

DEVELOPING AND VALIDATING A COMBINED ATTENTION
SYSTEMS TEST

by

Michael A. Lawrence

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

at

Dalhousie University
Halifax, Nova Scotia
August 2018

© Copyright by Michael A. Lawrence, 2018

To my wife, whose strength and support are a constant inspiration.

Table of Contents

List of Tables	vi
List of Figures	vii
Abstract	xi
List of Abbreviations and Symbols Used	xii
Acknowledgements	xiii
Chapter 1 Introduction	1
1.1 Measurement of attention	7
1.2 Initial explorations to improve the measurement of attention	10
1.3 The Combined Attention Systems Test	14
Chapter 2 Experiment 1: Young Adults	19
2.1 Methods	19
2.1.1 Participants	19
2.1.2 Materials	20
2.1.3 Procedure	20
2.2 Results	21
2.2.1 Data pre-processing	21
2.2.2 Statistical modelling	23
2.2.3 Primary effects and interactions	24
2.2.4 Effects of Session	27
2.2.5 Correlations	29
2.2.6 Reliabilities	32
2.2.7 Effect sizes	32
2.3 Discussion	35
Chapter 3 Experiment 2: Children	37
3.1 Methods	37
3.1.1 Participants	37
3.1.2 Materials	38
3.1.3 Procedure	38

3.2	Results	39
3.2.1	Data pre-processing	39
3.2.2	Statistical modelling	39
3.2.3	Primary effects and interactions	39
3.2.4	Effects of Session	41
3.2.5	Correlations	41
3.2.6	Reliabilities	47
3.2.7	Effect sizes	47
3.3	Discussion	49
Chapter 4	Development of attention as observed by the CAST	50
4.1	Results	50
4.1.1	Data pre-processing	50
4.1.2	Statistical modelling	50
4.1.3	Effects of age	50
4.1.4	Correlations	51
4.2	Discussion	55
Chapter 5	Conclusion	59
5.1	Other ANT-based tests of attention	60
5.2	Limitations and future directions	63
5.3	Final remarks	67
Bibliography	68
Appendix A	Task instructions	74
A.1	Initial task instructions	74
A.2	Subtest instructions for “N-first” group	75
A.2.1	N subtest instructions	75
A.2.2	X subtest instructions	76
A.3	Subtest Instructions for “X-first” group	77
A.3.1	X subtest instructions	77
A.3.2	N subtest instructions	78
Appendix B	Figures of fixation period distributions	79
Appendix C	E1 Plots	81

Appendix D	E2 Plots	99
Appendix E	Correlations from full developmental analysis	117

List of Tables

2.1	Conditional and marginal percentages of double-response types and accuracies of first responses therein.	22
-----	---	----

List of Figures

1.1	Time-course and stimuli of the Attention Network Test.	8
1.2	Example target stimuli common to both CAST subtests.	16
1.3	Time-course of stimuli in the X subtest.	16
1.4	Time-course of stimuli in the N subtest.	17
2.1	Primary effects and interactions in the CAST.	25
2.2	Effects of Session (session 2 minus session 1) on primary effects and interactions in the CAST.	28
2.3	Interaction between xT and nE and in each session.	29
2.4	Credibly non-zero correlations between outcomes (ER & RT) within a CAST effect.	30
2.5	Credibly non-zero correlations involving the intercept.	31
2.6	Credibly non-zero correlations between different attention systems.	31
2.7	Reliability for each of the primary measures in the CAST.	32
2.8	Between- and within-participant effect sizes for primary measures of the CAST.	33
3.1	Primary effects and interactions in the CAST.	40
3.2	Effects of Session (session 2 minus session 1) on primary effects and interactions in the CAST.	42
3.3	Interaction between xT and nE and in each session.	43
3.4	Interaction between nS and nE and in each session.	43
3.5	Interaction between nS and xE and in each session.	44
3.6	Credibly non-zero correlations between outcomes (ER & RT) within a CAST effect.	44
3.7	Credibly non-zero correlations involving the intercept.	45

3.8	Credibly non-zero correlations between different attention systems.	46
3.9	Reliability for each of the primary measures in the CAST. . .	47
3.10	Between- and within-participant effect sizes for primary measures of the CAST.	48
4.1	Distribution of participant ages from E1 and E2.	51
4.2	Effects of Age (change per year) on primary effects and interactions in the CAST.	52
4.3	Interaction between age and nS	53
4.4	Interaction between age and nE	53
4.5	Interaction between age and xE	54
4.6	Interaction between age, nE and xE	54
4.7	Credibly non-zero correlations between outcomes (ER & RT) within a CAST effect.	55
4.8	Credibly non-zero correlations involving the intercept.	56
4.9	Credibly non-zero correlations between different attention systems.	56
B.1	Histogram of all fixation periods from all participants in E1 . .	80
B.2	Histogram of all fixation periods from each of a random selection of participants in E1	80
C.1	Main effect of the nT manipulation on RT and ER	82
C.2	Main effect of the xT manipulation on RT and ER	83
C.3	Main effect of the nS manipulation on RT and ER	84
C.4	Main effect of the xS manipulation on RT and ER	85
C.5	Main effect of the nE manipulation on RT and ER	86
C.6	Main effect of the xE manipulation on RT and ER	87
C.7	Interaction of effects of nT and nS manipulations on RT and ER.	88
C.8	Interaction of effects of xT and xS manipulations on RT and ER.	89

C.9	Interaction of effects of nT and nE manipulations on RT and ER.	90
C.10	Interaction of effects of nT and xE manipulations on RT and ER.	91
C.11	Interaction of effects of xT and nE manipulations on RT and ER.	92
C.12	Interaction of effects of xT and xE manipulations on RT and ER.	93
C.13	Interaction of effects of nS and nE manipulations on RT and ER.	94
C.14	Interaction of effects of nS and xE manipulations on RT and ER.	95
C.15	Interaction of effects of xS and nE manipulations on RT and ER.	96
C.16	Interaction of effects of xS and xE manipulations on RT and ER.	97
C.17	Interaction of effects of xE and nE manipulations on RT and ER.	98
D.1	Main effect of the nT manipulation on RT and ER	100
D.2	Main effect of the xT manipulation on RT and ER	101
D.3	Main effect of the nS manipulation on RT and ER	102
D.4	Main effect of the xS manipulation on RT and ER	103
D.5	Main effect of the nE manipulation on RT and ER	104
D.6	Main effect of the xE manipulation on RT and ER	105
D.7	Interaction of effects of nT and nS manipulations on RT and ER.	106
D.8	Interaction of effects of xT and xS manipulations on RT and ER.	107
D.9	Interaction of effects of nT and nE manipulations on RT and ER.	108
D.10	Interaction of effects of nT and xE manipulations on RT and ER.	109
D.11	Interaction of effects of xT and nE manipulations on RT and ER.	110

D.12	Interaction of effects of xT and xE manipulations on RT and ER.	111
D.13	Interaction of effects of nS and nE manipulations on RT and ER.	112
D.14	Interaction of effects of nS and xE manipulations on RT and ER.	113
D.15	Interaction of effects of xS and nE manipulations on RT and ER.	114
D.16	Interaction of effects of xS and xE manipulations on RT and ER.	115
D.17	Interaction of effects of xE and nE manipulations on RT and ER.	116
E.1	Correlations subfigure 1.	118
E.2	Correlations subfigure 2.	119
E.3	Correlations subfigure 3.	120
E.4	Correlations subfigure 4.	121
E.5	Correlations subfigure 5.	122
E.6	Correlations subfigure 6.	123
E.7	Correlations subfigure 7.	124

Abstract

Modern research on the phenomena of attention has motivated an increasingly nuanced view of its subsystems and their relations. However, popular tools for the measurement thereof are limited in their design in ways that hinder ongoing exploration of both typical and atypical operation of attention. This work describes a new tool that seeks to remediate these deficiencies, the Combined Attention Systems Test (CAST), and details initial efforts to validate its use, including deployment to study the operation of attention in populations of young adults and children. We observe reliable and stable measurement of a variety of phenomena of attention, some expected from prior literature and some newly discovered as a function of the CAST's improved design over existing tools. While we remark on areas of further potential improvement, this work argues strongly for the use of the CAST over previously popular tools.

List of Abbreviations and Symbols Used

- *ANT*: Attention Network Test
- *ANT-I*: Attention Network Test - Interaction
- *CAST*: Combined Attention Systems Test
- *RT*: response time
- *ER*: error rate
- *I*: intercept
- *nT*: endogenous temporal attention
- *xT*: exogenous temporal attention
- *nS*: endogenous spatial attention
- *xS*: exogenous spatial attention
- *nE*: endogenous executive attention (i.e. flanker effect)
- *xE*: exogenous executive attention (i.e. spatial Stroop)

Acknowledgements

This work would not have been possible without the intellectual, moral and financial support of Dr. Raymond Klein.

Dr. Shannon Johnson provided key inspiration and data collection resources in the development of the CAST.

I am also indebted to Dr. Tracy Taylor-Helmick and Dr. John Christie, who provided critical assessment and helpful suggestions for this document.

Chapter 1

Introduction

First observed in writing from the 14th century, the word “attention” has a latinized etymology from the words “tendere” (“to stretch”) and “ad” (“to” or “toward”), and was used specifically in reference to the mind, as if to stretch one’s mind toward something or exert one’s mind to a certain goal. While a similar colloquial usage is employed today, progress in the science of mind has resulted in a more specific and multifaceted understanding of attention as a mental phenomenon. While some preliminary work delineating the features of attention was undertaken by early luminaries such as Wundt (1874), James (1890), Ribot (1898), Titchener (1908), and Pillsbury (1908), the subsequent dominance of the behaviourist perspective left the field of attention stagnant for several decades. Following the cognitive revolution of the 1950s and consequent revival of experimental exploration and theoretical consideration of attention, Posner and Boies (1971) ventured a summary of thinking to date by delineating three “components” of attention: alertness, selection, and capacity.

By “alertness”, Posner and Boies (1971) intended to denote the phenomena associated with the maintenance of sensitivity to stimulation over periods of time in the often boring tasks employed by experimentalists, including the phenomenon whereby experience of a warning stimulus prior to onset of a target stimulus enhanced the observer’s sensitivity to detect and discern the properties of that target stimulus. By “selection” (also termed “orienting” in subsequent literature), they intended to denote phenomena associated with scenarios within which multiple stimuli or information sources are present yet response-pertinent information manifests at a subset of these sources, requiring the participant to discern pertinent sources and ignore irrelevant sources. The necessity of such selection implies the third component, “capacity”, which Posner and Boies identified with phenomena associated with the inability of the mind to simultaneously process multiple inputs and execute multiple outputs.

Later work (Baddeley & Hitch, 1974; Norman & Shallice, 1980) delineated an “executive” system responsible for coordinating behaviour amidst such capacity limitations, and subsequent neuroanatomical evidence (Posner & Peterson, 1990; Posner & DiGirolamo, 2000) bolstered the computational and physical separability of alerting, orienting and executive aspects of attention.

More recently, Klein and Lawrence (2011) proposed a framework of attention with two orthogonal dimensions: domains and modes. In this framework, domains roughly map to the categories delineated by Posner and colleagues but with a slightly different nomenclature, referring to the domains of time, space and task.¹ In this framework, each domain of attention can be engaged via two different modes: exogenous and endogenous. To engage a domain endogenously involves the learning of contingencies local to an experimental task while the exogenous mode involves either innate responses or stimuli with very well-learned extra-experimental associations to responses. This framework thus relies on a large literature addressing the distinctions between exogenous and endogenous spatial attention, extending this distinction to both temporal and executive attention as well.

Evidence for both exogenous and endogenous modes of spatial attention became most concrete with the experiments of Posner (1980) and Jonides (1982) in which “targets” for behavioural response were presented at random spatial locations but preceded by “cues” that could vary in their properties, depending on whether exogenous, endogenous, or both modes of spatial attention were to be engaged. A typical cue to engage exogenous spatial attention would be a stimulus that is distinct from the set of possible targets that appears at a location that is completely random with respect to the location of the subsequent target. With such cues, it is typically observed that when the target follows the cue after a short (100-200ms) interval, responses to the target are faster when the target appears at the same location as the cue, as if attention is drawn to the location of the cue despite its complete lack of predictive

¹Note that Klein & Lawrence opted for the term “task” over the more common “executive”, on the basis that “task” might be considered more descriptive and dissociates from a broad use of the term “executive”; for example, some literatures where the construction and working-through of plans are termed “executive” tasks, without clear relation to the use of “executive” in the attention literature. However, while these motives are sound, for the mundane reason that brevity in the results sections to follow is best achieved by having the terms associated with each domain starting with a unique letter, this report will revert to using the more conventional “executive” attention when referring to what Klein and Lawrence termed “task-based” attention.

information with regards to the location of the subsequent target. A typical cue to engage endogenous spatial attention would be a stimulus that appears at a central location (where targets never appear) but has some identity property (ex. color, shape) that is manipulated to relatively consistently predict the location of the subsequent target (ex. diamond-shaped cues are usually followed by right-located targets, square-shaped cues are usually followed by left-located targets). With such cues, it is typically observed that when the target follows the cue after a longer interval (500-1000ms), responses to the target are faster when the target appears at the location predicted by the cue. In the decades following these initial demonstrations, there has been much work elaborating the distinctions among exogenous and endogenous cues of various specific forms yielding evidence for a variety of influences on their effects including: the cue-target interval, the existence of extra-experimental spatial associations with the cue stimuli, the magnitude of experimentally-manipulated spatial associations, and instruction set to the participant. While a more nuanced picture of exogenous and endogenous spatial attention is now emerging (Christie, Chun, Wylie & Klein, submitted) relating the two modes on a continuum associated with these manipulations, it remains useful to examine the relative extremes of this continuum with the distinct cue stimuli described above.

