling	63.82	10.91	32.00	80.00
Processing Speed	509.00*	243.28	230.00	2703.00
PTSD	10.65	13.1	0.00	60.00

Note: PLAT = Predictive Looking at Action Task; VW = Visual World

*The median for the processing speed measure.

Scores on the Raven's Progressive Matrices task were also summed to generate a total score for each participant. For the Letter Sets and Paper Folding tasks, participants received one point for every correct response and lost .25 points for every incorrect response. An error in the Experiment Builder code for the Letter Sets and Paper Folding tasks caused the program to skip through items and report an extremely short reaction time of less than 100 ms. This affected twenty participants for the Letter Sets task (one participant was missing three items, eight participants were missing two items, and eleven participants were missing one item), and twenty participants for the Paper Folding task (one participant was missing five items, three participants were missing four items, one participant was missing three items, eight participants were missing two items, and seven participants were missing one item). Missing responses for each task were

imputed using the mean score for the non-missing responses[1]. The split-half reliability for the

Letter Sets task was 0.44 and the split-half reliability for the Paper Folding task was 0.65. The

low split half reliability for these tasks are quite low, which may be due, in part, to the small

number of items for each task (15 for the Letter Sets task and 20 for the Paper Folding task). See

Table 4 for descriptive statistics for these tasks.

3.1.3   Prediction Battery

For the PLAT, three growth curve models were tested: 1) a model that only allowed the

intercept to vary by subject, 2) a model that allowed the intercept and linear slope to vary by

subject, and 3) a model that allowed the intercept and both the linear and quadratic slopes to vary

by subject. A comparison of these models suggested that the model with random intercepts and

linear and quadratic slopes provided the best fit for the data ($\chi^2(3) = 970.26$, $p < .001$). For this

model, the fixed linear ($t = -26.49$, $p < .001$) and quadratic ($t = 38.23$, $p < .001$) effects of time

bin were significant. In addition, the random intercept for movie had a variance of 0.09 (SD =

.29), and the residual variance was 0.22 (SD = .46). The random intercept for subject had a

variance of 0.11 (SD = .34), the random linear slope for subject had a variance of .02 (SD = .16),

and the random quadratic slope for subject had a variance of .17 (SD = .41). The intercept and

random linear slope were correlated with $r = -0.74$, the intercept and quadratic random slope

were correlated with $r = 0.89$, and the linear and quadratic random slopes were correlated with r

---

[1] Multiple imputation was not performed because participants were allowed to skip items, which

likely resulted in non-random missing data. To ensure that imputing using the mean did not

significantly affect the results, a follow-up analysis was performed in which the missing data

were not imputed, and instead were treated as missing. All of the models were then tested using

the resulting data. The pattern of results was identical to that reported below.

= -0.97. Because the linear and quadratic random slopes were so highly correlated, only the

quadratic random slope was used as an individual difference measure in later analyses. A split

half reliability test found that this individual difference measure had adequate reliability (r =

0.79, *p* < .001). (See Figure 10 for subject level data for this task and Table 4 for descriptive

statistics for this task.)
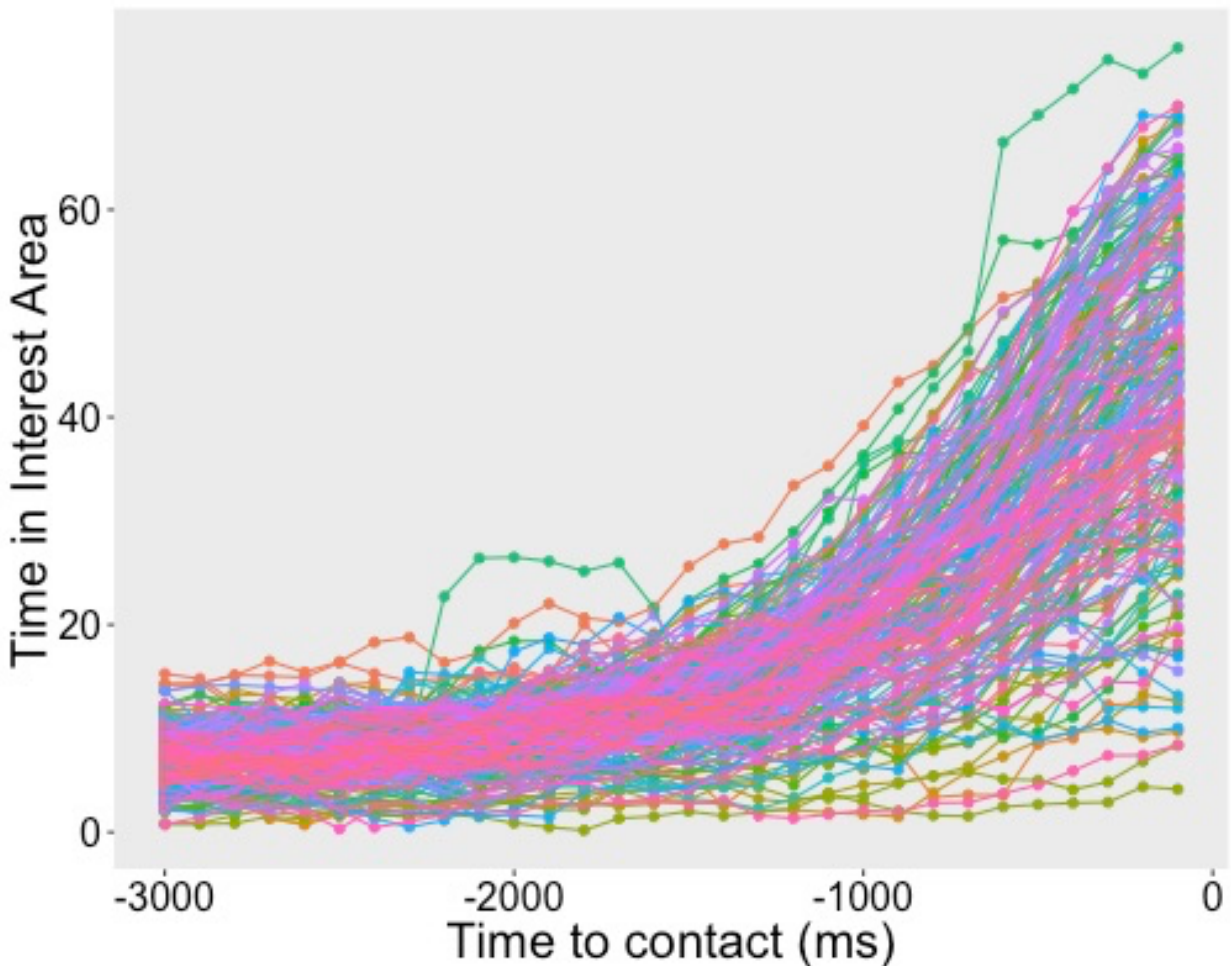
Figure 10. Participant level data for the PLAT. Each colored line represents one participant's
performance on this task. The x-axis represents the 3000 ms before contact for the interest areas,
divided into 100 ms bins. The y-axis represents the number of milliseconds participants spent
looking in the interest areas during each 100 ms bin.

For the visual world task, two sentences were dropped from further analysis because they

engendered very low predictive looking across participants, with participants only looking at the

target item 3% or 4% of the time during the predictive sentences. Then, the data from the remaining 22 items were entered into a mixed-effects model with the time spent looking at the target in the predictive sentences as the dependent variable and the time spent looking at the target in the unpredictive sentences as the independent variable. Only the intercept was allowed to vary by subject. The residuals from this model represented the time spent looking at the target item, controlling for the saliency of the item. The residuals for each item were averaged for each participant to get a predictive looking measure. Split-half reliability for the predictive looking measure was .53, and although this split-half reliability was lower than what is typically preferred, other scoring methods produced even lower reliability. (See Table 4 for descriptive statistics for the predictive looking measure for this task.) This predictive looking measure was entered into the structural equation models as the individual differences measure for the visual world task.

In addition, to ensure that the results in this study replicated results from other studies that have used the visual world paradigm, predictive looking results for the predictive and unpredictive sentences were plotted in Figure 11. It is evident from this figure that participants generally spent more time looking at the target object when the verb was predictive of only the target object compared to sentences in which the verb could reference all of the objects in the display. These results replicate those found in other studies of the visual world paradigm (e.g., Altmann and Kamide, 1999).
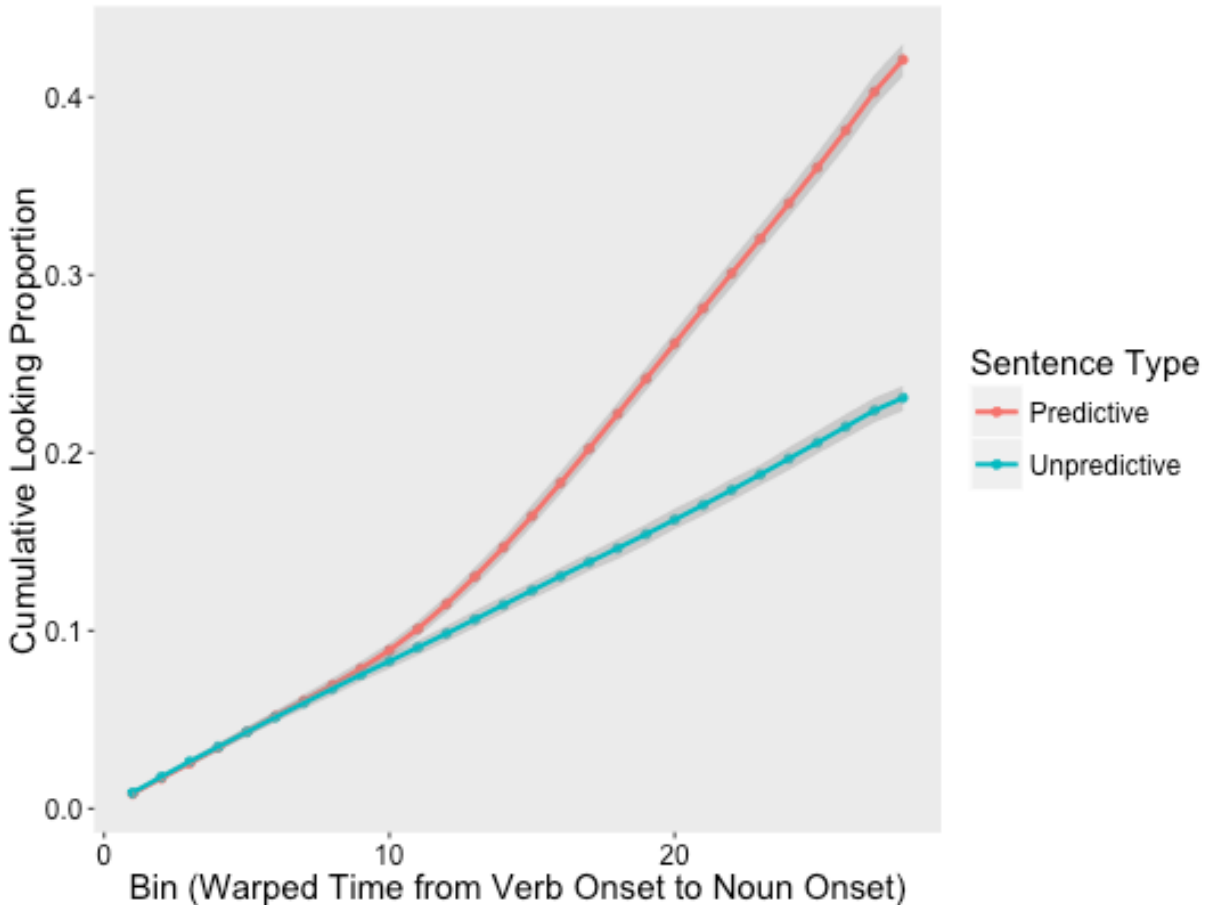
Figure 11. Predictive looking results for the visual world task. The time from verb onset to noun onset was warped into 28 bins with varying numbers of milliseconds in each bin, and these 28 bins are on the x-axis. The y-axis represents the cumulative proportion of time participants spent looking at the target object. The shaded regions around each line represent the 95% confidence intervals.