In contrast with the decades of research on the distinction between endogenous and exogenous spatial attention, it is only more recently that the domain of temporal attention has been delineated as affected by these modes. While a variety of phenomena associated with temporal attention have long been observed and explored since the cognitive revolution (Klemmer, 1956; Karlin, 1959; Bertelson, 1967; Bertelson & Tisseyre, 1968; Niemi & Näätänen, 1981), these phenomena had not been given careful consideration with regard to their relation to attention more generally until the work of Lawrence and Klein (2013). We, who introduced a pair of novel experimental manipulations that combined to provide evidence for exogenous and endogenous forms of temporal attention. Specifically, previous research demonstrated that presentation of a non-target “warning” stimulus prior to the onset of a target improved detection and discrimination performance for the subsequent target. Improved compared to what? While a number of manipulations had been explored to provide various comparisons, including warning presence/absence, warning intensity,

and the time between the warning and target, Lawrence and Klein demonstrated that when one considers the paradigm from the perspective of separable endogenous and exogenous modes of attention, all previous manipulations either left these modes confounded, engaging both at the same time, or failed to sufficiently manipulate each.

For example, studies might precede the target by a warning on some trials but present the target with no warning on other trials, and compare performance across these conditions to measure the effect of the warning. However, as typically implemented, this comparison is confounded insofar as it should engage both endogenous and exogenous temporal attention. That is, such studies typically establish a fixed temporal interval between the warning and target, and thus the warning provides information with respect to the subsequent time of target presentation, a contingency that would engage endogenous temporal attention. Yet Lawrence and Klein (2013) propose that the mere presentation of a stimulus preceding the target is sufficient to engage exogenous temporal attention. One might consider varying the warning-target interval from trial to trial to diminish the temporal information manifest by the warning, thereby diminishing engagement of endogenous temporal attention and leaving exogenous temporal attention the focus of study; however, this approach is limited by the demonstrable sensitivity of human participants to the remaining temporal information even in such interval-varying conditions (Baumeister & Joubert, 1969). Specifically, after experiencing a number of trials with varying warning-target intervals, participants can use the experienced distribution of intervals to modulate their readiness to respond following a warning to optimize performance for the likely target presentation times.

For example, if two discrete intervals are presented randomly and equiprobably across trials, on any given trial the participant will be completely uncertain as to the likely time of target onset during the time between the onset of the warning and the time of the end of the shortest possible interval. However on trials on which the shortest interval elapses and no target has been presented, participants can be perfectly certain that the target will appear at the termination of the longer interval. Thus, responses to targets on short-interval trials reflect performance under more temporal uncertainty than responses to targets on long-interval trials. In such scenarios, we expect participants to perform better on long-interval trials than short-interval trials,

and well-replicated observation of this precise pattern of performance demonstrates the dynamic engagement of endogenous temporal attention in such paradigms. Neither the addition of more discrete intervals nor sampling intervals continuously affect this pattern. Indeed, even changing the distribution of intervals from equiprobable/uniform to other distributions yields performance that matches expectations if participants are learning the distribution of intervals and dynamically re-weighting their readiness within trials to optimize performance given these distributions.

To fully eliminate the engagement of endogenous temporal attention, Lawrence and Klein (2013) proposed the application of techniques developed by Rescorla (1969) who established gold-standard controls for similar problems of contingency in the field of animal learning. In this method, the very idea of a “trial” itself is eliminated such that a fixed number of targets and warning stimuli are distributed randomly and independently within a block of time, thereby eliminating any temporal contingency between the stimuli. After data collection, the warnings and targets can be paired to inspect the effect of the warning and time-course thereof. While the use of this “Rescorla method” provides a means by which endogenous temporal attention can be minimized for the study of exogenous temporal attention, Lawrence and Klein deployed a second novel experimental manipulation to achieve the converse, minimizing the engagement of exogenous temporal attention for the study of endogenous temporal attention. Proposing stimulus intensity as a key stimulus property for engagement of exogenous temporal attention, Lawrence and Klein developed a stimulus that could achieve a minimal-yet-discernible intensity that could be manipulated to have increased intensity. This was achieved by use of the fact that when presented with auditory white noise, humans can quickly discern changes in the between-ear correlation of noise (Boehnke, Hall, & Marquardt, 2002), with the most easily discerned change being that from mono noise (the same noise source for both ears) to stereo noise (an independent noise source for each ear). Note that with this stimulus, there is no change in the intensity of stimulation on the sensory effectors, but an intensity change can be added by accompanying the change in correlation with a change in volume of the noise.

By using a brief change in noise correlation as the warning stimulus, and manipulating warning intensity orthogonally to the use of either the Rescorla method or

fixed warning-target intervals, Lawrence and Klein (2013) were able to demonstrate the existence of both endogenous and exogenous temporal attention, and also show that these modes interact. Consequently², Klein and Lawrence (2011) proposed that these modes exist not only in spatial and temporal domains, but also in the executive domain. Little elaboration was made thereafter on the delineation between phenomena of executive attention associated with either endogenous or exogenous modes, but we will here propose such delineation. One phenomenon traditionally associated with executive attention is Eriksen and Erkisen’s (1974) “flanker” task, which involves the presentation of multiple stimuli, all of which are associated with a task-related response, with the instruction to execute the response associated only with the central stimulus. For example, participants might be instructed to indicate the direction of an arrow that is pointing either left or right that is situated at the center of an array of other “flanker” arrows that are also pointing either left or right. The stimuli in such tasks are typically manipulated such that all of the flanker stimuli have the same identity, which is either identical with the central target (“congruent flankers”) or opposite to it (“incongruent flankers”), and comparison of performance in target discrimination between these conditions reveals dramatically faster and more accurate responses in the congruent flankers condition. As it has been demonstrated (Miller, 1991) that the performance difference is attributable primarily to the fact that the flanker stimuli have identities associated with task-related responses, and that this effect persists when these associations are learned during the course of the experiment, it seems reasonable to infer that this phenomena is associated with an endogenous mode of executive attention.

Another phenomenon traditionally associated with executive attention is the Simon effect (Simon, 1969), which manifests in tasks where a lateralized response requirement is manipulated orthogonally with location of the target stimulus whose identity is mapped to that response requirement. For example, the task might require participants to indicate the identity of targets consisting of squares and diamonds, with instruction to indicate squares with a right-handed button push and diamonds with a left-handed button push. If this task furthermore presents the squares and

²Despite their official publication dates, the work of Lawrence and Klein (2013) preceded that of Klein and Lawrence (2011).

diamonds as randomly located to the left or right of fixation, performance for responses in the same location as the stimulus (ex. right-located squares, left-located diamonds) will be better than for responses opposite the location of the stimulus (ex. left-located squares, right-located diamonds). While there is learning involved in such experiments (mapping the target identity to response location), the source of the response conflict lies in the unlearned (or extra-experimentally well-learned) association between stimulus location and response (Lu & Proctor, 1995), and it therefore seems reasonable to infer that this phenomenon is associated with an exogenous mode of executive attention. A variant of the Simon effect is the “spatial Stroop” effect, so dubbed as it arose in a somewhat independent literature from the Simon effect but reflects the special case of the Simon effect that occurs when the identity property of the target stimulus used to map to response is itself spatial in nature (ex. arrows or the words “left” and “right”).

1.1 Measurement of attention

While the preceding introduction to the phenomena of attention (and theoretical taxonomy thereof) describes multiple experimental paradigms for their observation and measurement, for a variety of reasons it would be useful to devise a single experimental paradigm for the simultaneous measurement of as many of these phenomena as possible. This is advantageous in: (1) reducing the number of times participants have to be instructed as to the requirements of a new task, (2) reducing the overall time spent collecting data from a given participant, and (3) providing the opportunity to examine the interaction of the systems associated with the individual phenomena. It is presumably with these goals in mind that Fan, McCandliss, Sommer, Raz and Posner (2002) developed the Attention Network Test (ANT). The ANT is inspired by the tripartite framework of Posner and Peterson (1990) and as such proposes to measure its three systems (Alerting, Orienting, and Executive) and their interactions. The ANT presents arrows as target stimuli, and participants are tasked with indicating their direction, left or right. The ANT (see Figure 1.1) provides an “Executive” measure from a standard Eriksen and Eriksen (1974) manipulation of flanker arrows presented with the target arrow, while measures for “Alerting” and “Orienting” are derived from various cue stimuli that precede the presentation of the target. On

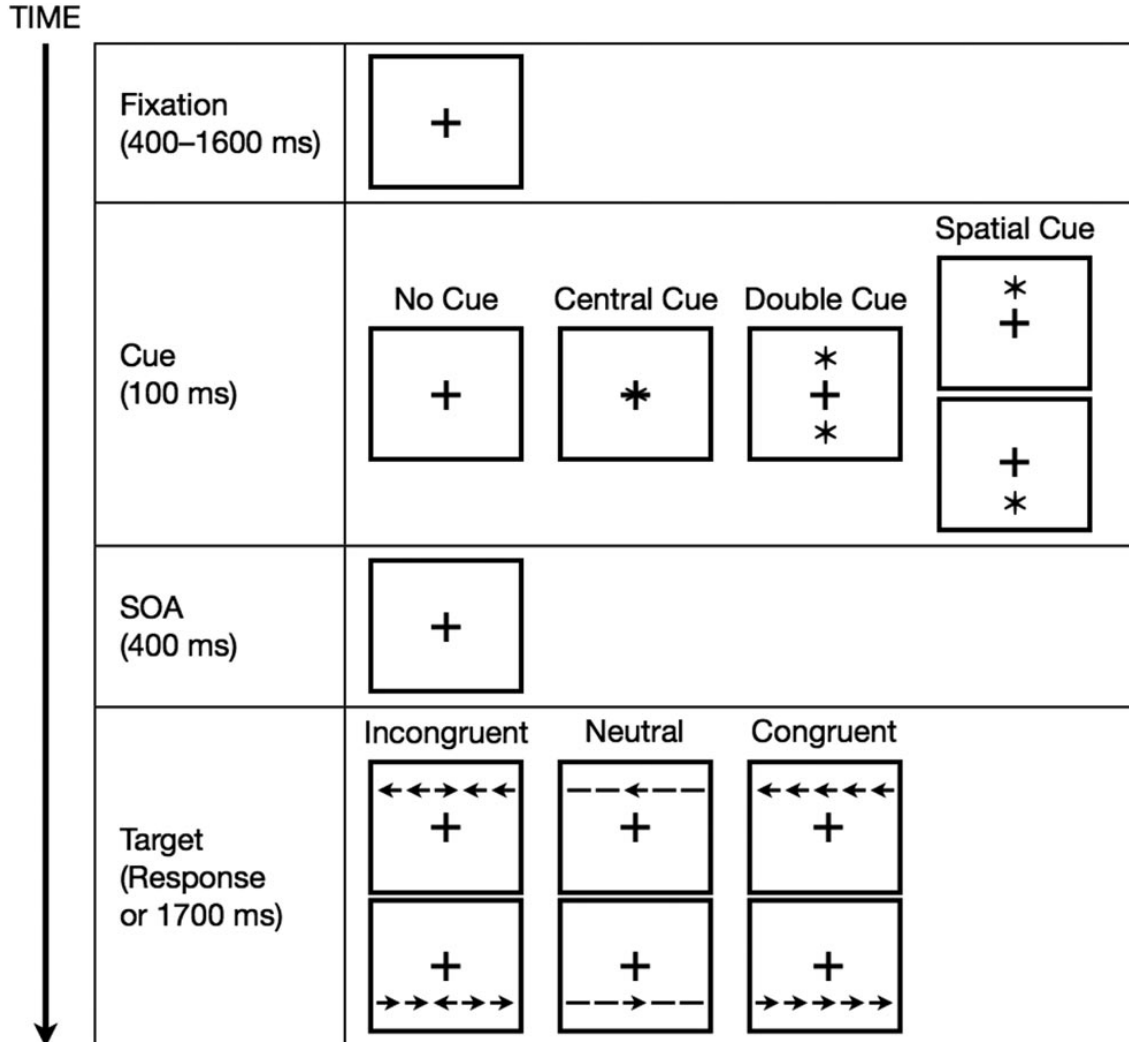


Figure 1.1: Time-course and stimuli of the Attention Network Test (reproduced from MacLeod et al. 2010).

some trials no cue stimuli precede the target, while on others a visual cue appears at fixation precisely 500ms before the target; on all trials there is a random fixation period preceding the moment when a cue might appear, and the ANT proposes to measure Alertness as the difference in performance between no-cue and center-cue performance. The targets appear randomly either above or below fixation, and on some trials a single visual (“spatial”) cue appears at whichever location will be the location of the subsequent target, while on other trials visual cues appear at both possible locations. By comparing performance among the spatial-cue and double-cue conditions, the ANT proposes to measure Orienting.

While it has become a frequently deployed tool since its introduction, the ANT has several deficiencies. At the level of experimental design, the manipulation of both Alertness and Orienting by a single dimension of the design, the nature of the cue stimulus, means that these constructs are being measured on separate trials, adding time to data collection, and cannot be evaluated for how they might interact. Furthermore, the fact that the spatial-cue stimulus is both peripherally-located and provides spatial information as to the likely location of the subsequent target means that the ANT's measure of Orienting confounds the well-established and distinct exogenous and endogenous forms of spatial attention. When such confounds exist and a measure obtains an effect of some treatment (or difference between groups), it is impossible to attribute this observation to one mode or the other. Furthermore, it is possible for a treatment to have an effect (or a set of groups to differ) in opposite directions on the two modes, yielding diminished power to detect any difference at all.

In the years since its first publication, several attempts have been made to modify and improve the ANT, and noteworthy examples of these efforts are discussed in relation to the present work in the General Discussion. The work of Callejas, Lupiáñez, and Tudela (2004), however, stands out as an immediate inspiration for the present efforts. Cognizant of the above noted deficiencies in the ANT, Callejas et al. developed a modified version of the ANT they dubbed the ANT-I (“I” for interaction). In the ANT-I, Alertness and Orienting are manipulated by orthogonal stimuli, permitting more efficient measurement of both as well as evaluation of their interaction. Orienting is measured similarly to the ANT whereby the target is preceded by a visual cue stimulus, but Callejas et al. eliminated the confound of the ANT by presenting only single cues and making their location random with respect to the subsequent target, thereby achieving a pure measure of exogenous spatial attention through comparison of performance between trials on which the target appears at the same versus opposite location as the cue. Alertness in the ANT-I is manipulated by an auditory warning stimulus that precedes the cue on half the trials, with the other half considered “no warning” trials. While the modifications manifested by the ANT-I were a genuine improvement over the ANT, the work reported here represents an attempt to make even further improvements.

Specifically, in eliminating the confound between endogenous and exogenous spatial attention manifest in the ANT, the ANT-I opts to measure solely exogenous spatial attention, yet surely there may be scenarios where measurement of endogenous spatial attention would be useful. Furthermore, the ANT-I's manipulation of Alertness can be improved. That is, to the extent that Alertness as operationalized by the ANT-I involves endogenous temporal attention whereby the presentation of the warning stimulus permits the participant to predict the time of the subsequent target, this role is undermined by subsequent presentation of the cue stimulus, which would also provide this temporal information. Furthermore, in using the contrast between a temporally informative tone and silence to operationalize Alertness, this measure is likely to confound endogenous and exogenous modes of temporal attention. We therefore sought further modifications to the ANT/ANT-I tasks to ameliorate these remaining deficiencies.

1.2 Initial explorations to improve the measurement of attention

Several pilot studies were conducted to explore a variety of options for improving the measurement of attention. In a first study, we sought to modify the ANT-I to:

- Add lateralized targets, permitting measurement of exogenous executive attention
- Achieve measurement of both endogenous and exogenous spatial attention
- Employ child-friendly cartoon fish as target stimuli (rather than arrows)

The latter modification was motivated by a desire to develop a test that could be more easily used to test attention among developmental populations and was inspired by Rueda et al. (2004), who employed similar cartoon fish stimuli.

In our initial explorations we were uncertain how inclusion of measurement of exogenous executive attention and consequent lateralization of target stimuli would affect the other measures of attention, and furthermore what configuration of flankers would be best to employ when targets were lateralized.³ We thus explored a between-blocks manipulation of flanker configurations. One option for the configuration of

³Note that in the General Discussion the work of Fan et al. (2009) will be discussed, which similarly implemented lateralized targets but was published after we began our own explorations.

flankers is to place flankers above and below the target in a column, in which case the central target should be easily localized (i.e. the target is in line with fixation with no intervening stimuli) as it is in the standard ANT/ANT-I configuration where the target appears above and below fixation with flankers to either side. Another option for the configuration of flankers would be to place flankers to the left and right of the target in a row, in which case there is always a flanker between fixation and the target, possibly leading to enhanced manifestation of effects of endogenous executive attention as it may be more difficult with this configuration to spatially filter the flankers. This row-configuration necessitates a larger distance between the target and fixation to prevent the flanker array from overlapping fixation, so to evaluate the effect of this increased distance we added a third configuration for comparison consisting of targets placed at the same distance as targets in the row-configuration but with column-configured flankers.

To satisfy our desire to measure both exogenous and endogenous spatial attention (c.f. ANT-I, which only measures exogenous spatial attention), we explored the feasibility of achieving measurement of each system on separate trials that were randomly intermixed within blocks, with exogenous spatial attention manipulated as in the ANT-I and endogenous spatial attention manipulated by a 75% valid central arrow. Furthermore, while we used a common cue-target interval for all trials, the duration of this interval was manipulated between participants in order to establish whether a relatively short interval (150ms) or a relatively long interval (900ms) would be best for the measurement of either system of spatial attention.

Twenty undergraduate participants completed this initial pilot study. Results indicated that the location and configuration of targets and flankers did not substantially affect measurement of the attention systems with the exception of the endogenous executive system, for which there was a larger effect for the row-configured flankers. Additionally, and expected from prior research on the time-course of both exogenous and endogenous spatial attention, we observed larger exogenous spatial attention scores in the short cue-target interval condition and larger endogenous spatial attention scores in the long cue-target interval condition. Finally, whereas the lateralization of targets was motivated by a desire to measure exogenous executive attention, we failed to observe a credibly non-zero score for this system.

Having observed larger endogenous executive scores for the farther-located targets with row-configured flankers, we ran a second pilot study using only this target/flanker combination and manipulating cue-target interval between-blocks rather than between-subjects, permitting computation of spatial attention scores for individual participants at both intervals. After 25 undergraduate participants completed the study with this design, we obtained a mix of expected and puzzling results. We observed the expected effects of endogenous executive attention and both exogenous and endogenous spatial attention, including the expected interval effects on the spatial attention scores. However, consistent with the first pilot study, we failed to observe non-zero effects of exogenous executive attention. Furthermore, where the first pilot study demonstrated non-zero effects of the temporal attention stimulus in both cue-target interval conditions, in this second pilot study we observed these effects only in the longer cue-target interval condition.

These inconsistent/unexpected results led us to rethink the task's design, including the previously mentioned deficiency of the general ANT-I design whereby the stimulus to manipulate temporal attention is presented before the stimulus to manipulate spatial attention. Thus, the next step was to explore a design whereby the spatial stimulus precedes the temporal stimulus. Specifically, to minimize the temporal information manifest by presentation of the spatial stimulus, we opted to present the temporal stimulus after an interval that varied randomly from trial to trial with an exponential ("non-aging") distribution. This change meant that the interval between the spatial stimulus and target would be necessarily be long and variable, in which case we considered manipulation of exogenous spatial attention unlikely to yield reliable measurement. Thus we sought to manipulate endogenous spatial attention alone, using an arrow presented at the beginning of trials that accurately predicted the subsequent target's location on 75% of trials. As in the ANT-I, a tone was used to manipulate temporal attention, with half of trials containing no tone and measurement of a temporal attention score obtained by comparison of performance on tone-present and tone-absent trials. We additionally explored whether we could achieve measurement of what might be called "cued temporal attention", a phenomenon first demonstrated by Kingstone (1992) whereby temporal stimuli precede target stimuli with varying intervals and some property of the temporal stimulus

(ex. tone pitch) predicts the likely duration of the subsequent interval, in which case performance is superior on trials when this prediction is accurate (i.e. the target appears after the predicted interval) compared to when it is inaccurate. To achieve this in our design, the pitch of the tone was varied on tone-present trials such that a high tone predicted that the target would appear after a 1s interval while a low tone predicted that the target would appear after a 2s interval. We also ran a second group of participants with a central visual temporal stimulus where the shape (diamond vs square) of the visual stimulus provided the interval prediction.

Finally, motivated to further explore our failure to observe exogenous executive attention, we added “neutral” conditions for manipulations of endogenous and endogenous executive attention. The neutral condition for the manipulation of endogenous executive attention was achieved by presentation of the target without flanker stimuli (as is common in variants of the ANT, but a condition not included in our previous pilot studies), while the neutral condition for the manipulation of exogenous executive attention was achieved by presentation of the target above or below central fixation in addition to left or right of fixation. With both horizontal and vertical target locations, we were concerned that when using a row configuration of target and flankers these conditions differed in whether a flanker was present in the space between central fixation and target. To eliminate this potential confound, we added flankers above and below the target in addition to those to each side. With this new design, we obtained data from 22 participants and observed strong effects for all manipulations with the exception of the attempt to measure cued temporal orienting, which appeared to yield no effect in both auditory and visual temporal stimulus groups (despite the stimulus-present/stimulus-absent measure providing a strong effect for both auditory and visual stimulus groups). Critically, with the inclusion of neutral trials we discovered that exogenous executive attention does indeed manifest, but interacts strongly with endogenous executive attention such that exogenous executive scores are largest on trials with no flanker stimuli, smaller on trials with congruent flanker stimuli, and reverse in sign (worse performance on trials on which the stimulus and response locations match) on trials with incongruent flanker stimuli.

Having failed to observe cued temporal orienting with either visual or auditory cue stimuli, we decided that the complex combination of manipulations was likely

overwhelming the endogenous systems responsible for discerning contingencies and dynamically deploying attention in response to these contingencies. We therefore opted to eliminate the manipulation of temporal stimulus identity and employ a single 1s stimulus-target interval, using a tone as a temporal stimulus and comparing performance on tone-present trials to tone-absent trials as a measure of temporal attention. With this reduced design, we collected data from 34 participants and observed strong effects for all manipulations.

We then engaged one final round of design optimization, including replacement of tone as a temporal stimulus with the white noise stimulus of Lawrence and Klein (2013), which permits manipulation of endogenous temporal attention with minimal engagement of exogenous temporal attention, bolstering attribution of the performance difference between stimulus-present and stimulus-absent trials selectively to endogenous temporal attention. Additionally, after observing from prior experiments little benefit from inclusion of neutral trials for the manipulation of exogenous executive attention (i.e. trials with targets presented above or below fixation), we eliminated these to reduce the total task time (n.b. neutral trials in the manipulation of endogenous executive attention, on which no flankers appeared, were kept in the design as these are necessary for the measurement of exogenous executive attention). We similarly sought a reduction of task time by decreasing the ratio of valid-to-invalid trials in the manipulation of endogenous spatial attention from 3:1 to 2:1. Finally, and more substantially, motivated by a desire to provide measurement of exogenous modes of both spatial and temporal attention, we added a second subtest wherein these could be measured. A full description of this final design follows.

1.3 The Combined Attention Systems Test

To reflect the numerous improvements over existing tasks, our new test abandons the “ANT” nomenclature and is dubbed the Combined Attention Systems Test (CAST). For brevity, the following abbreviations will be used hereafter:

- nT : **e**ndogenous **t**emporal attention
- xT : **e**xogenous **t**emporal attention
- nS : **e**ndogenous **s**patial attention

- xS : exogenous spatial attention
- nE : endogenous executive attention (i.e. flanker effect)
- xE : exogenous executive attention (i.e. spatial Stroop)

Furthermore, the subscript RT and ER will be appended to these abbreviations to indicate the measurement of the above as manifest in response time and error rates, respectively.

The CAST maintains the original core task of the ANT/ANT-I whereby participants are instructed to indicate the direction of a stimulus, with numeric feedback presented following responses to indicate the time it took to respond and motivate rapid responses. Emulating Rueda et al. (2004), the CAST employs child-friendly colorful cartoon fish as target stimuli (rather than arrows), which can appear alone or embedded in what can be described to participants as a “school” of flanking fish. Manipulating whether the flanking fish point in the same or opposite direction as the target fish permits measurement of nE (see Figure 1.2). In contrast to the ANT/ANT-I, which locates targets either above or below fixation, the CAST presents targets to the left or right of fixation. By having participants indicate the direction of the target fish using left and right fingers, and with the additional lateralization of targets, we expect to observe spatial Stroop, which the introduction above has argued provides a measure of xE . Throughout the CAST participants wear headphones through which they hear white noise. All trials begin with a shifted random exponential fixation interval (lasting at least 1s, with a mean of 2s and re-sampled for a maximum of 10s; see Appendix B for figures of this distribution); an interval is designed to minimize the temporal information manifest by the trial onset.

Further description of the CAST requires separate discussion of its “N” and “X” subtests (for “eNdogenous” and “eXogenous”, respectively), which take approximately 15 minutes and 7 minutes to complete, respectively, and differ in the nature of the stimuli presented on each trial prior to the appearance of the target fish.

In the X subtest (see Figure 1.3), fixation consists of a black cross on a white background and at the end of the fixation interval a black dot is presented for 100ms at a location that is completely random with respect to the location of the subsequent target. The target appears 200ms following the onset of the dot, and comparison of

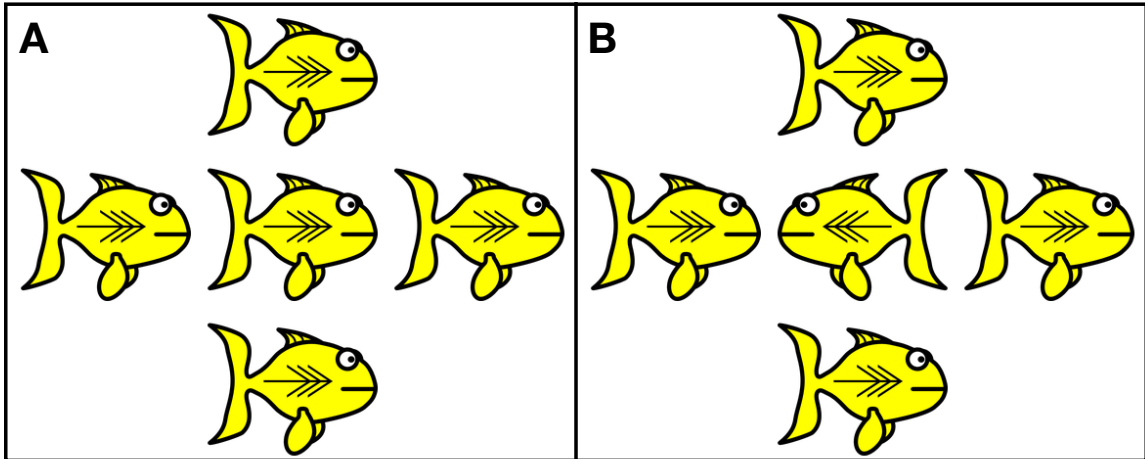


Figure 1.2: Example target stimuli common to both CAST subtests. Panel A shows a congruent Flanker configuration while Panel B shows an incongruent Flanker configuration.