For the weather prediction task, the results from the logistic growth curve model suggested that the fixed effect of the logistic slope was significant ($z = 8.001$, $p < .001$). In addition, the random intercept had a variance of 0.25 (SD = .50), and the random logistic slope had a variance of .05 (SD = .22). The random logistic slope was used as the individual differences measure in subsequent analyses. (See Figure 12 for the logistic curves fit for each participant and Table 4 for descriptive statistics for this task.)
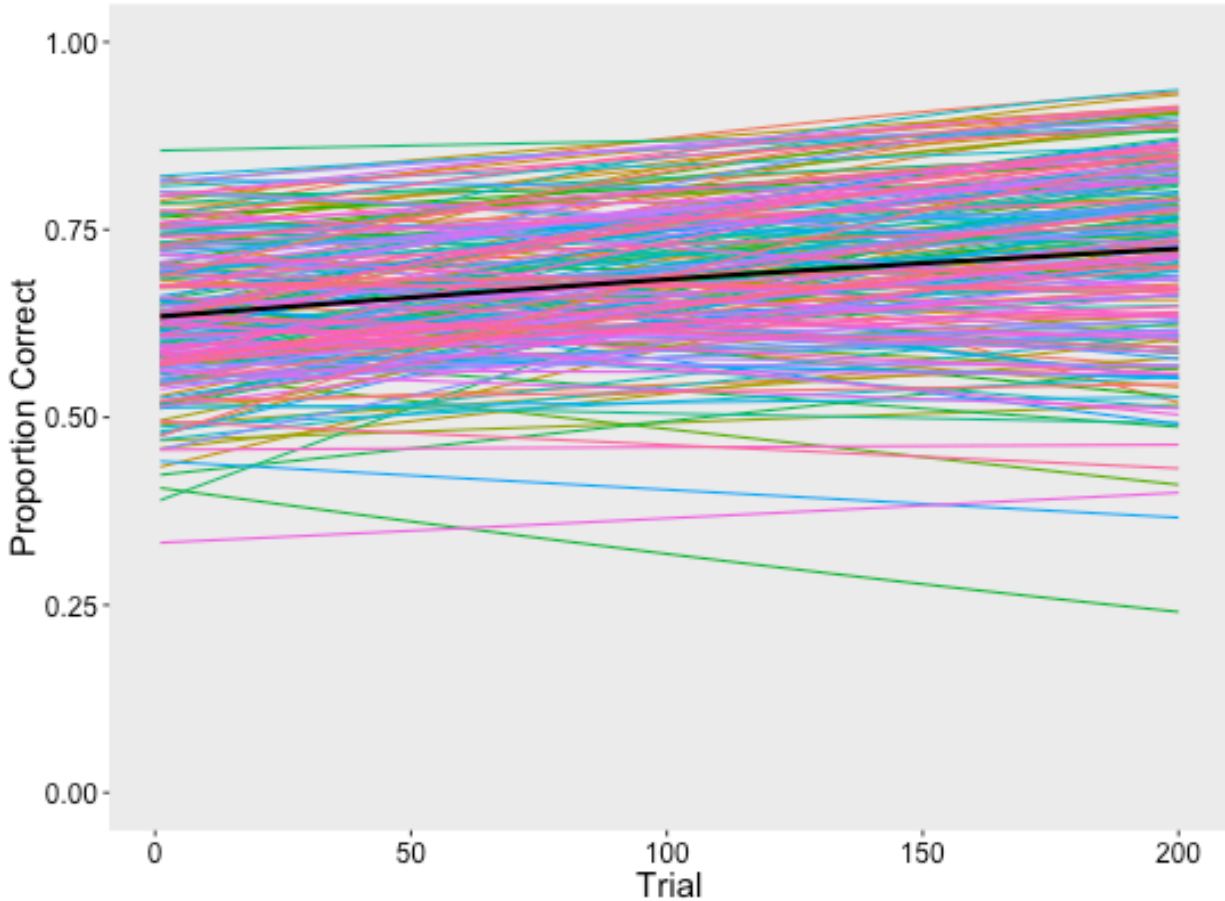
Figure 12. Participant level logistic general linear model results for the weather prediction task. Each colored line in this figure represents the best-fit logistic slope for an individual participant. The black line represents the overall model estimate for the logistic slope.

For the Iowa gambling task, responses were scored as correct if the participant chose from one of the decks that resulted in a higher long-term payout (decks C or D). A final score for this task was calculated by summing the total number of correct responses over the one hundred trials of this task. (See Table 4 for descriptive statistics for this task.)

3.1.4   Processing Speed

First, response times for visual world, weather prediction, synonym vocabulary, and antonym vocabulary tasks were calculated for each participant using the mean response time across items. For the weather prediction tasks, only response times for the last 50 trials of the

task that also only had one card presented were used for the processing speed measure. However, most of the correlations were very low (see Table 5). To ensure that the low correlations were not due to very long response times for some trials, the median response time and the 30th percentile response time was calculated for each item and then averaged within participants to obtain a potential processing speed measure. However, a similar pattern of correlations emerged from these analyses as well (see Table 6). Therefore, the median response time for the visual world task was used as a measure of processing speed for future analyses, as this seemed, a priori, to be the best measure of processing speed available from the tasks included in this study. (See Table 4 for descriptive statistics for this measure of processing speed.)

Table 5: Correlations of Means for Processing Speed Tasks

	VW	Weather	Synonym	Antonym
VW	1.00	0.04	-0.10	-0.08
Weather	0.04	1.00	0.15	0.17
Syonym	-0.10	0.15	1.00	0.76
Antonym	-0.08	0.17	0.76	1.00

Note: VW = Visual World Task

Table 6: Correlations of Medians and 30th Percentiles*

	VW	Weather	Synonym	Antonym
VW	1.00/1.00	0.11/0.12	0.13/0.12	0.08/0.11
Weather	0.11/0.12	1.00/1.00	0.15/0.18	0.16/0.25
Synonym	0.13/0.12	0.15/0.18	1.00/1.00	0.79/0.75
Antonym	0.08/0.11	0.16/0.25	0.79/0.75	1.00/1.00

*Number before the slash is the correlation of the medians; Number after the slash is the correlation of the 30th percentiles
Note: VW = Visual World Task

3.1.5   *Summary of Individual Differences Measures*

For the crystallized and fluid intelligence tasks, final score on these measures were entered into the structural equation models as the individual differences measure. For the PLAT, the

quadratic random slopes representing the time each participant spent looking at the target object over the three seconds before contact were used as the individual differences measure in the structural equation models. For the visual world task, the individual differences measure was the average time each participant spent looking at the target objects, controlling for the saliency of the objects. The individual differences measure for the weather prediction task was the random logistic slope representing subject level growth in performance over the trials of the task. For the Iowa gambling task, the individual differences measure was the total number of correct responses over all trials of the task. Finally, for the models that included PTSD severity score, total PTSD severity was used as the individual differences measure.

3.2 Modeling

First, a confirmatory factor analysis was conducted to determine whether responses on the synonym, antonym, and Information tasks loaded onto a crystallized intelligence latent factor and whether responses on the Raven's Progressive Matrices, letter sets, and paper folding tasks loaded onto a fluid intelligence latent factor. This model also included a term for the correlation of the crystallized and fluid latent factors. A second model in which the crystallized and fluid latent factors loaded onto a general intelligence latent factor (g) was also tested.

The same model fit indices as discussed above were used to determine whether these models provided a good fit for the data. For the model without the general intelligence latent factor, the TLI was .984, the SRMR was .035, the CFI was .991, and the RMSEA was .037. The Swain correction was then applied to correct for potential bias of model fit estimators. Using this correction, the TLI was .98, the CFI was .99, and the RMSEA was .036. All of these measures suggest that the model provided a good fit for the data. Figure 13 displays the factor weights for this model. However, the correlation between the crystallized and fluid latent factors was 0.09,

suggesting that these factors would not load onto a single general intelligence latent factor.

Therefore, the second model that included a single general intelligence latent factor was not

tested, and a latent factor for overall cognitive functioning was not included in any of the

following models.
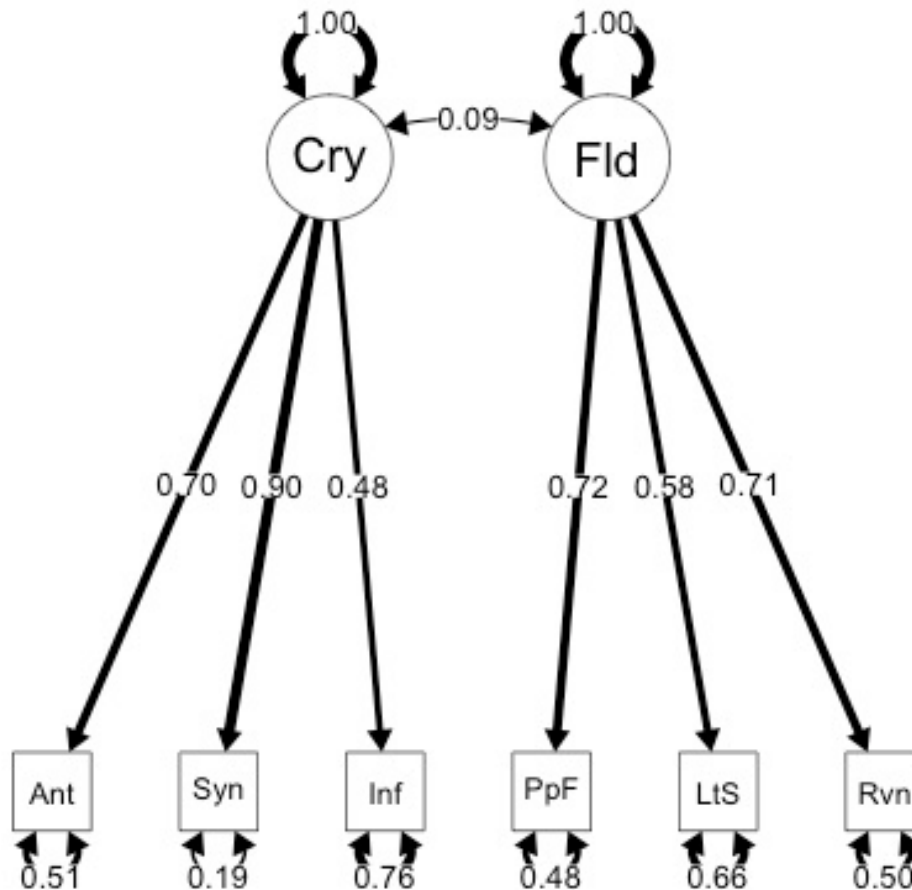
Figure 13. Results of the confirmatory factor analysis for the general intelligence measures. Ant = Antonym Vocabulary task; Syn = Synonym Vocabulary task; Inf = Information task; PpF = Paper Folding task; LtS = Letter Sets task; Rvn = Raven's Progressive Matrices; Crys = crystallized intelligence latent factor; Fld = fluid intelligence latent factor.

The correlations among the four prediction tasks were very low, with none of the

correlations above .11. (See Table 7 for the correlations among these tasks.) Although it was

therefore unlikely that the four tasks would load onto a single latent factor, a confirmatory factor

analysis was performed. Although the model converged, the fit indices suggested that this model was misspecified. For this model, the CFI was 1.00, the TLI was -14.58, the RMSEA was 0.00, and the SRMR was .01. Inspecting the loadings onto the latent factor revealed that none of the prediction tasks significantly loaded onto the latent variable, which was unsurprising given the low correlations among the tasks.

Table 7: Correlations among the prediction tasks

	PLAT	VW	Weather	IGT
PLAT	1.00	0.07	0.07	0.10
VW	0.07	1.00	0.02	0.11
Weather	0.07	0.02	1.00	0.01
IGT	0.10	0.11	0.01	1.00

Note: PLAT = Predictive Looking at Action Task; VW = Visual World Task

Because the prediction tasks did not load onto a single latent factor, the planned structural equation models could not be tested. Therefore, the model in Figure 14 was tested to determine whether fluid intelligence, crystallized intelligence, and processing speed predicted performance on each of the individual prediction tasks. The simple correlations among these measures are given in Table 8. For this model, the TLI was .95, the SRMR was .04, the CFI was .97, and the RMSEA was .04. After applying the Swain correction, the TLI was .95, the CFI was .97, and the RMSEA was .04. These fit indices suggest that this model provided a good fit for the data. For this model, the standardized regression coefficient for the relationship between crystallized intelligence and the PLAT was negative and significant (-0.17, $p = .02$), which was quite surprising given that knowledge of what tends to happen in a given situation would be expected to improve performance on the PLAT. The standardized regression coefficient for the relationship between fluid intelligence and the weather prediction task was significant (0.18, $p = .02$). Furthermore, the standardized regression coefficients for the relationships between the Iowa

gambling Task and crystallized intelligence (0.17, p = .03) and fluid intelligence (.16, p = .04) were significant. Finally, the relationship between processing speed and the PLAT (-.13, p = .047) was significant. All of the other regression coefficients were not significant. (See Figure 14 for the path diagram with all of the standardized estimates.)

Table 8. Simple correlations between the prediction tasks and the crystallized and fluid intelligence tasks.

	Antonym	Synonym	Information	Paper Fold.	Letter Sets	Ravens
PLAT	-0.12	-0.12	-0.02	0.04	0.00	0.06
Visual World	-0.08	-0.02	0.05	0.08	0.11	-0.05
Weather Pred.	-0.03	0.06	-0.03	0.16	0.08	0.11
Iowa Gambling	0.13	0.18	0.20	0.08	0.08	0.20
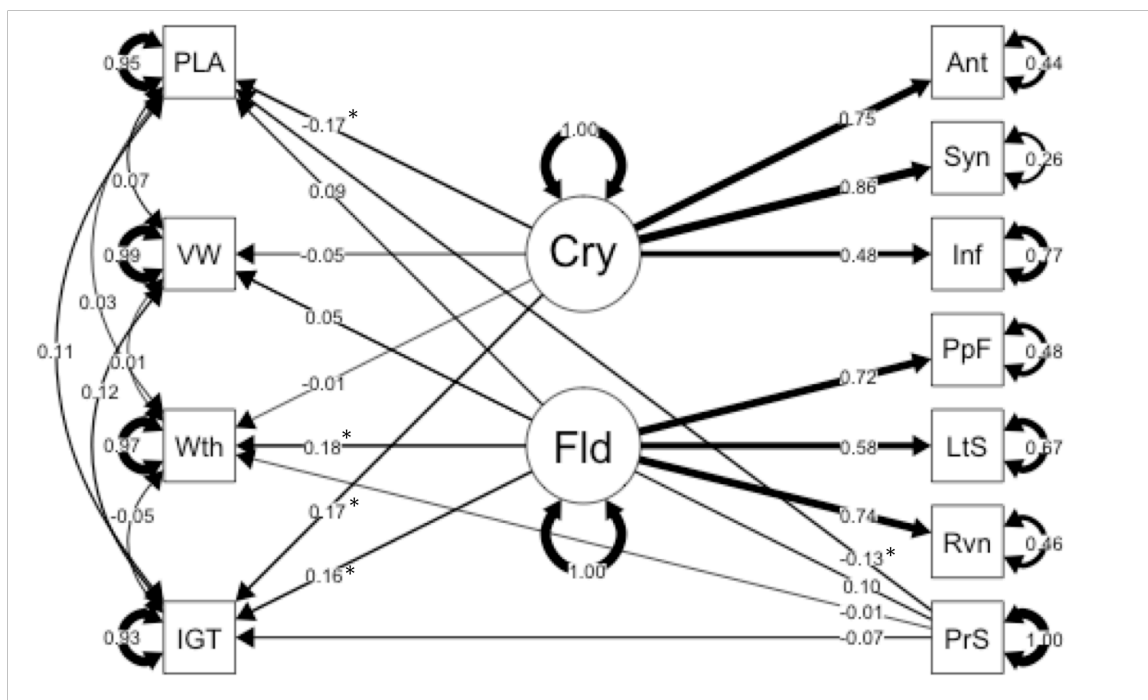
Figure 14. Path diagram. PLA = Predictive Looking at Action Task, VW = Visual World task, Wth = Weather Prediction, IGT = Iowa Gambling Task, Cry = Crystallized Intelligence, Fld = Fluid Intelligence, Ant = Antonym Vocabulary, Syn = Synonym Vocabulary, Inf = Information, PpF = Paper Folding, LtS = Letter Sets, Rvn = Raven's Progressive Matrices, PrS = Processing Speed. The weights of the arrows represent the magnitude of the path coefficients. A star next to a regression coefficient indicates a significant beta weight.

The final model included PTSD symptom severity as a predictor of performance on the

crystallized and fluid intelligence latent factors as well as each of the four prediction tasks. For

this model, the TLI was .95, the SRMR was .04, the CFI was .97, and the RMSEA was .03. After

applying the Swain correction, the TLI was .96, the CFI was .98, and the RMSEA was .03. These

fit indices suggest that this model also provides a good fit for the data. However, none of the

regression coefficients for the relationships between PTSD symptom severity and the other tasks

and latent variables were significant (PLAT: 0.02, $p = .73$; VW: -0.02, $p = .76$; Weather: -0.08, p

= .24; IGT: -0.04, $p = .50$; Fluid: -0.01, $p =. 91$; Crystallized: -0.04, $p = .61$), suggesting that

PTSD symptom severity was not related to performance on the other tasks included in this study.

(See Figure 15 for the path diagram with all of the standardized estimates.)
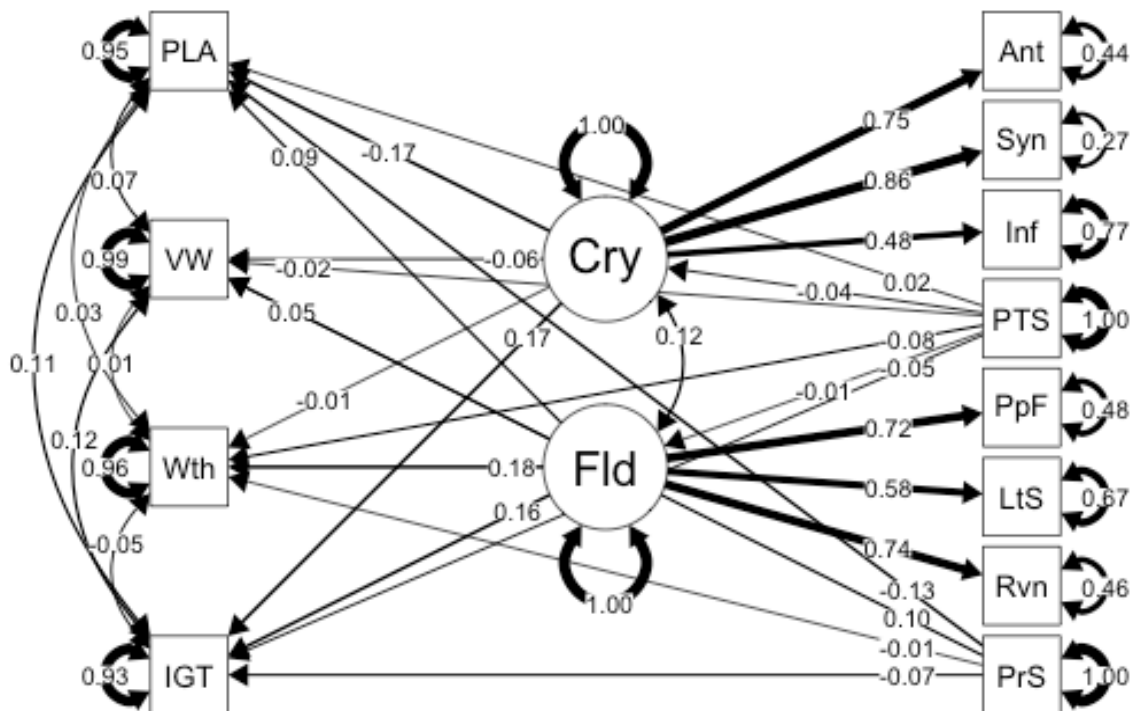
Figure 15. Path diagram for the model that included PTSD as a predictor of performance on the tasks. PLA = Predictive Looking at Action Task, VW = Visual World task, Wth = Weather Prediction, IGT = Iowa Gambling Task, Cry = Crystallized Intelligence, Fld = Fluid Intelligence,

Ant = Antonym Vocabulary, Syn = Synonym Vocabᵃulary, Inf = Information, PTS = PTSD, PpF = Paper Folding, LtS = Letter Sets, Rvn = Raven's Progressive Matrices; PrS = Processing Speed. The weights of the arrows represent the magnitude of the path coefficients.

Chapter 4: Discussion

4.1 No Evidence for a Prediction Construct

The current study was the first to directly investigate the question of whether there is a single higher-order integratory prediction mechanism in the brain, and the data provide evidence against the existence of such a mechanism. Specifically, if there were a higher order integratory prediction mechanism, performance on tasks that require prediction formation should have been correlated, regardless of the task modality. However, in the current study, the PLAT, visual world, weather prediction, and Iowa gambling tasks did not load onto a single latent factor, and the correlations among these tasks were uniformly extremely low. This is in contrast with current theories including the predictive coding theory (Friston, 2005) and Event Segmentation Theory (Zacks, Speer, Swallow, Braver, & Reynolds, 2007), which both imply the existence of a higher-order integratory prediction mechanism. The results of the current study also differ from the implications of previous neurophysiological and neuroimaging studies that found that similar signals are sent from lower-order to higher-order brain areas and vice versa in many systems in the brain (e.g, Tap & Bar, 2005; Tanaka et al, 2004; Dikker & Pylkkänen, 2003).

On the other hand, in concert with the results of the current study, neurophysiological data provide evidence against an integratory prediction mechanism, as there are many different prediction error signals in the brain (e.g., ERN, MMN, P300, N400, P600) rather than a single error signal used by all brain systems that is sent to an integrative prediction mechanism. In addition, in two of the very few studies that investigated performance on the weather prediction and Iowa gambling tasks in the same participants, performance across these tasks was not as similar as expected. For example, in a study of HIV positive participants with a history of

substance dependence, there was no significant relationship between the weather prediction task and the Iowa gambling task (Gonzales, Wardle, Jacobus, Vassileva, & Martin-Thormeyer, 2010). Similarly, in a study investigating the effects of antipsychotic medications in patients with schizophrenia, performance on the weather prediction task did not track performance on the Iowa gambling task (Wasserman, Barry, Bradford, Delva, & Beninger, 2012). Although neither of these previous studies included a control group of healthy adults, they do provide some converging evidence in for the lack of correlations among the four prediction tasks in the current study. The results of the current study therefore provide initial evidence against an integrative prediction mechanism and suggest that current theories involving prediction may need further examination.