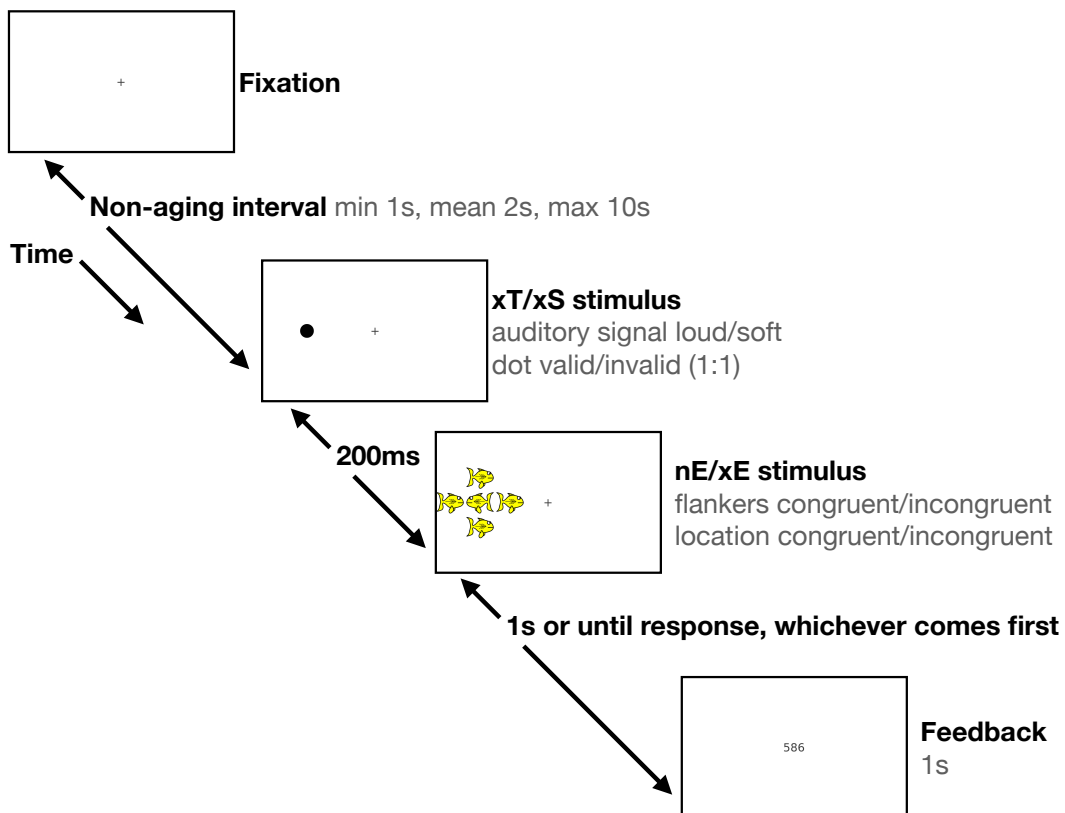


Figure 1.3: Time-course of stimuli in the X subtest.

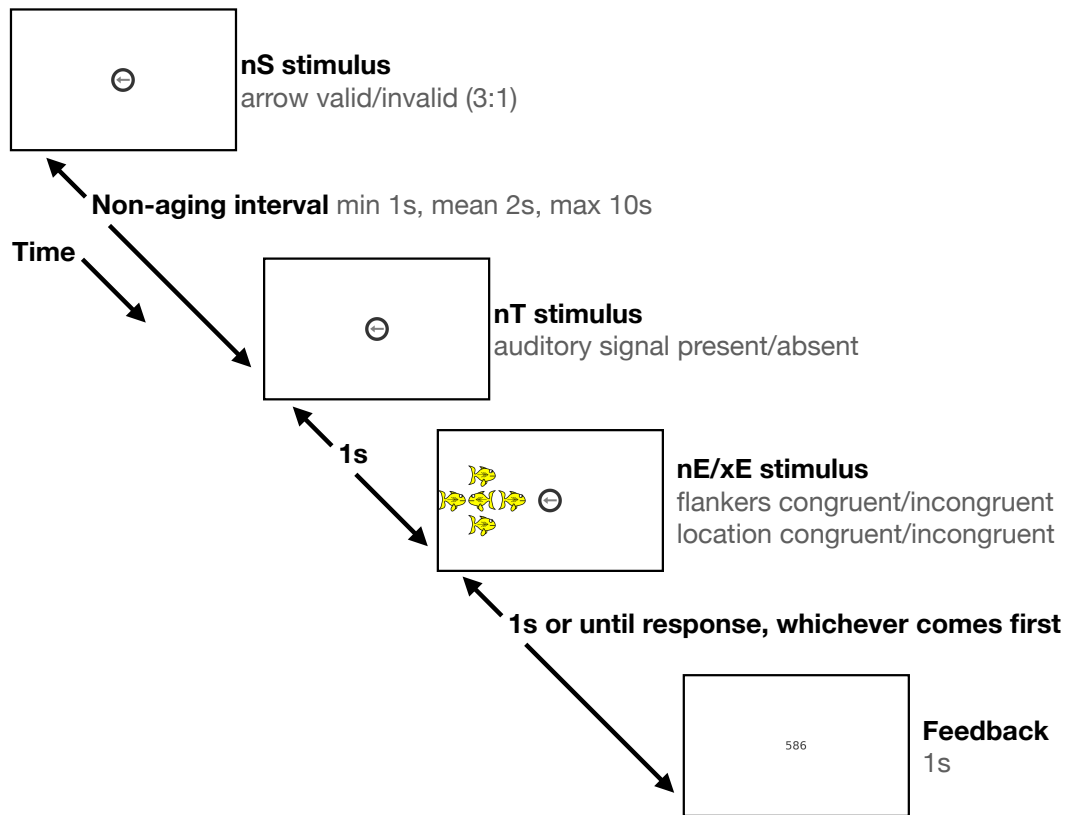


Figure 1.4: Time-course of stimuli in the N subtest.

performance on trials in which the target appears at the same (“valid cue”) versus opposite (“invalid cue”) location as the dot provides a measure of xS . Synchronous with presentation of the dot stimulus, the auditory noise switches from mono to stereo for 100ms, and on half of trials this change is accompanied by a substantial increase in volume of the noise, thus comparison of trials including a noise volume change versus no volume change provides a measure of xT .

In the N subtest (see Figure 1.4), fixation consists of an arrow pointing either left or right, with the direction of this arrow predicting the likely location of the subsequent target (with 66% accuracy); comparison of trials on which the target appears at the arrow-indicated location (“valid cue”) versus the opposite location (“invalid cue”) provides a measure of nS . On half of trials in the N subtest the end of the fixation interval is accompanied by a warning stimulus consisting of a 100ms change in the auditory noise from mono to stereo (with no change in volume). The

target appears 1s following the end of the fixation interval, and the comparison of performance on trials in which the warning was present vs absent provides a measure of nT . (Note that, to account for the possible influence of uncertainty with regards to what kind of event participants should expect next, the fixation arrow is surrounded by a circle on trials that will include a warning stimulus, and a square on trials that will not include a warning stimulus.)

The CAST therefore proposes to measure all cells of the Klein and Lawrence (2013) framework: both exogenous and endogenous modes of temporal, spatial, and executive attention. The present work reports on initial attempts to deploy the CAST in two samples, young adults and middle-to-older children, evaluating the phenomena and psychometric properties of the CAST in each.

Chapter 2

Experiment 1: Young Adults

Experiment 1 sought to establish the baseline psychometric properties of the CAST in a healthy young adult population, a popular demographic for the study of attention. It may be expected that attention in this population is operating at peak efficiency, having achieved maturity whilst not yet diminished by the effects of aging into older adulthood. Participants in this experiment completed the CAST twice in a single appointment lasting approximately 45-minutes, with the second repetition conducted immediately after the first, so the stability of its measures across repeated testing could be evaluated.

2.1 Methods

2.1.1 Participants

Participants without self-reported diagnosed cognitive deficits were recruited from a local undergraduate Research Experience Participation Pool, and awarded bonus course credit as compensation for their time. While the target N was 80 participants, 82 data sets were collected, 1 was removed after reporting a diagnosed attention deficit, and 11 were removed for failing to complete both repetitions of the CAST. Further performance-based exclusions included: 4 participants removed for performing at or below chance in more than half of blocks, 1 participant removed for failing to respond on 20% of trials (c.f. next highest miss rate: 2%), and 1 participant removed as a bivariate outlier in mean response time and error rate (ie. responded both slowly and with a high error rate). Of the remaining 64 participants, ages ranged from 18-24 years, 38 were female and 5 were left-handed.

2.1.2 Materials

The experiment was run on an Apple Mac Mini with visual stimuli presented on an LCD with a resolution of 1280x1024 pixels running at 60Hz. The experiment was coded in the Python programming language. Participants were seated approximately 100cm from the screen and wore earphones through which auditory stimuli were presented at a comfortable volume. Participants made responses using their right and left index fingers on the analog triggers of an xbox360 gamepad configured to record all movement of the triggers but that required a movement of at least half its maximum to register a response for feedback. One motivation for use of analog triggers, rather than traditional binary input modes like keyboard keys or gamepad buttons, is that they permit observation of nearly the full timecourse of response, including the phase of response prior to the trigger achieving the criterion for registering a response. This is helpful as we have observed in prior unpublished work that participants make partial responses where a response is initiated but then terminated before reaching the registration criterion on about 10% of trials, often followed by a second response that is much delayed compared to trials with no initial partial responses present. Such delayed second responses contribute to the notorious heavy-tail skew of RT data. If a binary input mode like a keyboard key or gamepad button were used in these cases, no evidence of such partial responses would be recorded.

2.1.3 Procedure

Participants were alternately assigned to groups specifying the order in which they experienced the two CAST subtests (N & X), completing both tests and then repeating them in that same order in a single appointment lasting approximately one hour. Participant-terminated breaks were provided between subtests and after every 24 trials (approximately 3 minutes) during the subtests themselves.

Participants were provided with an initial set of instructions common to both subtests followed by subtest-specific instructions as they first experienced each subtest. The subtest-specific reminder instructions were provided when participants repeated each subtest in the second half of the experiment. See Appendix A for instruction scripts.

The stimuli and procedure of the CAST were as described in the previous chapter.

Feedback after responses consisted of the time it took for the participant to respond, conveyed in milliseconds. If participants failed to respond to the target fish within 1s, the word “Miss!” was presented as feedback. If participants made responses prior to the appearance of the target fish, the trial was immediately terminated and the words “Too soon!” was presented as feedback. All feedback lasted for 1s before the experiment continued to the next trial.

Each repetition of the CAST consisted of the two CAST subtests (X & N), where each subtest consisted of 48 unique trial types (2 target locations \times 2 target directions \times 3 flanker conditions \times 2 *xT-or-nT* stimulus conditions \times 2 *xS-or-nS* stimulus conditions). However, for the N subtest, the “valid” *nS* stimulus condition must be represented twice to achieve a 2:1 ratio of valid:invalid trials, yielding a total of 72 trials. For each subtest, the full set of trials for that subtest (48 for the X subtest, 72 for the N subtest) was repeated twice in successive but independent random orders during that subtest, yielding a total of 96 trials for the X subtest and 144 trials for the N subtest.

2.2 Results

2.2.1 Data pre-processing

Data collected during the experiment consisted of trial summaries containing trial type information as well as the timing and identity of any responses that surpassed the criterion for registering a response for feedback. In addition to these summaries, a trial-by-trial record of the full movement of the analog triggers was collected. Inspection of these records revealed that on a considerable proportion of trials (5%), participants made multiple responses consisting of an initial response that was initiated after presentation of the target but failed to reach criterion, followed by a subsequent response that exceeded criterion. Among these double-responses, about half included a switch from one response to another while the other half of double-responses maintained the initial response identity. Among the switch double-responses, the majority (92%) consisted of a switch from an error to an accurate response, while the majority of non-switch double-responses (78%) were accurate responses. This pattern (see Table 2.1) is consistent with an account whereby continuous but noisy information

		Double Response kind:		
		Switch	Non-switch	
First response is...	Accurate	4%	39%	43%
	Error	46%	11%	57%
		50%	50%	

Table 2.1: Conditional and marginal percentages of double-response types and accuracies of first responses therein.

processing of the target throughout the trial can lead to an initial erroneous identification of the target and subsequent initiation of an erroneous response that is followed by continued processing, revision of the identification, cancellation of the erroneous response, and initiation of the correct response.

While explicit modelling of the double-response phenomenon is possible (especially in a process model of speeded choice tasks like the diffusion model of Ratcliff, 1979), for simplicity we opt to instead use these trials to re-code both response time and accuracy of the double-response trials. That is, we take the perspective that, for the purposes of examining the influence of our experimental manipulations on initial information processing absent the intervention of higher-order response-monitoring processes, we use the time and identity of the first response initiated, regardless of whether it surpasses the criterion for feedback, as the response for that trial.

After this recoding, the collected data were pre-processed to remove all trials on which:

- Participants failed to make a response (0.1% of trials)
- Responses were made prior to target appearance (0.6% of trials)
- Response times were faster than 200ms (0.3% of trials)

The latter exclusion criterion was determined by both prior experience with RT data suggesting that responses faster than 200ms tend to be anticipatory responses unrelated to target processing, as well as application of a generalized additive model of trial accuracies predicted by trial RTs, showing that only above about 200ms do responses rise above chance performance.