If, as the results of the current study suggest, there is no higher-order integrative prediction mechanism that allows for performance on prediction tasks that require multiple modalities, how might people successfully perform these tasks? It is possible that when people engage in prediction within a particular modality, a network of regions, including regions that are specific to the primary task modality, is activated. If the task requires the integration of multiple sensory modalities, the network of regions that is activated may simply include the additional regions necessary for the second modality. While some of these areas would likely overlap, it could be the activity in the separate areas that drives behavioral performance on the tasks. For example, the Iowa gambling task, which requires visual processing, has been found to activate the occipital cortex, medial frontal gyrus, and orbitofrontal cortex, among other areas, when participants chose a risky deck compared to a safe deck (Lawrence, Jollant, O'Daly, Zelaya, and Phillips, 2009). In addition to requiring predictions within the visual modality, the Iowa Gambling task also involves reward processing, as participants obtain rewards and losses

throughout the task. Tasks involving reward prediction often activate the striatum and orbitofrontal cortex (e.g., Tanaka et al, 2004; Ernst et al., 2004). In fact, (Lawrence, Jollant, O'Daly, Zelaya, and Phillips, 2009) found activation in the orbitofrontal cortex when participants completed the Iowa gambling task compared to a control task in which participants were told which choices to make. It is therefore possible that the activation patterns seen when participants complete the Iowa gambling task are due to the activation of two separate brain networks—a visual prediction network and a reward prediction network—that operate in parallel and therefore appear to be a single activated network. Thus, overall performance on the Iowa gambling task would depend on prediction ability in the two separate modalities, and performance might very well not be correlated with a different task that requires predictions in a different modality (e.g., auditory predictions) which might activate yet another overlapping but different brain network. This hypothesis of overlapping but separate brain networks that are activated by the distinct predictive processing modalities involved in each task could explain the lack of correlations among the prediction tasks included in the current study. For example, if a particular participant tends to be better at making visual predictions than reward predictions, this participant might obtain a higher score on the PLAT, which is primarily a visual task, than on the Iowa gambling task, which also involves a reward prediction component. If participants have strengths and weaknesses that drive performance differently on each task, the low correlations among the tasks in the current study would be expected.

In addition to differences in the prediction modalities required by each task in the current study, method variance across the tasks may provide a less interesting source for the lack of correlations among the tasks, as individual differences in abilities other than prediction could have overshadowed prediction ability as a driver of performance. For example, the weather

prediction task may have required the ability to use spatial information while performing the task, as the spatial configuration of the stimuli predicted the correct response. Evidence from fMRI studies of probabilistic classification tasks provide support for this possibility. For example, an fMRI study of a probabilistic classification task found activation in the parietal cortex, an area that is often implicated in tasks involving spatial processing, during the time that participants were likely making predictions about the category of the stimulus (Aron et al., 2004). In fact, a second fMRI study of a different probabilistic classification task also found activation in the parietal lobe when participants experienced uncertainty about their response (Huettel, Song, & McCarthy, 2005). If some participants were better at learning spatial information, these participants may have performed better on this task regardless of their ability to make predictions, which would lower overall correlations among the tasks.

Similarly, each of the other tasks also includes components, distinct from prediction, which could drive performance. For example, Li, Lu, D'Argembeau, Ng, and Bechara (2010) suggest that the Iowa gambling task requires the processing of risk, finding that when people perform the task, the amygdala signals the presence of risk to the orbito-frontal/venteromedial prefrontal cortex. Though risk evaluation likely plays a role in making predictions in the Iowa gambling task, the other tasks in the current study did not heavily involve the evaluation of risk, which could explain the low correlations between this task and the other prediction tasks.

The visual world task very clearly involves a language processing component that is not relevant for any of the other tasks in the current study. Therefore, if participants varied in their ability to process language information, this may have masked the effect of individual differences in prediction ability on performance of the visual world task. An imaging study on a task that is conceptually similar to the visual world paradigm, in which participants listened to

sentences and chose which of three or four line drawings in a array best represented the content of the sentences, provides evidence that language comprehension is necessary for successful performance on the task (Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004). The authors found that lesions in the posterior middle temporal gyrus and anterior superior temporal gyrus, which are areas involved in language comprehension, strongly affected performance on the task. Therefore, it is possible that performance on the visual world task did not correlate with performance on the other tasks because difference in language processing dominated prediction ability in driving individual differences on this task.

Finally, successful performance on the PLAT requires participants to quickly process human action, an ability that was not required by any of the other tasks. Imaging studies have found evidence that when participants watch another person perform a task, the participants create a motor program for completing a task that is very similar to the motor program participants use when performing the task themselves (e.g., Flanagan & Johansson, 2003, but see Caramazza, Anzellotti, Strnad, & Lingnau, 2014 for an evaluation of this and another potential mechanism). If individual differences in people's ability to create a motor program dominated individual differences in prediction ability, low correlations between performance on the PLAT and the other prediction tasks might be expected. Overall, given that each of the four prediction tasks included in this study likely involve different mechanism (e.g., spatial processing, risk processing, language processing, and motor planning), in addition to prediction, individual differences in each of these other abilities may have resulted in the lack of correlations among the tasks observed in the current study.

There is, however, another possible explanation for the finding in the current study: Perhaps the tasks used in the current study did not require the intervention of an integrator or arbiter

because information from various sensory modalities did not conflict in these tasks. Maybe an integratory process only becomes active, and therefore only drives performance, when conflict resolution is necessary. For example, on the visual world task, both the visual information participants saw and the auditory information participants heard drove eye movements to the same object. There were no trials in which the auditory information directed eye movements to one object but visual information directed eye movements to a different object. If the tasks did not, in fact, require conflict resolution, the conclusion that there is no higher order prediction mechanism that drives performance on various types of prediction tasks is likely generally true, but there may also be a separate mechanism that arbitrates conflict across many types of tasks. In fact, there is suggestion in the literature that the anterior cingulate cortex may play a conflict resolution role across a wide variety of tasks (e.g., Botvinick, Cohen, & Carter), although there is also evidence that conflict resolution does not adequately describe the full function of the anterior cingulate cortex (e.g., Brown and Braver, 2005).

In addition, it is possible that there is a higher-order prediction mechanism that integrates information from various modalities, but that people, or at least the participants in the current study, do not differ in their ability to integrate information in order to make predictions. This potential lack of individual differences could result in low correlations among the tasks, because individual differences are necessary in order to find correlations. The fact that participants had a wide range of scores on each of the individual tasks used in this study at least suggests that there were individual differences on the tasks, but this does not necessarily mean that people differed in their ability to combine the information they gained from different modalities and use this integrated information to make predictions. However, it would likely be difficult to determine whether there are individual differences in the ability to integrate information from various

68

modalities and make predictions based on this combined information. For example, it is difficult

to find a study design that would be able to differentiate between true integration of information

from multiple sensory modalities versus separate representations of information from each

modality that are nevertheless all used to guide behavior. Imaging studies could potentially

identify networks that are activated by tasks that require integration of information, and these

networks could be compared to the networks activated by tasks that only require the use of a

single modality. If a multi-modality task requires brain regions that are not activated by separate

tasks that require each of the modalities included in the multi-modality task, this would provide

some evidence that performance on multi-modality tasks requires more than just the concurrent

activation of networks specialized for the various modalities. On the other hand, it seems

unlikely that participants would display such high variability on all of the individual tasks and no

variability on prediction integration. Furthermore, it seems likely that individual differences in

prediction integration (if such integration occurs) would exist, given that so many other higher

cognitive functions, including working memory, attention, and executive function, do show

individual differences. Therefore, the results of this study are more consistent with the absence of

a higher-order integratory mechanism than with an integratory process that does not vary across

individuals.

4.2 Measures of General Intelligence Predict Performance on Some Prediction Tasks

Although the prediction tasks did not load onto a single latent factor, the measures of general

intelligence did correlate with performance on some of the prediction tasks. In particular, higher

crystallized intelligence predicted worse performance on the PLAT and better performance on

the Iowa gambling task. In addition, higher fluid intelligence predicted better performance on the

weather prediction task and the Iowa gambling task. Finally, there was a negative relationship

between performance on the PLAT and processing speed. Although hypotheses about these relationships are post-hoc, the positive relationships between fluid intelligence and performance on the weather prediction and Iowa gambling tasks seem reasonable, as the ability to process new information and use that new information to complete tasks does seem related to fluid intelligence. In fact, previous research on the Iowa gambling task has found some support for the relationship between this task and general intelligence (e.g., Monterosso, Ehrman, Napier, O'Brien, & Childress, 2001), though other studies have found no relationship between performance on this task and IQ (e.g., Bechara et al., 2001; Brand, Recknor, Grabenhorst, & Bechara, 2007). In addition, the negative relationship between performance on the PLAT and processing speed (where a smaller processing speed score means faster responses) suggests that people who have faster processing speeds perform better on the PLAT, potentially because they are faster at predicting which object the actor is about to touch.

On the other hand, the positive relationship between the Iowa gambling task and crystallized intelligence was surprising because the ability to learn from new information, more than prior knowledge, would be expected to drive performance on this task. However, results from a previous study provide support for the current finding of the relationship between crystallized intelligence and performance on the Iowa gambling task: in a study of undergraduate students, participants who scored higher on the vocabulary measure also performed better on the Iowa gambling task (Damaree, Burns, & DeDonno, 2010. One potential explanation for the positive relationship is that participants with higher crystallized intelligence may have had more prior experience with the Iowa gambling task than participants with lower crystallized intelligence. For example, most of the participants in the current study were undergraduate students, and students who were further along in their education may both have had higher crystallized

intelligence and have taken more psychology classes in which they were exposed to the Iowa gambling task. Participants were not debriefed about their prior experience with the tasks, so it is not possible to determine whether prior experience drove the relationship. Similarly, the negative relationship between crystallized intelligence and performance on the PLAT was unexpected. Performance on the PLAT would be expected to be driven, at least in part, by schemas that include information about what typically happens in similar situations. Therefore, more prior knowledge about situations and people's typical behavior in given environments would be expected to improve, rather than impair, performance. However, the negative relationship between crystallized intelligence and the PLAT suggests that either prior schemas are not related to crystallized intelligence or that schemas are less important for successful performance on the PLAT than would be expected. Overall, replications are necessary to determine whether the relationships among the general intelligence constructs and the prediction tasks represent the true states of the relationships.

4.3 Posttraumatic Stress Disorder Did Not Predict Performance on Other Tasks

Although the model in Figure 15 that included PTSD as a predictor of the prediction tasks did provide a good fit for the data, PTSD severity did not predict performance on the general intelligence constructs or the prediction tasks. One possibility for the lack of significant relationships is that most of the participants in this study reported low levels of PTSD severity, with only a small number of participants in the middle to high range of severity, as can be seen in Figure 9. In fact, only 19 participants were above the current recommended cut-point of 33 on this measure. Therefore, there may not have been enough variability to see individual differences in performance on the other tasks based on PTSD severity scores. It is also possible that prediction ability is not impaired in PTSD. There have not been any studies investigating

performance on the visual world task, the Iowa gambling task, or the weather prediction task in people with PTSD, however, a recent study in our laboratory found that people with clinical levels of PTSD performed worse than controls on the PLAT (Eisenberg, Zacks, Rodebaugh, & Flores, in prep). Additional studies on the relationship between PTSD and various prediction tasks are necessary to determine whether the restricted range of PTSD severity scores in this study drove the lack of relationship between PTSD symptom severity and the prediction tasks.

4.4 Impact of Findings on Current Theories

4.4.1 Predictive Coding Model

As discussed in the introduction, the predictive coding model (Friston, 2005) suggests that predictions occur in a hierarchical fashion, with higher-order areas using past experience to make predictions and then sending those predictions to lower order areas. These lower order areas compare the predictions to sensory information from the environment. When there is a mismatch between the sensory information and the predictions, the lower-order areas send prediction error signals to the higher-order areas, which then either update their predictions or change sampling behavior so that incoming sensory information matches their predictions. Most of the research on this theory has studied individual systems, but Adams, Friston, and Bastos (2015) argue that because prediction errors can lead to both sensory and motor changes, the sensory and motor systems should be considered "a single active inference machine" (p. 100). This is a strong statement in support of a unified prediction mechanism in the brain, and Adams, Friston, and Bastos support this statement with findings that the laminar, topographic, and physiological characteristics of the sensory and motor cortices are quite similar.

However, given the results of the current study, it does not seem likely that Adams, Friston, and Bastos (2015) are correct about the existence of a single active inference machine, despite

the structural similarities of the systems. If all of the sensory and motor areas truly did act as a single unified system, performance on the tasks included in the current study, which all required activation of various sensory and motor areas, should have been correlated. Perhaps, instead of a unified inference machine, each system separately engages in feed-forward and feedback signals that allow for prediction within individual modalities. The structural similarities of the various cortical areas that Adams, Friston, and Bastos (2015) use to support their proposition of a unified inference machine may have developed as a parsimonious solution to developing complex cortical structures, but their structural similarities do not necessitate that all of the individual areas cohere into a unified prediction mechanism. Thus, while the current study does not provide evidence against the entirety of the predictive coding model of the human brain, the results reported here do suggest that a single unified prediction mechanism does not operate in the human brain.

4.4.2   Event Segmentation Theory

Event Segmentation Theory (EST; Zacks, Speer, Swallow, Braver, & Reynolds, 2007) provides a model of how people comprehend ongoing, dynamic activity. It proposes that people use their existing knowledge about typical situations (event schema) to create a representation of the current situation, and use this event model along with incoming sensory information to make predictions about what is going to happen next. When mismatches between predictions and incoming sensory information develop, the event model is updated to better represent the current situation. People perceive boundaries between events when this event model updating occurs (Zacks, Speer, Swallow, Braver, & Reynolds, 2007), and better perception of these event boundaries has been linked to increased memory for the events (Zacks, Speer, Vettel, & Jacoby, 2006).

EST postulates that a single event model that encompasses modalities is used to generate predictions about what will happen next. Although in any given situation, an event model could include information from only a single sensory modality, it should also be able to incorporate information from multiple modalities, as it is rare in everyday life for only a single modality to be relevant. For example, while watching an actor complete an everyday activity, an observer's event model would likely include visual information about the current activity, auditory information based on experience with the typical sounds generated by the activity, and motor information that would include the motor sequences necessary to generate a similar action. Therefore, EST proposes that multi-modality event models should be represented in some way in the brain, and that these event models should then be used to generate predictions about the future.

Most research on EST has studied event comprehension and prediction by having participants watch movies of everyday activities or read narratives about people engaging in activities, and it is possible that EST is limited to only these modalities. However, Zacks, Speer, Swallow, Braver, & Reynolds (2007) suggest that people use event models as a basis for prediction in every modality. For example, in the context of the current study, it might be possible to explain performance on the weather prediction task using EST: When people start the weather prediction task, they begin to learn the relationships between the geometric patterns on the cards and the outcome of the trial. Over time, they begin to create an event model that represents their current knowledge of these relationships. They make predictions about whether there will be rain or sun on the basis of this developing event model, and when their predictions are inaccurate, they update their event model to include the new information. Similar processes likely occur as people perform the Iowa gambling task and the visual world task. The PLAT,

though a novel task, is most similar to previous studies that have found evidence for EST. One study using this task found that healthy adult participants took longer to look at the target object when contact occurred around event boundaries than when contact occurred within an event (Eisenberg & Zacks, in prep), suggesting that predictions failed more often at event boundaries than within events. Performance on the PLAT, therefore, likely relies on the formation of event models that are updated at times of high prediction error.

If an integrated event model that is used to drive predictions does exist, one would expect that performance on the prediction tasks included in the current study would be correlated, as performance on all of these tasks would be based, at least in part, on the ability to form accurate event models; people who are better at forming event models should perform better on all of the tasks, and people who are worse at forming event models should perform more poorly. However, the low correlations among the four prediction tasks included in this study provide evidence against the existence of such an event model. This leaves open four possibilities: (1) multi-modal event models exist, but people do not vary in their ability to use event models to make predictions, (2) event models are only used when people are processing naturalistic activity, (3) event models consist of multiple separate representations from each of the separate sensory and motor systems, and (4) multi-modal event models exist, but individual differences are driven by modality specific prediction mechanisms that operate upstream of the event models.

First, it is possible that a unified event model does exist but people do not vary in their ability to use such an event model to make predictions. This possibility is very similar to the supposition, mentioned earlier, that people do not vary in their ability to integrate information in order to make predictions. Yet, as previously discussed, there was adequate variability across subjects on all of the tasks used in the current study. In particular, participants varied in their

ability to perform the PLAT, and performance on this task has previously been found to track the locations of event boundaries (Eisenberg and Zacks, in prep). Of the tasks included in the current study, the PLAT should most strongly involve the creation of event models, suggesting that participants were not identical in their ability to use event models to make predictions. Therefore, low variability cannot explain the results of the current study or preserve the concept of a unified event model.

Another possibility is that event models are limited to the domain of comprehending naturalistic activity. In this case, performance on tasks that all involve the comprehension of naturalistic activity should be correlated, even if performance relies on different sensory modalities. For example, performance on the PLAT should be correlated with performance on tasks in which participants listen to narratives of everyday activities (where prediction could be measured through predictive looking at arrays of images representing characters or objects that will soon be mentioned in the narrative) and on tasks in which participants read narratives of everyday activities (where prediction could be measured by pausing reading and asking participants to make explicit predictions about what will happen next). Although there have not been studies that test prediction performance on these different types of event comprehension tasks, the results of the current study and the principles of parsimony suggest that performance would not be correlated across different types of event comprehension tasks. Specifically, if unitary event models did exist, performance on all of the tasks used in this study should have relied on such an integrative event model, as it is unlikely that an integrative event model would be created solely during event comprehension tasks (e.g., the PLAT). There seems no reason for integrative event models to be used solely for tasks that involve the comprehension of naturalistic activities, when a similar mechanism could be used for many other types of tasks. In

fact, it is arguably the case that all of the tasks included in the current study were everyday events for the participants in the study, and therefore should have relied on integrative event models. Participants completed the tasks in and among all of the other daily activities in which they participated. Therefore, participants should have treated the tasks in the study as events, created unified event models of their perceptions of each activity, and then used these unified event models to make predictions that integrated information across modalities. In fact, it is very likely that if other people were asked to watch a movie that included a period of time in which an actor performed exactly the same tasks that were included in the present study, viewers would identify event boundaries at the beginning and completion of each task. Consequently, it seems unlikely that multi-modal event models that are used to make integrated predictions actually exist.

Therefore, a third possibility is that event models consist of separate representations from each sensory modality. Specifically, each of the sensory and motor systems may represent the state of the environment separately. Then, when a task requires the use of multiple sensory modalities, the necessary brain regions for each separate system may be activated in concert, which would result in the likely incorrect perception that active integration has occurred. For example, for the PLAT, visual areas and motor areas are likely both activated, and separate prediction processes within the respective areas operating in parallel could create an illusion of integration. A spike in prediction errors in one modality would lead to an updating of the representation of that modality, and people would experience this updating as an event boundary. If prediction errors spiked in multiple modalities at the same time (due to rapid changes in both visual and auditory information, for example) people might experience an even stronger perception of an event boundary.

The existence of separate representations rather than a unified event model is consistent with previous research evidence that prediction is more difficult around event boundaries (Zacks, Kurby, Eisenberg, & Haroutunian, 2011): prediction in the modality or modalities relevant to the rapid changes in the environment would be more difficult around event boundaries, though prediction in other modalities would not be affected. The existence of separate event models is also compatible with evidence that memory is updated at event boundaries (Swallow et al., 2011), as memory for visual information, for example, might be updated around event boundaries that are driven by visual changes in the environment, whereas memory for auditory information might not be updated in response to visual changes. This account is also in accord with previous research demonstrating that, at least during reading, components of event models can be updated independently. For example, Curiel & Radvansky (2014) found that spatial shifts and character shifts both slowed down reading time but the effects did not interact, suggesting that they did not influence one another. Though the results of this study were interpreted to be consistent with a unified event model that is updated incrementally rather than globally, the results of this study can also be interpreted as evidence against a unified event model, where representations of each type of shift are independent of one another and are used separately to make predictions. However, this account is less consistent with recent research on working memory updating. For example, Bailey & Zacks (2015) had participants read narratives that included shifts in characters and locations and answer recognition memory probes interspersed throughout the text. When these probes came after a shift in either character or location, people were slower to answer questions about either dimension, suggesting that participants primarily engaged in global event model updating. In addition, Radvansky, Tamplin, & Krawietz (2010) found that word pairs that were unrelated to the visual environment were less well remembered

after a visual and motor event boundary of walking through a doorway, which provides additional evidence for an integrated event model and against separate representations of information in each modality.

This leaves open the fourth possibility that unified event models exist and that people vary in their ability to use event models to make predictions, but individual differences in prediction are driven by modality-specific prediction mechanisms that operate upstream of a multi-modal event model. In this case, each brain system would make predictions independently of one another, and spikes in prediction error in any one modality could cause the unified event model to update its representation of the current situation. This would mean that predictions might not be correlated across task modality if prediction ability within each brain system differs within individuals. For example, if a person makes very accurate predictions when tasks require predictions based on visual information but experiences difficulty making predictions using auditory information, that person would display quite different performance on prediction tasks requiring each of the modalities. The lack of correlations among the four tasks used in the current study is consistent with this possibility.