2.2.2 Statistical modelling

Bayesian inference was achieved using the Stan (Carpenter et al., 2017; Stan Development Team, 2017) probabilistic programming language via the RStan package for R (R Core Team, 2017). Response time and accuracy from both subtests were modelled simultaneously, where trial-by-trial accuracy was modelled as a binomial event (Dixon, 2008; Jaeger, 2008) and trial-by-trial response time was modelled as having log-normal measurement noise (Luce, 1986; Ratcliff, 1993; Baayen, 2010). Within a given participant, the influence of the manipulated variables on accuracy was modelled as affecting the log-odds of error while their influence on the response time was modelled as affecting the log-mean response time; the scale of the log-normal measurement noise was also modelled for each participant. The full set of coefficients relating a given participant to their trial-level data was modelled as varying across participants through a multivariate normal distribution in a hierarchical model that sought inference on the population-level coefficient means, variabilities, and correlations (i.e. a “maximal” mixed effects model; Barr et al. 2013). Notably, by modelling the response time and accuracy data in the same model, we achieve more accurate/informed inference on their associated coefficients at both the participant and population level to the degree that there are correlations among them manifest in the population, which is a strong expectation for these measures (ex. slower participants tend to be more accurate; participants with larger effects on response time tend to have larger effects on response accuracy, etc.). Weakly informed priors were used for all population-level parameters, and an explicitly exploratory approach was taken in formulating these priors as independent from one another.¹

Note that the output of such Bayesian analysis is a “posterior distribution” conveying the relative credibility of values for each parameter in the model. While we embrace the perspective that these distributions reflect the ultimate description of the results from this work, in order to discuss these results, particularly in relation to previous work, we will adopt a quantitative and linguistic shorthand whereby we will compute 95% Credible Intervals (CrIs) for each parameter reflecting the interval

¹Future confirmatory work might reasonably employ hierarchical priors on related parameters to deploy partial-pooling and thereby reduce rate of type-M errors that may otherwise be of concern with the large number of coefficients in this model.

whose endpoints reflect the narrowest interval² of the posterior distribution containing 95% of samples, and apply the arbitrary decision to consider 0 relatively non-credible as a value if it falls outside this interval. Additionally, we will compute 50% CrIs and apply the arbitrary decision to consider 0 relatively credible as a value if it falls within this interval. In those cases where 0 falls inside the 95% CrI but outside the 50% CrI, we will express indecision as to the binary credibility of 0. Finally, as a linguistic shorthand it is useful to invert this categorization scheme and refer to the credibility of a given effect, where a credibly-non-zero effect is one where 0 falls outside the 95% CrI and a credibly-zero effect is one where 0 falls within the 50% CrI, and again expressing indecision for those effects where 0 falls within the 95% CrI but outside the 50% CrI. Of course, even though this trinary-state representation of the results is a step above the binary representation (“significant/non-significant”) found in typical reports, it is important to reiterate that the language used is merely a shorthand, connected to the full posterior by arbitrary quantitative summaries, and the reader is encouraged to accompany their reading of the prose with careful consideration of the graphical depictions of the distributions themselves.

2.2.3 Primary effects and interactions

Figure 2.1 displays the posterior distributions associated with the primary measures of the CAST as well as their interactions. There is both confirmation of expected results and novelty in these distributions. Confirmed are the expectations on the primary measures of the CAST such that they all achieve posterior distributions that are both narrow (i.e. confident) and far from zero for one of the dependent variables, if not both. Also consistent with prior expectation (ex. Fan, Gu, Guise, Liu, Fossella, Wang & Posner, 2009) is the manifestation of a credible $nE:xE$ interaction such that (see Figure C.17) the xE effect manifests as expected (congruent performance better than incongruent performance) in the nE -congruent condition but manifests the opposite pattern (congruent performance worse incongruent performance) in the nE -incongruent condition (equivalently: a larger nE effect in the xE -congruent condition than in the xE -incongruent condition).

²This “narrowest” interval is sometimes known as the “highest posterior density interval” (Hyndman, 1996).

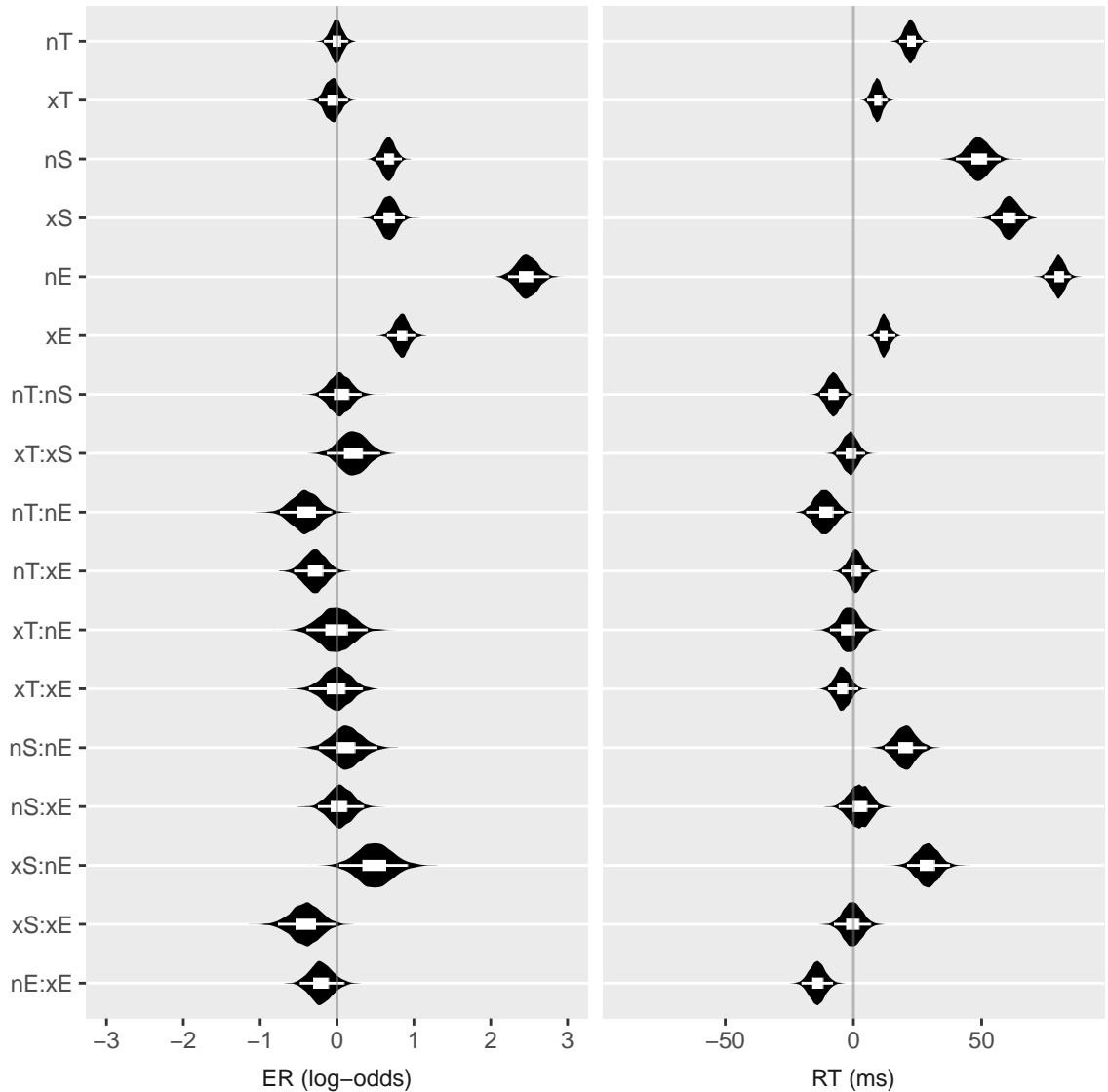


Figure 2.1: Primary effects and interactions in the CAST. Violin plots represent full posterior distribution for each effect while boxplot conveys the 95% CrI (thinner bounds) and 50% CrI (thicker bounds). These visual conventions repeated in all subsequent figures. Plots of each effect’s constituent conditions can be found in Appendix C. Readers viewing this document electronically should be able to click any row of the plot to jump to that effect’s corresponding figure in Appendix C. Effects are coded such that positive values among the simple effects reflect better performance in the condition in which we expect better performance (ex. a positive nT effect reflects better performance in the nT congruent condition than in the nT incongruent condition). Positive values among the interactions reflect a larger effect of one variable in the condition of the other variable in which we expect worse performance (ex. the positive $nS:nE$ interaction which manifests as a larger nS effect in the nE incongruent condition and a larger nE effect in the nS invalid condition.)

As prior research has failed to provide rigorous and thorough evaluation of the subsequent interactions, this will be the first time they can be properly evaluated. For example, while previous research has suggested that temporal and spatial attention interact (Callejas, Lupiáñez & Tudela, 2004; c.f. Fernandez-Duque and Posner, 1997), the present data suggest that this is credibly the case only for their endogenous flavours ($nT:nS$, not $xT:xS$).³ As seen in Figure C.7, the $nT:nS$ interaction manifests such that the nT effect is larger in the nS -valid condition than in the nS -invalid condition (equivalently: the nS effect is larger in the nT -present condition than in the nT -absent condition).

While previous research has suggested that temporal attention and nE should interact (Posner, 1994; Fan, McCandliss, Sommer, Raz & Posner, 2002; Callejas, Lupiáñez & Tudela, 2004; Callejas, Lupiáñez, Funez & Tudela, 2005), the present data suggests this is only credibly the case for $nT:nE$, not $xT:nE$.⁴ As seen in Figure C.9, the $nT:nE$ interaction manifests such that the nT effect is larger in the nE -congruent condition than in the nE -incongruent condition (equivalently: the nE effect is larger in the nT -present condition than in the nT -absent condition), a pattern that replicates that observed previously for this interaction.

Finally, where previous work (Funez & Lupiáñez, 2003) observed an interaction between temporal attention and xE , here we observe that this is only credibly the case for $nT:xE$.⁵ As seen in Figure C.10, the $nT:xE$ interaction manifests such that the nT effect on ER is larger in the xE -congruent condition than in the xE -incongruent condition (equivalently: the xE effect is larger in the xT -present condition than in the nT -absent condition), a pattern that replicates that observed previously for this interaction.

Prior literature is a bit more clear on predictions for the interactions involving spatial attention, having a longer history with the distinction between exogenous and endogenous modes thereof. As previously observed with the ANT-I (Callejas,

³Note that the posterior on the difference between the $nT:nS$ effect and $xT:xS$ effects, both as manifest in RT, was ambiguous, with a 95%CrI of -13.9 – 1.15 and 50%CrI of -9.20 – -3.79.

⁴Note that the posterior on the difference between the $nT:nE$ and $xT:nE$ effects were ambiguous for both RT and ER: RT 95%CrI= -20.0 – 0.121; RT 50%CrI= -13.2 – -6.43; ER 95%CrI= -0.895 – 0.140; ER 50%CrI= -0.579 – -0.222

⁵Note that the posterior on the difference between the $nT:xE$ and $xT:xE$ effects were ambiguous for both RT and ER: RT 95%CrI= -2.46 – 12.6; RT 50%CrI= 2.52 – 7.68; ER 95%CrI= -0.725 – 0.132; ER 50%CrI= -0.408 – -0.115

Lupiáñez & Tudela, 2004), we here observe a credible interaction for $xS:nE$. As seen in Figure C.15, the $xS:nE$ interaction manifests such that the xS effect on RT is larger in the nE -incongruent condition than in the nE -congruent condition (equivalently: the nE effect is larger in the xS -invalid condition than in the xS -valid condition), a pattern that replicates that observed previously for this interaction. While previous work (Goldberg, Maurer & Lewis, 2001) failed to observe a $nS:nE$ interaction when using target locations above and below fixation (we were unable to find prior work exploring this interaction with lateral target locations), here we observe a credible $nS:nE$ interaction similar to that between $xS:nE$. As seen in Figure C.13, the $nS:nE$ interaction manifests such that the nS effect on RT is larger in the nE -incongruent condition than in the nE -congruent condition (equivalently: the nE effect is larger in the nS -invalid condition than in the nS -valid condition).

Finally, we here observe a credible $xS:xE$ interaction as has been previously reported (Funez & Lupiáñez, 2003; Funez, Lupiáñez & Milliken, 2007), but in contrast to those reports fail to observe a credible $nS:xE$ interaction.⁶ As seen in Figure C.16, the $xS:xE$ interaction manifests such that the xS effect on ER is larger in the xE -congruent condition than in the xE -incongruent condition (equivalently: the xE effect is larger in the xS -valid condition than in the xS -invalid condition), a pattern that replicates that observed previously for this interaction.

2.2.4 Effects of Session

Examination of how the above measures vary from one session to the next provides insight into the stability of these effects through repeated testing. These are depicted graphically in Figure 2.2. Evident from these is a credible reduction of overall RT, an increase in the magnitude of nT , a reduction of nS , an increase of nE , and an increase of the interaction between nE and xT . This latter effect reflects reversal of direction of an indecisively credible interaction from one session to the next, as depicted in Figure 2.3. It is notable that the direction of the $xT:nE$ interaction in session 1 is negative, matching that observed for the $nT:nE$ and predicted by Posner

⁶Note that the posterior on the difference between the $nS:xE$ effect and $xS:xE$ effects, both as manifest in ER, was ambiguous, with a 95%CrI of -0.0248 – 0.884 and 50%CrI of 0.280 – 0.596.