If it is indeed the case that individual differences are driven by modality-specific prediction mechanisms and that a multi-modal event model is updated incrementally whenever there is a spike in prediction error within any modality, some changes must be made to the EST model. Specifically, instead of sensory information entering a single perceptual processing node, sensory information from each modality would enter a perceptual processing node specific to the modality. Each perceptual processing node would receive information from a unified event model and would use this information to make predictions. Separate error processing mechanisms would monitor these predictions and when errors are signaled from any of these

error processing mechanisms, people would experience a subjective event boundary and the unified event model would update to capture the changes in the environment. The process would then repeat until the next error signal causes the event model to update again. For example, if visual information were changing rapidly, predictions about future visual information would likely be incorrect. This would cause the event model to update to better represent the new state of the visual information. Once the new information is integrated into the event model, relevant information from the event model would be used by all of the brain systems to continue making predictions. If there were changes in multiple modalities at the same time, people would still experience a single event boundary, but the event model would be updated to capture the changed information from all of the relevant modalities. Figure 16 provides a potential representation of such a model of event comprehension and prediction.
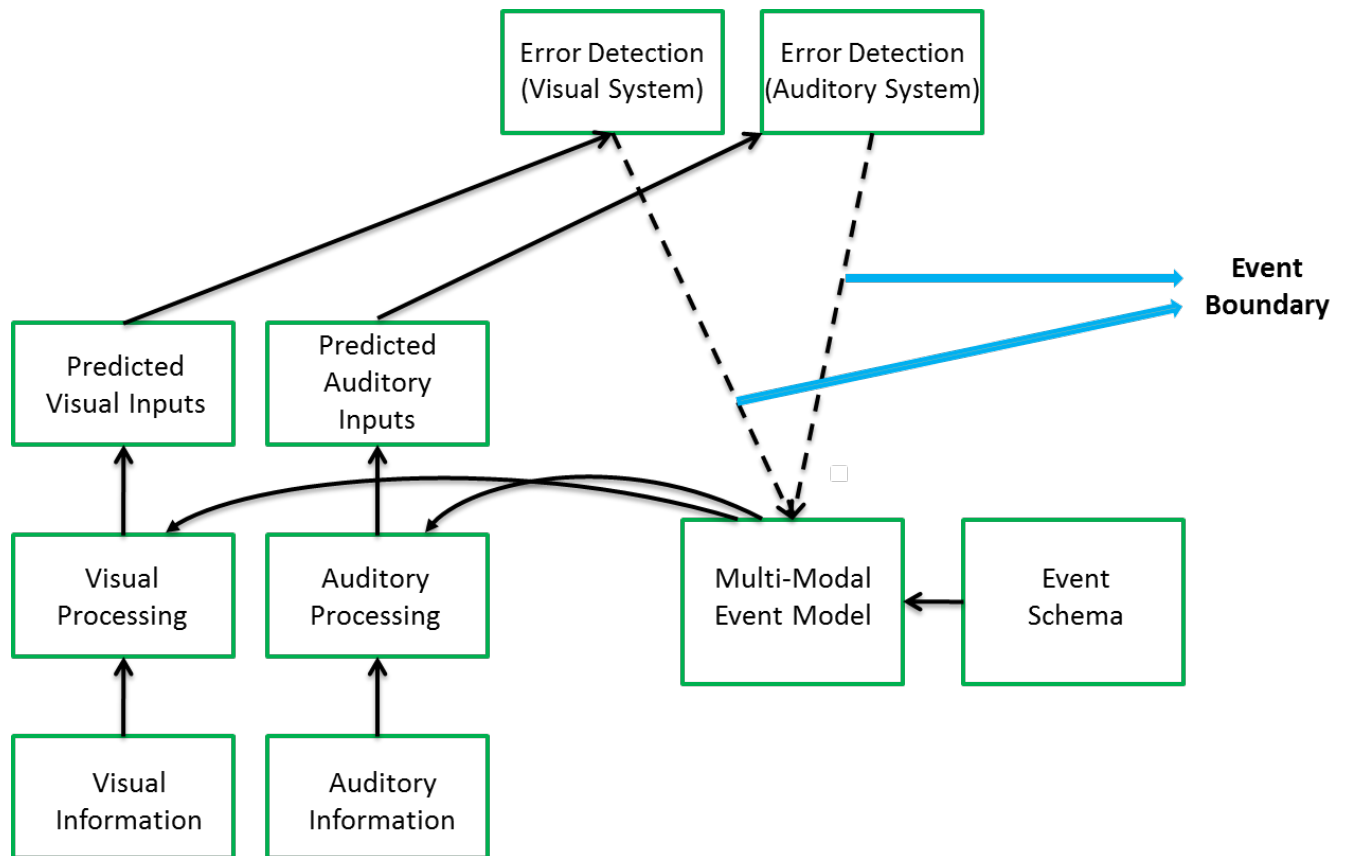
Figure 16. Suggested model of Event Segmentation Theory if individual differences are driven by modality-specific prediction mechanisms. In this case, error signals from any of the modalities would lead to the perception of a subjective event boundary and would reset a unified event model. Information from the unified event model would then be used by separate perceptual processing systems specific to each modality to make new predictions. Only the visual and auditory systems are represented here for the sake of simplicity, but there are likely many more separate systems that are involved in making predictions.

4.5. Limitations of the Current Study

There are some limitations of the current study that are important in interpreting the results. First, the reliability of most of the tasks included in the current study was relatively high, except for the split-half reliability of the visual world task ($r = 0.53$), the Letter Sets task ($r = 0.44$) and the Paper Folding task ($r = 0.67$). There is therefore some chance that the low reliability of these tasks prevented them from correlating with the other tasks included in the current study.

However, other methods of scoring the visual world task resulted in even lower split-half reliability scores. Specifically, when growth curves for looking time were calculated separately for predictive and unpredictive trials, and the slopes from the unpredictive trials were subtracted from the slopes for the predictive trials to obtain a difference score, the split half reliability was 0.16. Similarly, when the cumulative looking time on unpredictive trials was subtracted from the cumulative looking time on predictive trials to obtain a difference score, the split-half reliability was 0.18. One potential explanation for the low split-half reliability is the relatively small number of trials participants completed during this task; participants only completed 24 experimental trials in this task, half of which were unpredictive sentences and half of which were predictive sentences. This means that for the split-half reliability testing, each half of the data only included six trials of each type. Therefore, the split-half reliability found in this study may not replicate in other studies, and a study with a larger number of trials would be necessary to determine the actual reliability of this task. Nevertheless, it is unlikely that higher reliability of the visual world task would have dramatically changed the results of the current study. None of the other prediction tasks were correlated with one another, and the correlations between performance on the visual world task and the other prediction tasks were so low that even if a portion of the signal from the visual world task were correlated with performance on the other tasks, it still is unlikely that the noise masked more than a small correlation among the tasks. Similarly, the low split-half reliability for the Letter Sets and Paper Folding tasks was likely due, in part, to the small number of trials in each task. A study using larger numbers of items for each of these tasks would be necessary to determine whether the low split-half reliability is an intrinsic feature of these tasks. However, it is unlikely that the low split-half reliabilities of the Letter Sets task and the Paper Folding task dramatically affected the results of the current study.

All three of the fluid intelligence tasks loaded strongly on a fluid intelligence latent factor, meaning that the fluid intelligence latent factor likely captured the signal present in all of the tasks without being significantly affected by the noise in the tasks.

In addition, although I examined the relationships between PTSD severity and performance on the prediction tasks in the study, there were few participants with high scores on the PTSD scale. This lack of variability may have obscured the true relationships between PTSD severity and prediction performance. As mentioned above, Eisenberg, Zacks, and Flores (in prep) examined performance on the PLAT in a clinical sample of people diagnosed with PTSD and found that people with PTSD performed more poorly on the task than control participants without PTSD. However, no relationship between PTSD severity and performance on the PLAT was found in the current study. Similarly, previous studies have found relationships between PTSD and performance on tasks of general cognitive functioning (e.g., Vasterling et al., 2002; Bremner, Vermetten, Afzal, & Vythilingam, 2004), whereas there were no significant relationships between PTSD severity and crystallized or fluid intelligence in the current study; however, the previous studies included more participants with clinical levels of PTSD. Therefore, additional research examining the relationship between PTSD severity and performance on various types of prediction tasks is necessary to determine whether PTSD impacts performance on prediction tasks other than the PLAT.

Furthermore, as discussed earlier, it is possible that the low correlations among the prediction tasks could have been due to differences in methods across the tasks, as the tasks all required different abilities in addition to prediction. One potential method for controlling for potential individual difference in abilities other than prediction would be to test participants' performance on tasks that require each of the non-prediction abilities and then control for these

83

individual differences in the final analyses. For example, participants could complete tasks that are relatively pure measures of spatial ability, risk processing, language processing, and motor planning, and scores on these measures could then be entered as covariates in analyses of the correlations among the prediction tasks.

In addition to potentially requiring different abilities, the tasks also differed in the directions given to participants for each task. For example, for both the weather prediction and the Iowa gambling tasks, the directions to participants were very explicit and explained the true nature of the task: Participants were told to make predictions about whether there would be rain or sun in the weather prediction task, and participants were told to choose decks in a way that made them the most money in the Iowa gambling task. On the other hand, for both the PLAT and the visual world task, the directions did not explain the ultimate goal for the task: For the PLAT, participants were simply told to pay attention to the movie, and for the visual world task, participants were told to respond based on whether the sentence that they heard applied to any of the images on the screen, without any reference to the predictive nature of the task. It is possible that if participants had been told to actively make predictions during all of the tasks, that performance across the tasks would have been more similar. On the other hand, it is highly likely that people make predictions during everyday life without realizing that they are doing so, and it is therefore likely that people made predictions while they performed the visual world task and the PLAT as well. In addition, the pairs of tasks that used similar directions were not correlated with one another, again suggesting that differences in directions across the tasks cannot explain the low correlations among the tasks.

Another source of difference across the tasks is how performance was measured for the four tasks. Performance on the weather prediction and Iowa gambling tasks was measured using the

accuracy of responses, whereas performance on the PLAT and the weather prediction task was measured using oculometric data. Accuracy data only provides a measure of the final decision a participant makes about a particular item, whereas eye tracking data can be much more sensitive, providing information throughout the decision making process. It is possible that if eye tracking data had been collected while participants performed the weather prediction and Iowa gambling tasks, the increased sensitivity would have allowed correlations among the tasks to emerge.

In addition, especially for the Iowa gambling task, an accuracy measure may not have been the best choice for examining individual differences on this task. While individual variation certainly existed within the data set, many previous studies using this task used measures of skin conductance to determine how well participants had learned the task and when they began making predictions based on their knowledge. For example, Bechara, Damasio, Tranel, & Damasio (1997) identified a hunch period, in which participants were unable to report that they knew some decks were bad, but experienced an increased skin conductance response when choosing from the bad decks. Predictions during this hunch period might more closely resemble the predictions made in the PLAT and visual world task. Although skin conductance responses have not typically been used in the weather prediction task, a similar effect might exist in which participants experience an increased skin conductance response when making an incorrect choice, while still experiencing the feeling of guessing on the trial. If eye tracking were combined with skin conductance measures, it might be possible to use the time spent looking at the correct choice during the hunch phase as a measure of prediction performance on the task, and that measure might correlate more strongly with performance on the PLAT and the visual world task. However, if the measurement modality were the driver of the lack of correlations among the four tasks, one would expect that the tasks with more similar measurement modalities

would be most highly correlated with one another, which was not the case; all of the tasks had very low correlations with one another, and most of the correlations among the pairs of tasks with the same measurement modalities were actually slightly lower than the correlations among pairs of tasks with different measurement modalities. Therefore, measurement modality is unlikely to have been the only reason performance on these tasks was uncorrelated.

Chapter 5: Conclusion

In this study, performance on four different types of prediction tasks was uncorrelated, providing evidence against a single higher-order prediction mechanism. While the results of this study suggest that existing theories about prediction require some alteration, particularly the predictive coding theory (Friston, 2005) and Event Segmentation Theory (Zacks et al., 2007), the results are not directly in opposition with these theories. The predictive coding theory does not necessitate a unified prediction mechanism, though such a mechanism was proposed due to the similar anatomical structure of many systems in the brain (Adams, Friston, & Bastosm 2015). In fact, the results of this study are in accord with the hierarchical predictive brain proposed by Friston (2005), with systems that process information from each modality sending predictive signals from higher order areas to lower order areas and vice versa. In addition, the results of this study certainly do not provide evidence against the whole of Event Segmentation Theory, but rather suggest that certain elements of the theory need revision. Overall, instead of an integratory prediction mechanism, it appears most likely that predictions are formed and acted upon separately within each sensory modality and that the resulting behavior creates an illusion of integration. This may be represented in the brain by the activation of separate but overlapping brain regions that, when activated at the same time by different components of a task, creates the appearance of an integratory prediction mechanism. Finally, these results can also inform future studies on prediction, including those investigating prediction ability in psychopathology: Attempting to generalize from performance on one type of prediction task to another will likely result in only erroneous *predictions*.