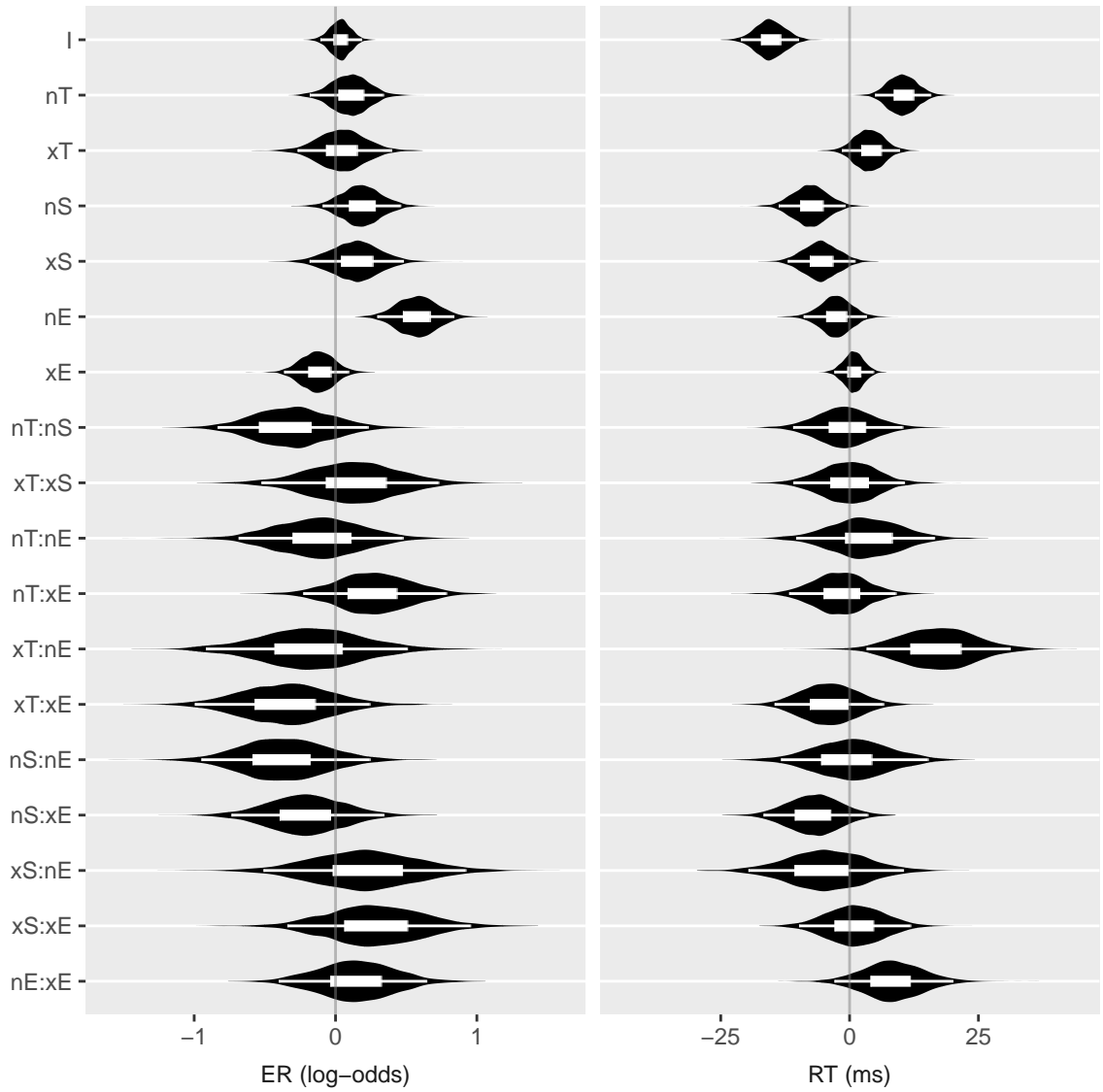


Figure 2.2: Effects of Session (session 2 minus session 1) on primary effects and interactions in the CAST. In addition to the abbreviations noted in the introduction, “I” is used here to indicate the intercept (i.e. overall ER and RT).

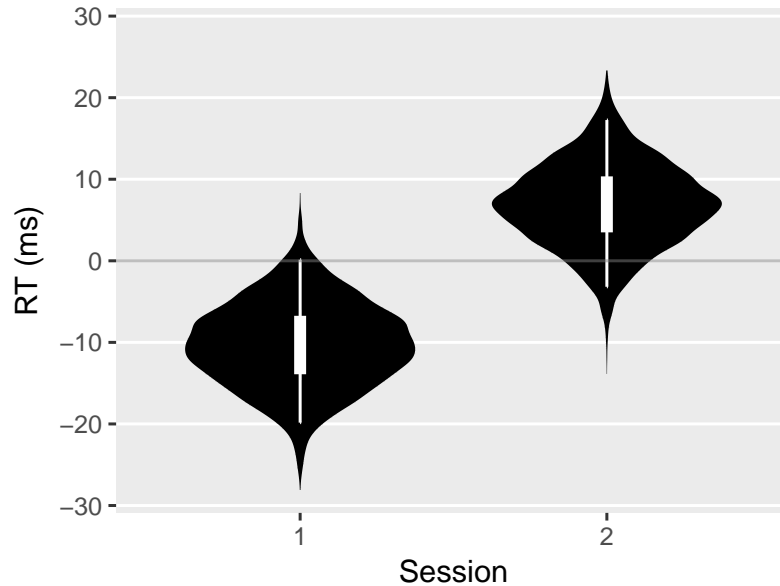


Figure 2.3: Interaction between xT and nE and in each session.

(1994) such that the xT effect is larger in the nE -congruent condition than in the nE -incongruent condition (equivalently: the nE effect is larger in the high-intensity- xT condition than in the low-intensity- xT condition).

2.2.5 Correlations

With two outcomes (RT & ER) and seven primary measures (including the Intercept), there are 91 correlations to evaluate⁷; for brevity Figures 2.4 – 2.6 display those with credibly non-zero values. Figure 2.4 displays credible correlations involving the two outcome measures of a given CAST effect, Figure 2.5 displays correlations involving the intercept, and Figure 2.6 displays correlations between different CAST measures.

Present in Figure 2.4 is the expected negative correlation between overall RT and overall ER, reflecting the fact that slower participants tend to be more accurate, as well as the positive correlations between the manifestation of effects in RT and ER for nS , nE and xE , reflecting that for these measures, participants with larger effects in RT also had larger effects in ER.

Figure 2.5 shows the credibly non-zero correlations involving the intercepts (overall RT & overall ER), which convey the degree to which the measures of the CAST

⁷There are actually nearly 20000 correlations in the model, including among the various higher-order interactions, but only correlations between the primary measures are considered here.

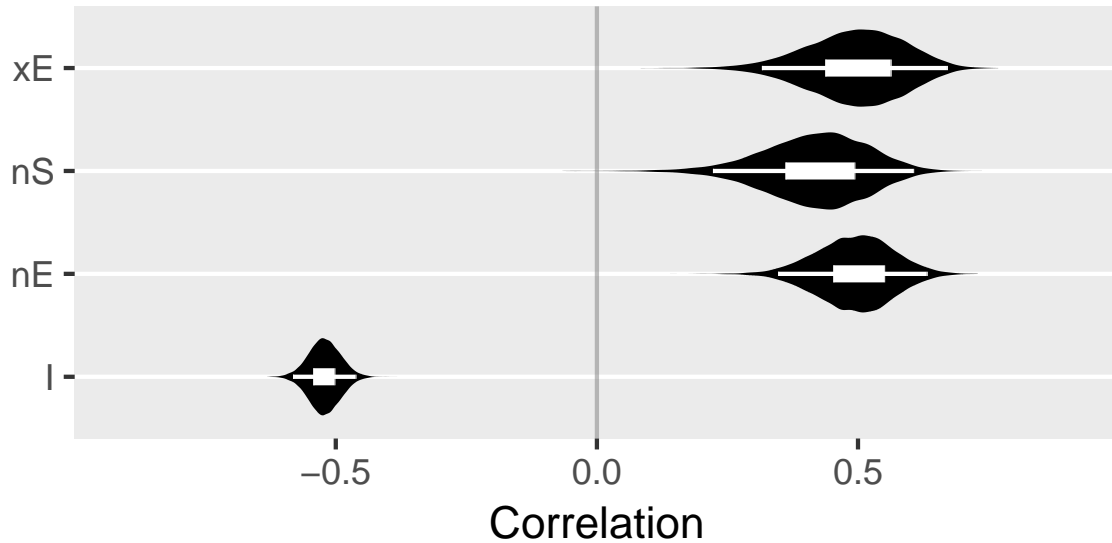


Figure 2.4: Credibly non-zero correlations between outcomes (ER & RT) within a CAST effect.

correlate with overall performance on a speeded-choice task outside the effects of experimental manipulations. We observe here that all attention subsystems measured by the CAST involve at least one correlation with overall performance, with the exception of xT .

Figure 2.6 shows the credibly non-zero correlations among the CAST measures, including a negative correlation between xS_{RT} and nE_{ER} , and a positive correlation between nT_{RT} and nE_{ER} . Insofar as the absence of such correlations may be taken to indicate isolable cognitive systems, the observation of merely 3 credibly non-zero correlations (two of which involve the same CAST measures) bolsters the utility of the CAST as a tool to measure multiple such isolable systems. This conclusion is supported by the observation that of the 60 correlations among the measures, 41 have posterior distributions in which zero falls within the 50% credible interval, and that the median interval of these ranged from -0.08 to 0.09, a relatively narrow range; taken together, these observations indicate that the failure to observe credibly non-zero correlations is not a product of insufficient statistical power but truly zero-or-nearly-zero correlations among these CAST measures.

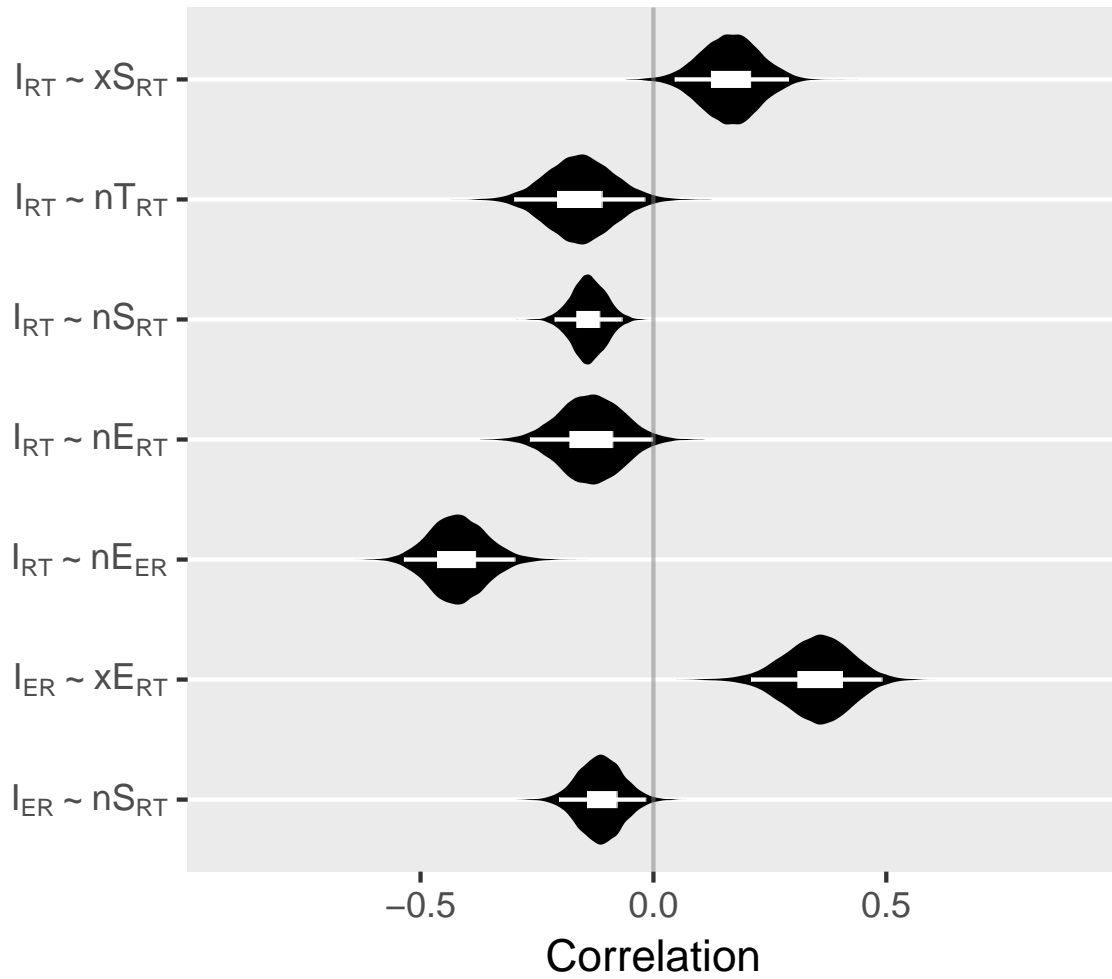


Figure 2.5: Credibly non-zero correlations involving the intercept.

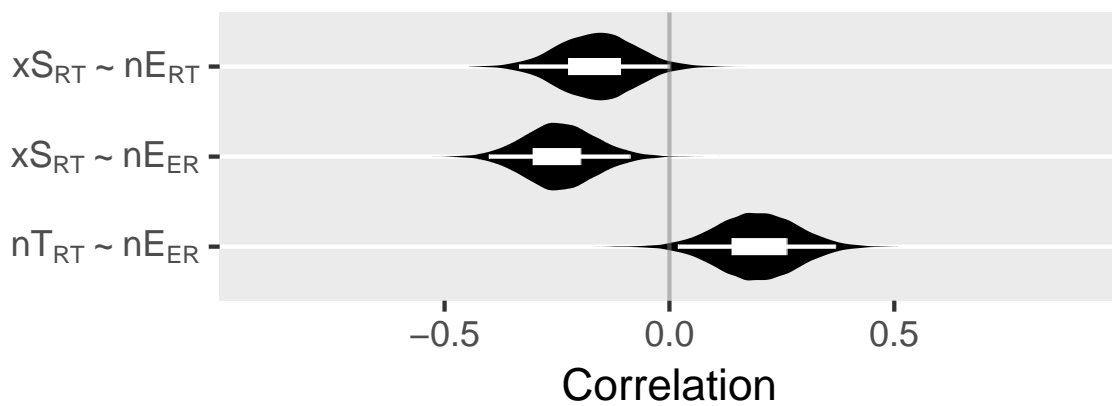


Figure 2.6: Credibly non-zero correlations between different attention systems.