References

Adams, R. A., Friston, K. J., & Bastos, A. M. (2015). Active inference, predictive coding, and cortical architecture. In Casanova, M. F. & Opris, I. (Eds.), *Recent Advances on the Modular Organization of the Cortex* (pp. 97–121). Netherlands: Springer. doi:10.1007/978-94-017-9900-3

Apicella, P., Scarnati, E., Ljungberg, T., & Schultz, W. (1992). Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *Journal of Neurophysiology*, *68*, 945–960. doi:1432059

Aron, A. R., Shohamy, D., Clark, J., Myers, C., Gluck, M. A., & Poldrack, R. A. (2004). Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*, *92*, 1144–1152. doi:10.1152/jn.01209.2003

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*, 247–264. doi:10.1016/S0010-0277(99)00059-1

Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*, 502–518. doi:10.1016/j.jml.2006.12.004

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. doi:10.1146/annurev-psych-120710-100422

Bailey, H. R. & Zacks, J. M. (2015). Situation model updating in young and older adults: Global versus incremental mechanisms. *Psychology and Aging, 30*, 232-244. doi: http://dx.doi.org/10.1037/a0039081

Barch, D. M., Braver, T. S., Akbudak, E., Conturo, T., Ollinger, J., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: Effects of response modality and processing domain. *Cerebral Cortex, 11*, 837–848. doi:10.1093/cercor/11.9.837

Bors, D. A., & Forrin, B. (1995). Age, speed of information processing, recall, and fluid intelligence. *Intelligence*, *20*(3), 229–248. http://doi.org/10.1016/0160-2896(95)90009-8

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*, 539–546. doi:10.1016/j.tics.2004.10.003

Brand, M., Recknor, E. C., Grabenhorst, F., & Bechara, A. (2007). Decisions under ambiguity and decisions under risk: Correlations with executive functions and comparisons of two different gambling tasks with implicit and explicit rules. *Journal of Clinical and Experimental Neuropsychology*, *29*(1), 86–99. http://doi.org/10.1080/13803390500507196

Brown, J. W., & Braver, T. S. (2005). Learned predictions of error likelihood in the anterior cingulate cortex. *Science*, *307*, 1118–1121. doi:10.1126/science.1105783

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01

Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia*, *39*(4), 376–389. http://doi.org/10.1016/S0028-3932(00)00136-6

Bendixen, A., SanMiguel, I., & Schröger, E. (2012). Early electrophysiological indicators for

    predictive processing in audition: A review. *International Journal of Psychophysiology*, *83*,

    120–131. doi:10.1016/j.ijpsycho.2011.08.003

Bledowski, C., Prvulovic, D., Goebel, R., Zanella, F. E., & Linden, D. E. J. (2004). Attentional

    systems in target and distractor processing: A combined ERP and fMRI study. *NeuroImage*,

    22(2), 530–540. doi:10.1016/j.neuroimage.2003.12.034

Bledowski, C., Prvulovic, D., Hoechstetter, K., Scherg, M., Wibral, M., Goebel, R., & Linden,

    D. E. J. (2004). Localizing P300 generators in visual target and distractor processing: A

    combined event-related potential and functional magnetic resonance imaging study. *The*

    Journal of Neuroscience, *24*, 9353–9360. doi:10.1523/JNEUROSCI.1897-04.2004

Blevins, C. A., Weathers, F. W., Davis, M. T., Witte, T. K., & Domino, J. L. (2015). The

    Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Development and initial

    psychometric evaluation. *Journal of Traumatic Stress, 28*, 489-498. doi: 10.1002/jts.22059

Bremner, J. D., Vermetten, E., Afzal, N., & Vythilingam, M. (2004). Deficits in verbal

    declarative memory function in women with childhood sexual abuse-related Posttraumatic

    Stress Disorder. *The Journal of Nervous and Mental Disease*, *192*, 643–649.

    doi:10.1097/01.nmd.0000142027.52893.c8

Brouwer, H., & Hoeks, J. C. J. (2013). A time and place for language comprehension: Mapping

    the N400 and the P600 to a minimal cortical network. *Frontiers in Human Neuroscience*, *7*,

    1-12. doi:10.3389/fnhum.2013.00758

Calamia, M., Markon, K., & Tranel, D. (2013). The robust reliability of neuropsychological

    measures: meta-analyses of test-retest correlations. *The Clinical Neuropsychologist*, *27*,

    1077–1105. http://doi.org/10.1080/13854046.2013.809795

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions.
Developmental Science, *15*, 292–298. doi:10.1111/j.1467-7687.2011.01127.x

Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror
neurons: A critical assessment. *Neuroscience, 37*, 1–15. http://doi.org/10.1146/annurev-
neuro-071013-013950

Chen, J., Hale, S. & Myerson, J. (2007). Predicting the size of individual and group differences
on speeded cognitive tasks. *Psychonomic Bulletin and Review*, *14*, 534–541.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain
response to morphosyntactic violations. *Language and Cognitive Processes*, *13*, 21–58.
doi:10.1080/016909698386582

Curiel, J. M., & Radvansky, G. a. (2014). Spatial and character situation model updating.
Journal of Cognitive Psychology, *26*(2), 205–212.
http://doi.org/10.1080/20445911.2013.879590

Czigler, I., Balázs, L., & Pató, L. G. (2004). Visual change detection: Event-related potentials are
dependent on stimulus location in humans. *Neuroscience Letters*, *364*, 149–153.
doi:10.1016/j.neulet.2004.04.048

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and
reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a Neural System for Error
Detection and Compensation. *Psychological Science, 5*, 303–305. doi:10.1111/j.1467-
9280.1994.tb00630.x

Demaree, H. A., Burns, K. J., & DeDonno, M. A. (2010). Intelligence, but not emotional intelligence, predicts Iowa Gambling Task performance. *Intelligence*, *38*(2), 249–254. http://doi.org/10.1016/j.intell.2009.12.004

Dikker, S., & Pylkkänen, L. (2013). Predicting language: MEG evidence for lexical preactivation. *Brain and Language*, *127*, 55–64. doi:10.1016/j.bandl.2012.08.004

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*, 145–177. doi:10.1016/j.cognition.2003.11.002

Ehlers, A., & Clark, D. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, *38*, 319–345. doi:10.1016/S0005-7967(99)00123-0

Eisenberg, M. L., Zacks, J. M., & Flores, S. Predictive Looking During Viewing of Naturalistic Activity. Manuscript in preparation, Department of Psychology, Washington University in St. Louis, St. Louis, MO.

Eisenberg, M. L., Zacks, J. M., Rodebaugh, T. L., & Flores, S. Prediction of Future Activity Impaired in Posttraumatic Stress Disorder. Manuscript in preparation, Department of Psychology, Washington University in St. Louis, St. Louis, MO.

Ekstrom, R. B., French, J.W., Harman, H. H., & Dermen, D. (1976). Kit of Factor-Referenced Cognitive Tests. Educational Testing Service. Princeton, New Jersey.

Elsner, C., Falck-Ytter, T., & Gredebäck, G. (2012). Humans anticipate the goal of other people's point-light actions. *Frontiers in Psychology*, *3*, 1–7. doi:10.3389/fpsyg.2012.00120

Ernst, M., Nelson, E. E., McClure, E. B., Monk, C. S., Munson, S., Eshel, N., … Pine, D. S. (2004). Choice selection and reward anticipation: An fMRI study. *Neuropsychologia*, *42*, 1585–1597. doi:10.1016/j.neuropsychologia.2004.05.011

Falck-Ytter, T., Gredebäck, G., & von Hofsten, C. (2006). Infants predict other people's action goals. *Nature Neuroscience*, *9*, 878–879. doi:10.1038/nn1729

Falconer, E., Bryant, R., Felmingham, K. L., Kemp, A. H., Gordon, E., Peduto, A., … Williams, L. M. (2008). The neural networks of inhibitory control in posttraumatic stress disorder. *Journal of Psychiatry & Neuroscience*, *33*, 413–22. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2527717&tool=pmcentrez&rendertype=abstract

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*, 447–455. doi:10.1016/0013-4694(91)90062-9

Flanagan, J. R., & Johansson, R. S. (2003). Action plans used in action observation. *Nature*, *424*, 769–771. doi:10.1038/nature01861

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*, 815–836. doi:10.1098/rstb.2005.1622

Fruchter, J., Linzen, T., Westerlund, M., & Marantz, A. (2015). Lexical Preactivation in Basic Linguistic Phrases. *Journal of Cognitive Neuroscience*, 1–24. doi:10.1162/jocn

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390. doi:10.1111/j.1467-9280.1993.tb00586.x

Gluck, M. A., Shohamy, D., & Myers, C. (2002). How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learning & Memory*, *9*, 408–418. doi:10.1101/lm.45202

Gonzalez, R., Wardle, M., Jacobus, J., Vassileva, J., & Martin-Thormeyer, E. M. (2010). Influence of procedural learning on Iowa Gambling Task performance among HIV+ individuals with history of substance dependence. *Archives of Clinical Neuropsychology*, *25*(1), 28–38. http://doi.org/10.1093/arclin/acp094

Haenschel, C., Vernon, D. J., Dwivedi, P., Gruzelier, J. H., & Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *The Journal of Neuroscience*, *25*, 10494–10501. doi:10.1523/JNEUROSCI.1227-05.2005

Hayhoe, M. M., Shrivastava, A., Mruczek, R., & Pelz, J. B. (2003). Visual memory and motor planning in a natural task. *Journal of Vision*, *3*, 49–63. doi:10.1167/3.1.6

Hikosaka, O., Sakamoto, M., & Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, *61*(4), 814–832.

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*(18), 2481–2484. http://doi.org/10.1097/01.wnr.0000099601.41403.a5

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*, 1–55. doi:10.1080/10705519909540118

Herzog, W., Boomsma, A., & Reinecke, S. (2007). The model-size effect on traditional and modified tests of covariance structures. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*, 361–390. doi:10.1080/10705510701301602

Hikosaka, O., Sakamoto, M., & Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, *61*, 814–832.

Huettel, S. A, Song, A. W., & McCarthy, G. (2005). Decisions under uncertainty: Probabilistic context influences activation of prefrontal and parietal cortices. *The Journal of Neuroscience*, *25*, 3304–3311. doi:10.1523/JNEUROSCI.5070-04.2005

Hunnius, S., & Bekkering, H. (2010). The early development of object knowledge: A study of infants' visual anticipations during action observation. *Developmental Psychology*, *46*, 446–454. doi:10.1037/a0016543

Johansson, R. S., Westling, G., Bäckström, A, & Flanagan, J. R. (2001). Eye-hand coordination in object manipulation. *The Journal of Neuroscience, 21*, 6917–6932.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133–156. doi:10.1016/S0749-596X(03)00023-8

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of Experimental Psychology: General*, *133*, 189–217. doi:10.1037/0096-3445.133.2.189

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning & Memory, 1*, 106–120. doi:10.1101/lm.1.2.106

Kok, P., Jehee, J. F. M., & de Lange, F. P. (2012). Less Is More: Expectation Sharpens Representations in the Primary Visual Cortex. *Neuron*, *75*(2), 265–270. http://doi.org/10.1016/j.neuron.2012.04.034

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and

    clinical neuroscience: Theoretical and methodological considerations. *Biological*

    Psychology, *74*, 1–19. doi:10.1016/j.biopsycho.2006.06.001

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400

    component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*,

    621–647. doi:10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect

    semantic incongruity. *Science, 207*, 203–205. doi:10.1126/science.7350657

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control

    of activities of daily living. *Perception*, *28*, 1311–1328. doi:10.1068/p2935

Linnman, C., Zeffiro, T. A, Pitman, R. K., & Milad, M. R. (2011). An fMRI study of

    unconditioned responses in post-traumatic stress disorder. *Biology of Mood & Anxiety*

    Disorders, *1*(8), 1–12. doi:10.1186/2045-5380-1-8

Liu, X., Hairston, J., Schrier, M., & Fan, J. (2011). Common and distinct networks underlying

    reward valence and processing stages: A meta-analysis of functional neuroimaging studies.