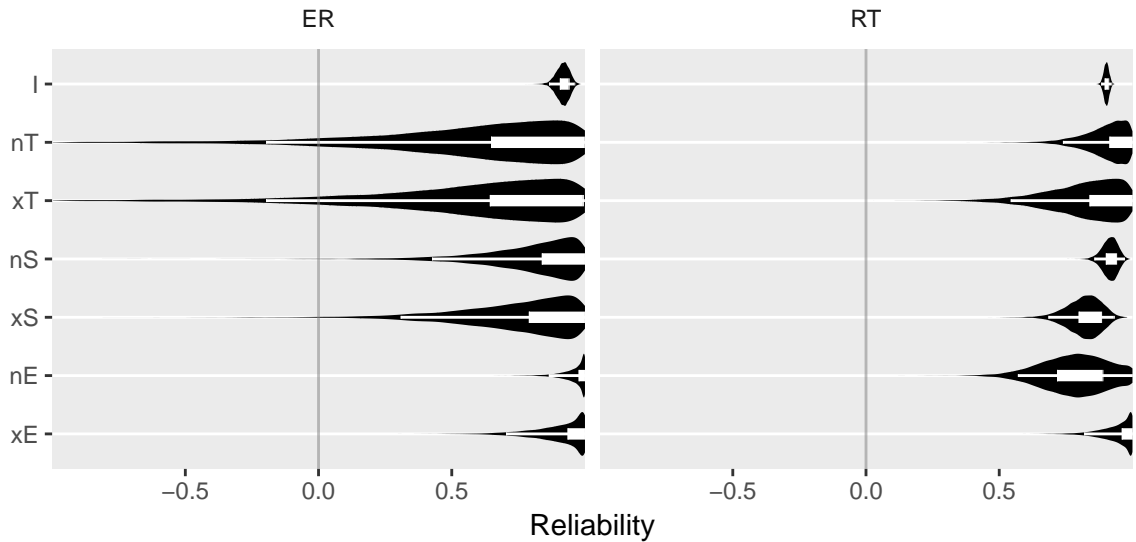


Figure 2.7: Reliability for each of the primary measures in the CAST.

2.2.6 Reliabilities

A benefit of the Bayesian framework is that even though the model was not explicitly parameterized for direct inference on the correlation between participants' scores for a given measure from one session to the next, a conventional measure of test-retest reliability, we are able to compute this quantity from the posterior samples.⁸ The subsequent posterior distributions for these correlations are shown in Figure 2.7, revealing relatively reliable measurement for all measures in RT and for all measures in ER, with the exception of those involving temporal attention, which is expected as they did not manifest a credible effect.

2.2.7 Effect sizes

For each effect of interest, we computed two measures of effect size; one that expresses the magnitude of the effect relative to the amount of between-participants variability associated with that effect, and one that expresses the magnitude of the

⁸The model includes terms for each participant's scores in each session, permitting posterior computation of the test-retest correlation across sessions for each score. While a frequentist evaluation of the same model would yield estimates for these same scores, it would not thereafter be straightforward to derive a confidence interval or p-value on the subsequent correlations. When evaluated in a Bayesian context, the posterior is obtained automatically by computing the correlations in every sample from the posterior.

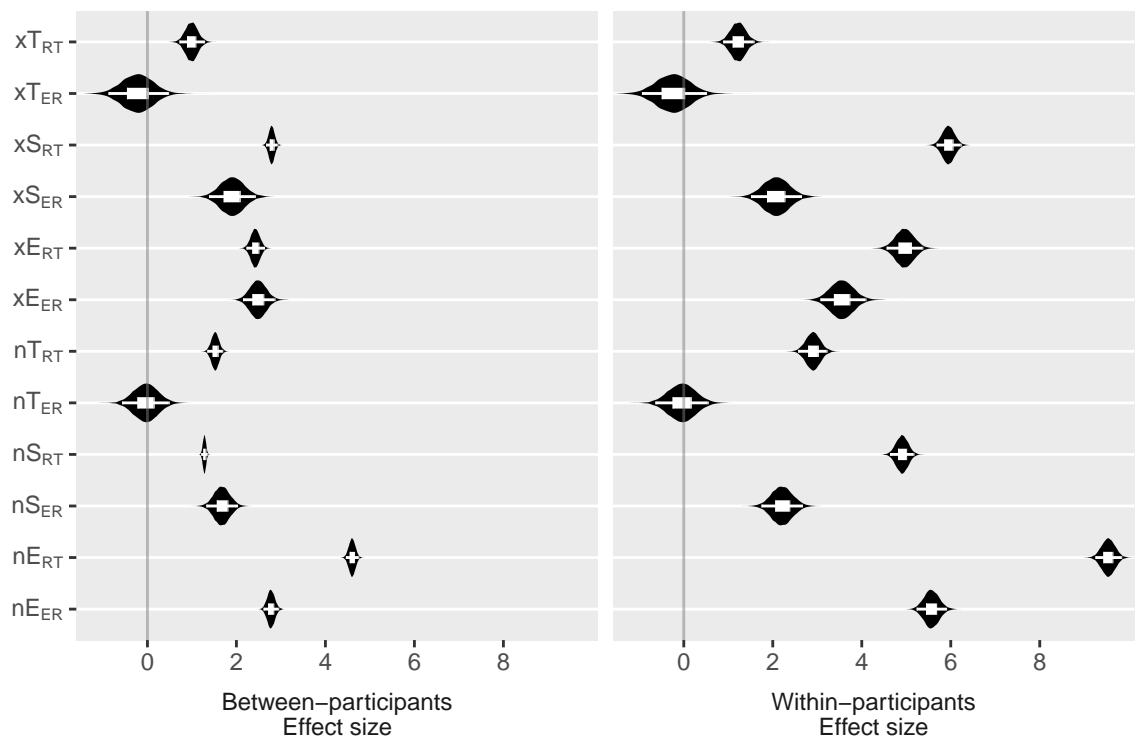


Figure 2.8: Between- and within-participant effect sizes for primary measures of the CAST.

effect relative to the amount of within-participants variability associated with that effect (MacLeod, Lawrence, McConnell, Eskes, Klein & Shore, 2010). The former “between-participants effect size” is useful to convey how consistently one might expect an effect to manifest across a population and the latter “within-participants effect size” is useful to convey the certainty of measurement for a given participant. The between-participants effect size is straight-forward to compute, as the model includes an estimate of the between-participants standard deviation for each effect; the value of a given effect is divided by this standard deviation to yield an effect size in each sample from the posterior. Thus computed, this between-participants effect size conveys the magnitude of the effect relative to the variability of scores expected to manifest across participants, a scale for which standard guidelines for interpretation (Cohen, 1988) apply such that values of 0.8 or greater are considered “large” effect sizes.

A more involved computation is necessary to derive the within-participants effect size as there are no explicit parameters in the model directly related to the within-participants standard deviation of each effect. Instead, we compute a within-participants standard deviation for a given effect by computing the standard deviation across posterior samples of estimates for that effect for each participant, then averaging the resulting per-participant standard deviations; the value of the effect is then divided by this average within-participant standard deviation to yield an effect size in each sample from the posterior. Thus computed, this latter within-participants effect size conveys the magnitude of the effect relative to the average uncertainty associated with scores from individual participants. As the precise computation of this within-participants effect size is (to our knowledge) novel, mapping to colloquial interpretation is not established, but we see no reason in the mathematics to deviate from the above noted guidelines where 0.8 values of 0.8 or greater are considered “large” effect sizes.

As shown in Figure 2.8, within-subjects effect sizes are generally larger than between-subjects effect size, with nE manifesting the largest effect sizes but otherwise relatively large effect sizes for all effects except for the effects of temporal attention on error rate, which is expected as they did not manifest a credible effect.

2.3 Discussion

We observed in this experiment that, when employed among young adults, the measures of the CAST are reliable, stable through repeated testing, and can be expected to yield relatively large effect sizes in both within-participants and between-participant experimental designs. We also observed that, through its improved theoretical foundations, the CAST is able to discern more nuanced interactions and correlations among the measured phenomena of attention. For example, the CAST isolates the previously observed correlation between Orienting and Executive to a correlation between the exogenous form of spatial attention and the endogenous form of executive attention. Similarly, while prior research using the ANT-I had suggested an interaction between Alerting and Orienting networks, through the CAST we are able to discern that this interaction manifests for the endogenous forms of these systems only. Previous research using the ANT also observed an interaction between Alerting and Executive networks, and again the CAST is able to discern that this interaction manifests for the endogenous forms of these systems only.

The relatively high reliabilities (50%CrIs that are $>.5$) observed across all credibly-non-zero primary CAST effects, contrasts with previous reports of relatively low reliabilities for their measures of Alerting and Orienting. Three accounts come to mind that may, individually or in some combination, explain this difference:

- The more theory-consistent manipulation and measurement of the CAST effects may reduce confounds that otherwise induce extraneous variability in the manifestation and measurement of effects in the ANT (ex. use of cues that engage both nS and xS).
- The attempts in the CAST to increase the game-like nature of the task (use of gamepads for response, speed incentive with feedback, etc.) may increase participant engagement and thereby reduce noise otherwise manifest when participants are less engaged.
- The analysis of RT as a log-normal outcome and ER as a binomial outcome may have reduced noise in the per-participant estimates for each measure produced by the model.

While evaluation of the first two accounts require additional experiments and data collection, re-analysis of data from mega-studies like that of MacLeod et al. (2010) would be able to assess the influence of analytic approach and is thereby recommended.

Having demonstrated the utility of the CAST for the measurement of attention among healthy young adults, we next repeat this evaluation in a population of healthy children, where the measurement of attention is of great theoretical and practical interest for the elaboration of its developmental trajectories and early diagnosis and remediation of its disorders.

Chapter 3

Experiment 2: Children

The data reported here as Experiment 2 are derived from a study conducted by Johnson, Snow, Lawrence and Rainham (submitted) who sought to explore the effects of nature versus urban environments on cognition in children. In this study, children completed the CAST both before and after a group-randomized intervention involving exposure to either nature or urban environments. As is reported in a separate manuscript currently under review, a selective effect of the intervention on endogenous modes of both spatial and temporal attention were observed. For the present purposes, use of this data set while ignoring intervention group provides the opportunity to investigate the psychometric properties of the CAST in children as achieved in E1 in young adults. As the age range collected here (8-16) nearly overlap with those of E1 (18-24), this analysis will be followed by a combined analysis of both data sets in Chapter 4, which will explore the developmental trajectories of attention.

3.1 Methods

3.1.1 Participants

Seventy children aged 8 to 16 years participated in this study. Participants were recruited through community bulletins, newsletters, paid advertisements, and emails distributed to families that had previously participated in studies in the Johnson laboratory. Children were required to have an estimated IQ equal or greater than 80 (established by prior testing conducted in the Johnson laboratory), normal or corrected-to-normal vision, no history of psychiatric/psychological diagnoses, no history of severe head injury, and no significant neurological disorders affecting the central nervous system. The study received ethical approval from the local Research Ethics Board. Participants received \$15 as an honorarium for study participation,

and parents were entered to win one of two \$50 gift cards to a bookstore as compensation for their time. Participants were assigned to one of two conditions: an urban walk or nature walk. The logistics of testing outside of a laboratory made true random assignment impossible for this particular study. However, participants blindly assigned themselves to study condition; participants were informed that there were two possible locations to which they could be assigned and then were asked to select a participation date, following which the testing location pre-scheduled for that date was revealed.

Thirty-eight children (16 males) participated in the urban condition and 32 children (15 males) participated in the nature condition. Participants were primarily Caucasian (82.6%). Of the 70 participants that completed the study, the CAST data of 10 participants were excluded due to either performing at chance during one of the sessions or failing to respond on more than 30% of trials. Thus, the analyses for the CAST were conducted on a sample of 60 participants, 30 in each condition, with 16 females in the nature condition and 15 females in the urban condition.

3.1.2 Materials

The experiment was run on multiple Macbook Pro laptops with varying screen sizes but placed such that all stimuli subtended the same visual angle for all participants. As in Experiment 1, participants wore earphones through which auditory stimuli were presented at a comfortable volume and made responses using the analog triggers on an xbox360 gamepad.

3.1.3 Procedure

As in E1, participants in this experiment completed two sessions of the CAST. Unlike E1, where the two sessions were completed one immediately after the other, in this experiment participants completed a 30 minute walk between the first and second sessions. Procedures were otherwise identical to those employed in E1.

3.2 Results

3.2.1 Data pre-processing

Data were pre-processed as in Experiment 1, including removal of all trials on which:

- Participants failed to make a response (0.1% of trials)
- Responses were made prior to target appearance (0.6% of trials)
- Response times were faster than 200ms (0.3% of trials)

As in Experiment 1, the latter exclusion criterion was determined by both prior experience with RT data suggesting that responses faster than 200ms tend to be anticipatory responses unrelated to target processing, as well as application of a generalized additive model of trial accuracies predicted by trial RTs, showing that only above about 200ms do responses rise above chance performance.

3.2.2 Statistical modelling

Data were modelled as in Experiment 1 with the addition of the between-participants effect of the nature/urban intervention on all coefficients. Even though they will not be discussed in detail below, it is necessary to include the effects of the intervention in the model to achieve an accurate estimate of between-participants variance in effects; omitting the effects of the intervention would yield inflated estimates of between-participants variance to the degree that such effects are non-zero.