    Neuroscience and Biobehavioral Reviews, *35*, 1219–1236.

    doi:10.1016/j.neubiorev.2010.12.012

Luft, C. D. B., Meeson, A., Welchman, A. E., & Kourtzi, Z. (2015). Decoding the future from

    past experience: learning shapes predictions in early visual cortex. *Journal of*

    Neurophysiology, *113*(9), 3159–71. http://doi.org/10.1152/jn.00753.2014

MacLeod, C. M., & MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An

    integrative review. *Psychological Bulletin*, *109*, 163–203. doi:10.1037/0033-2909.109.2.163

Mitchell, R. L. C. (2005). The BOLD response during Stroop task-like inhibition paradigms: Effects of task difficulty and task-relevant modality. *Brain and Cognition*, *59*, 23–37. doi:10.1016/j.bandc.2005.04.001

Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., & Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addiction*, *96*(12), 1825–1837. http://doi.org/10.1046/j.1360-0443.2001.9612182512.x

Mustovic, H., Scheffler, K., Di Salle, F., Esposito, F., Neuhoff, J. G., Hennig, J., & Seifritz, E. (2003). Temporal integration of sequential auditory events: Silent period in sound pattern activates human planum temporale. *NeuroImage*, *20*, 429–434. doi:10.1016/S1053-8119(03)00293-3

National Center for Posttraumatic Stress Disorder. (n.d.) *PTSD Checklist for DSM-5 (PCL-5)*. Retrieved July 2, 2016, from http://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp

Osterhout, L., & Holcomb, P. J. (1993). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, *31*, 785–806.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, *14*, 283–317. doi:10.1080/016909699386310

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2014). The development of a short domain-general measure of working memory capacity. *Behavior Research Methods*. 1-13. doi:10.3758/s13428-014-0543-2

Paavilainen, P., Arajärvi, P., & Takegata, R. (2007). Preattentive detection of nonsalient contingencies between auditory features. *Neuroreport, 18*, 159-163.

Paul, S. M. (1985). The Advanced Raven's Progressive Matrices: Normative data for an

    American university population and an examination of Spearman's g. *Journal of*

    Experimental Education, *54*, 95–100. http://doi.org/10.1080/00220973.1986.10806404

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical*

    Neurophysiology, *118*, 2128–48. doi:10.1016/j.clinph.2007.04.019

Polich (2012). Neuropsychology of P300. In Luck, S. J. & Kappenman, E. S. (Eds.), *The Oxford*

    handbook of event-related potential components (159-188). New York, NY: Oxford

    University Press.

Polich, J., & Margala, C. (1997). P300 and probability: Comparison of oddball and single-

    stimulus paradigms. *International Journal of Psychophysiology*, *25*, 169–176.

    doi:10.1016/S0167-8760(96)00742-8

Radvansky, G. A., Tamplin, A. K., & Krawietz, S. A. (2010). Walking through doorways causes

    forgetting: Environmental integration. *Psychonomic Bulletin & Review*, *17*(6), 900–904.

    http://doi.org/10.3758/PBR.17.6.900

Ramnani, N., Elliott, R., Athwal, B. S., & Passingham, R. E. (2004). Prediction error for free

    monetary reward in the human prefrontal cortex. *NeuroImage*, *23*, 777–786.

    doi:10.1016/j.neuroimage.2004.07.028

Raij, T., McEvoy, L., Mäkelä, J. P., & Hari, R. (1997). Human auditory cortex is activated by

    omissions of auditory stimuli. *Brain Research*, *745*, 134–143. doi:10.1016/S0167-

    8760(97)85548-1

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional

    interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–

    87. http://doi.org/10.1038/4580

Raven, J. C. (1990) Advanced Progressive Matrices: Sets I, II. Oxford Univ Press, Oxford.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, *28*, 164–171. doi:10.1027/1015-5759/a000123

Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *NeuroImage*, *12*, 14–19. doi:10.1006/nimg.2000.0591

Salthouse, T. (1993). Speed and knowledge as determinants of adult age differences in verbal tasks. *Journal of Gerontology: Psychological Sciences*, *48*, P29–36.

Salthouse, T. A. (2001). Structural models of the relations between age and measures of cognitive functioning. *Intelligence*, *29*(2), 93–115. http://doi.org/10.1016/S0160-2896(00)00040-4

Salthouse, T. & Babcock, R. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*, 763–777.

Schultz, W., Tremblay, L., & Hollerman, J. R. (1998). Reward prediction in primate basal ganglia and frontal cortex. *Neuropharmacology*, *37*, 421–429. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9704983

Service, E., Helenius, P., Maury, S., & Salmelin, R. (2007). Localization of syntactic and semantic brain responses using magnetoencephalography. *Journal of Cognitive Neuroscience*, *19*, 1193–1205. doi:10.1162/jocn.2007.19.7.1193

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial

    thinking and language processing: An individual differences approach. *Journal of*

    Experimental Psychology: General, *125*, 4–27.

Shin, L. M., & Liberzon, I. (2010). The neurocircuitry of fear, stress, and anxiety disorders.

    Neuropsychopharmacology, *35*, 169–91. doi:10.1038/npp.2009.83

Shohamy, D., Myers, C. E., Grossman, S., Sage, J., Gluck, M. A., & Poldrack, R. A. (2004).

    Cortico-striatal contributions to feedback-based learning: Converging data from

    neuroimaging and neuropsychology. *Brain*, *127*, 851–859. doi:10.1093/brain/awh100

Stam, R. (2007). PTSD and stress sensitisation: A tale of brain and body Part 1: Human studies.

    Neuroscience and Biobehavioral Reviews, *31*, 530–57.

    doi:10.1016/j.neubiorev.2006.11.010

Swaab, T. Y., Ledoux, K, Camblin, C. C., & Boudewyn (2012). Language-related ERP

    components. In Luck, S. J. & Kappenman, E. S. (Eds.), *The Oxford handbook of event-*

    related potential components (397-439). New York, NY: Oxford University Press.

Swallow, K. M., Barch, D. M., Head, D., Maley, C. J., Holder, D., & Zacks, J. M. (2011).

    Changes in events alter how people remember recent information. *Journal of Cognitive*

    Neuroscience, *23*(5), 1052–64. http://doi.org/10.1162/jocn.2010.21524

Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch negativity in the visual

    modality. *Neuroreport*, *10*, 3363–3367.

Tanaka, S. C., Doya, K., Okada, G., Ueda, K., Okamoto, Y., & Yamawaki, S. (2004). Prediction

    of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature*

    Neuroscience, *7*, 887–893. doi:10.1038/nn1279

Trapp, S., & Bar, M. (2015). Prediction, context, and competition in visual recognition. *Annals of the New York Academy of Sciences*, *1339*, 190–198. doi:10.1111/nyas.12680

Tse, C.-S., Balota, D. A., Yap, M. J., Duchek, J. M. & McCabe, D. P. (2010). Effects of healthy aging and early stage dementia of the Alzheimer's type on components of response time distributions in three attention tasks. *Neuropsychology*, *24*, 300–315.

Turner, M. L. & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, *28*, 127–154.

Ullsperger, M., & von Cramon, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, *14*, 1387–1401. doi:10.1006/nimg.2001.0935

van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182. doi:10.1006/jmla.1999.2641

van Berkum, J. J. A, Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*, 443–467. doi:10.1037/0278-7393.31.3.443

van Wingen, G., Geuze, E., Caan, M. W. A., Kozicz, T., Olabarriaga, S. D., Denys, D., … Fernández, G. (2012). Persistent and reversible consequences of combat stress on the mesofrontal circuit and cognition. *Proceedings of the National Academy of Sciences*, *109*, 15508–15513. doi:10.1073/pnas.1206330109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1206330109

Vasterling, J. J., Brailey, K., Constans, J. I., & Sutker, P. B. (1998). Attention and memory dysfunction in posttraumatic stress disorder. *Neuropsychology*, *12*, 125–133. doi:10.1037//0894-4105.12.1.125

Vasterling, J. J., Duke, L. M., Brailey, K., Constans, J. I., Allain, A. N., & Sutker, P. B. (2002). Attention, learning, and memory performances and intellectual resources in Vietnam veterans: PTSD and no disorder comparisons. *Neuropsychology*, *16*, 5–14. doi:10.1037//0894-4105.16.1.5

Wasserman, J. I., Barry, R. J., Bradford, L., Delva, N. J., & Beninger, R. J. (2012). Probabilistic classification and gambling in patients with schizophrenia receiving medication: Comparison of risperidone, olanzapine, clozapine and typical antipsychotics. *Psychopharmacology*, *222*(1), 173–183. http://doi.org/10.1007/s00213-011-2634-4

Weathers, F.W., Blake, D.D., Schnurr, P.P., Kaloupek, D.G., Marx, B.P., & Keane, T.M. (2013). The Life Events Checklist for DSM-5 (LEC-5). Instrument available from the National Center for PTSD at www.ptsd.va.gov

Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., & Schnurr, P.P. (2013). The PTSD Checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD at www.ptsd.va.gov.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale—Fourth Edition*. San Antonio, TX: Pearson Assessment.

Wei, J. H., Chan, T. C., & Luo, Y. J. (2002). A modified oddball paradigm "cross-modal delayed response" and the research on mismatch negativity. *Brain Research Bulletin*, *57*, 221–230. doi:10.1016/S0361-9230(01)00742-0

Weinberg, A., Dieterich, R., & Riesel, A. (2015). Error-related brain activity in the age of RDoC: A review of the literature. *International Journal of Psychophysiology*. doi:10.1016/j.ijpsycho.2015.02.029

Zacks, J. M. (2004). Using movement and intentions to understand simple events. *Cognitive Science, 28*, 979-1008.

Zacks, J. M., Kurby, C. a., Eisenberg, M. L., & Haroutunian, N. (2011). Prediction Error Associated with the Perceptual Segmentation of Naturalistic Events. *Journal of Cognitive Neuroscience*, *23*, 4057–4066. http://doi.org/10.1162/jocn_a_00078

Zacks, J. M., Speer, N. K., & Reynolds, J. R. (2009). Segmentation in reading and film comprehension. *Journal of Experimental Psychology: General, 138*, 307-327. doi: 10.1037/a0015305

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*, 273-293. doi: 10.1037/0033-2909.133.2.273

Zacks, J. M., Speer, N. K., Vettel, J. M., & Jacoby, L. L. (2006). Event understanding and memory in healthy aging and dementia of the Alzheimer type. *Psychology and Aging, 21*, 466-482. doi: 10.1037/0882-7974.21.3.466