3.2.3 Primary effects and interactions

Figure 3.1 displays the posterior distributions associated with the primary measures of the CAST as well as their interactions, and conveys that the primary effects manifest as expected, as does the $nE:xE$ interaction between. As in the young adult population of E1, here we observe both $nT:nE$ and $nT:xE$ interactions, as well as an $xS:nE$ interaction. All these interactions maintain the patterns of those observed in E1. In contrast, while both the children here and young adults of E1 manifest an $xS:xE$ interaction, they manifest in opposite patterns; in the young adults, the xS effect on ER is larger in the xE -congruent condition than in the xE -incongruent condition

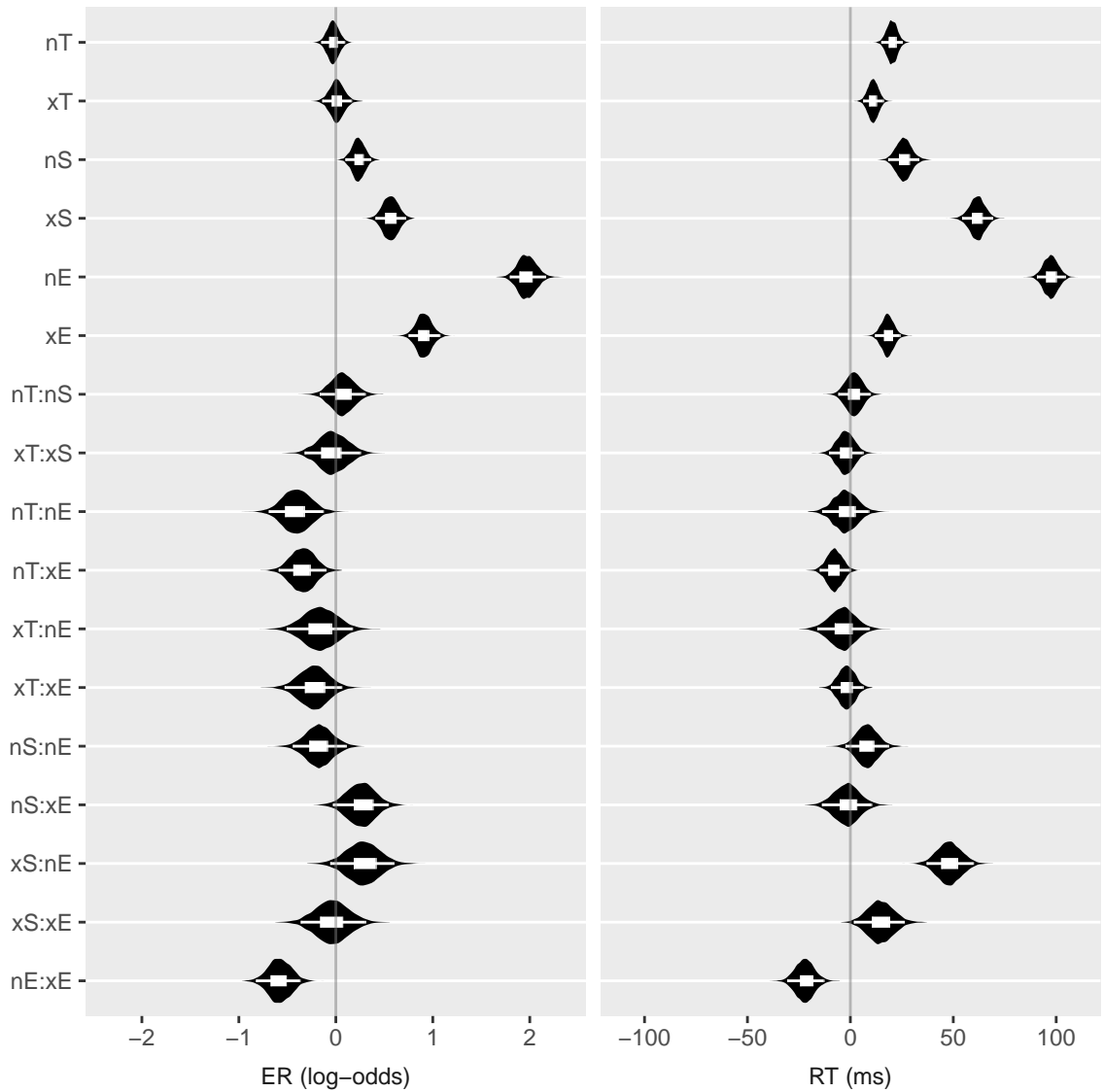


Figure 3.1: Primary effects and interactions in the CAST. Plots of each effect’s constituent conditions can be found in Appendix D. Readers viewing this document electronically should be able to click any row of the plot to jump to that effect’s corresponding figure in Appendix D.

(equivalently: the xE effect is larger in the xS -valid condition than in the xS -invalid condition), yet in the children the xS effect on RT is larger in the xE -incongruent condition than in the xE -congruent conditions (equivalently: the xE effect is larger in the xS -invalid condition than in the xS -valid condition). Interactions present in the adults but absent in the children include the $nT:nS$ and $nS:nE$ interactions (both notably involving nS). There are no interactions present in the children that were absent in the adults.

3.2.4 Effects of Session

Effects of session are shown in Figure 3.2. As in the young adult sample of E1, children show a reduction of overall RT as well as a $xT:nE$ interaction, this latter manifesting in children (see Figure 3.3) in a similar pattern as in adults (Figure 2.3).

The children manifest several credible effects of session that are not observed in adults, including: an effect of session on xS such that the effect of xS on RT is smaller (though still credibly-non-zero; 95%CrI= 48ms – 65ms) in the second session than the first (note though that the adults have an indecisively credible effect in the same direction); an effect of session on the $nS:nE$ interaction such that (see Figure 3.4) this interaction is only credibly present in the second session (manifesting in that session as larger nS effect on ER in the nE -congruent condition and a larger nE effect on ER in the nS -valid condition); and an effect of session on the $nS:xE$ interaction such that (see Figure 3.5) this interaction is credibly present in neither session alone but switches sign in it's direction from one session to the next.

Effects of session that were present in adults but are absent in children include the effects of session on nT , nS , and nE .

3.2.5 Correlations

Figure 3.6 shows that the set of effects in children where the effect on ER is correlated with the effect on RT replicates those observed in young adults with the exception that the correlation between nS_{RT} and nS_{ER} observed in young adults is not observed in children, and a correlation between nT_{RT} and nT_{ER} is observed in children but not young adults. Note also that the strength of the correlation between I_{RT} and I_{ER} is much stronger in young adults than in children.

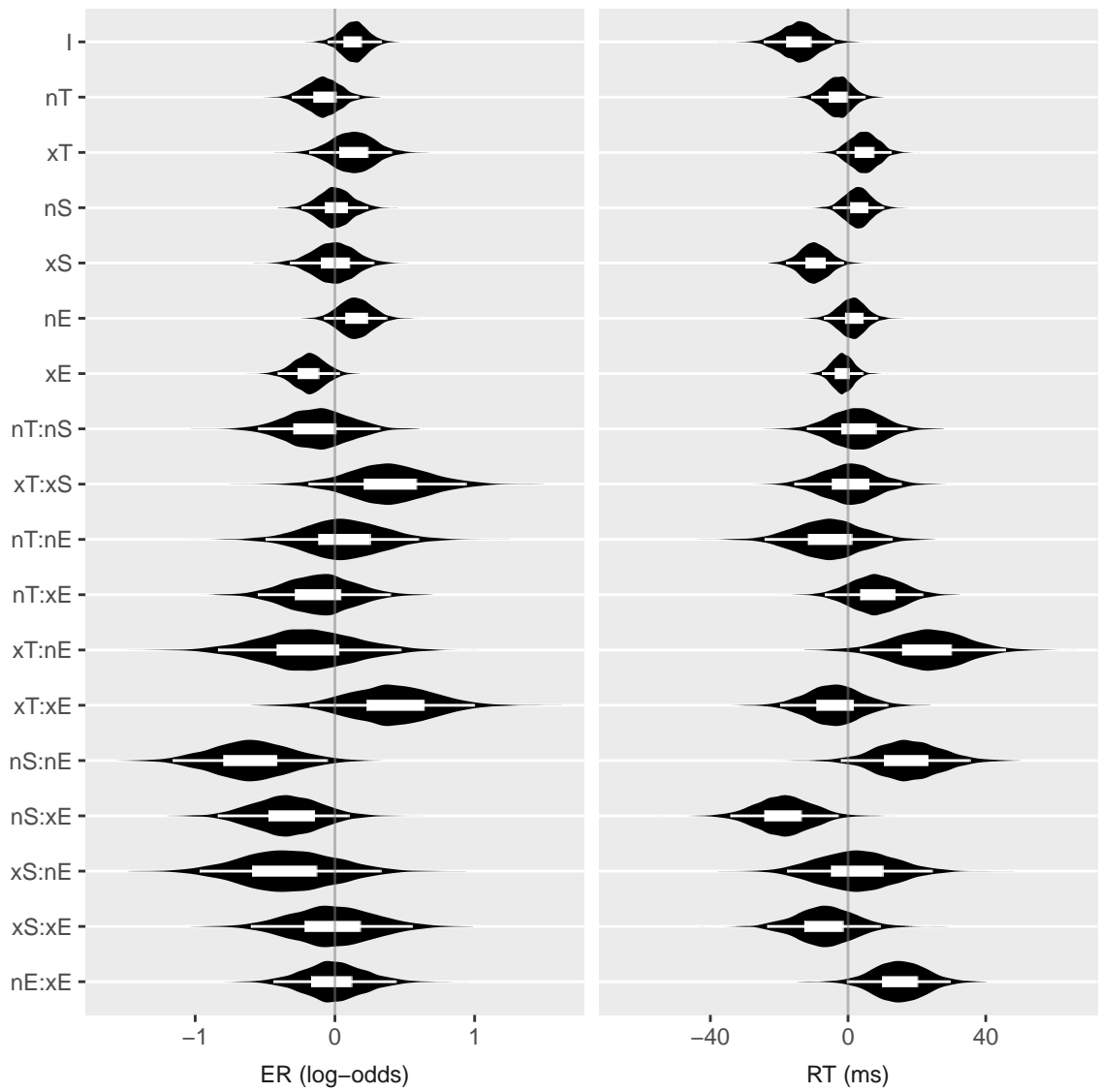


Figure 3.2: Effects of Session (session 2 minus session 1) on primary effects and interactions in the CAST.

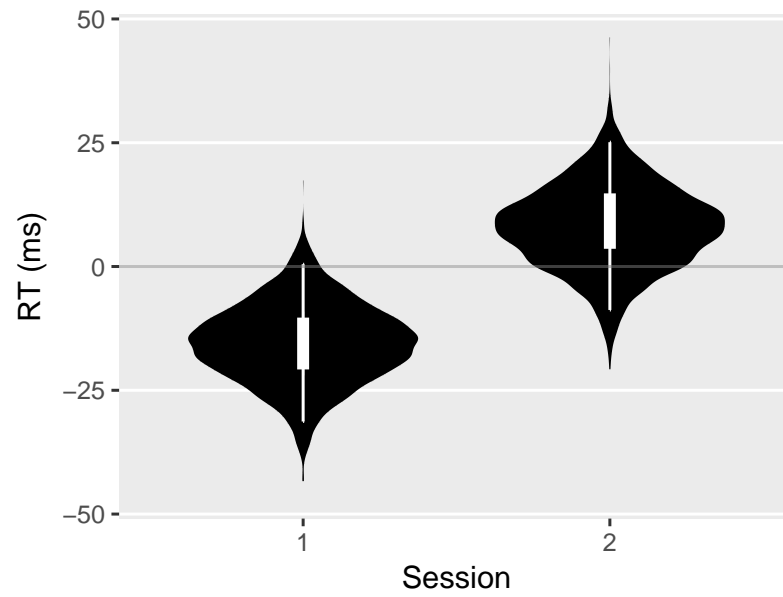


Figure 3.3: Interaction between xT and nE and in each session.

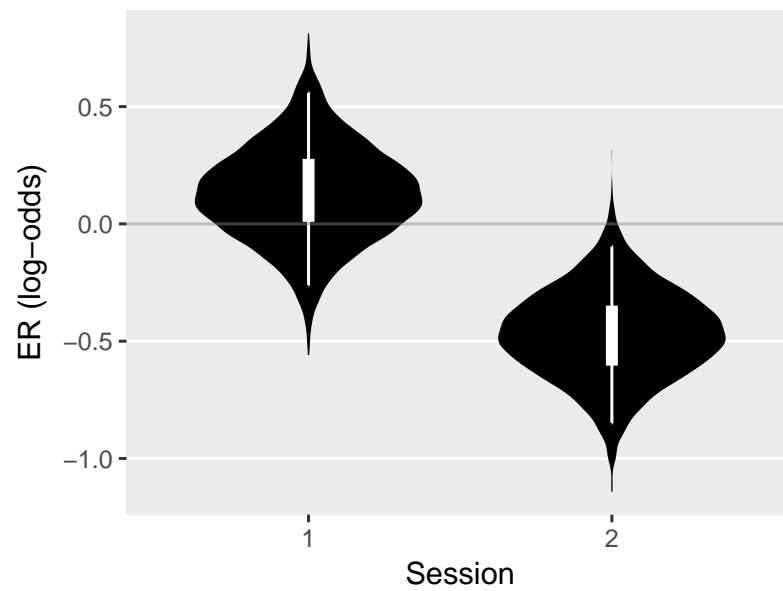


Figure 3.4: Interaction between nS and nE and in each session.

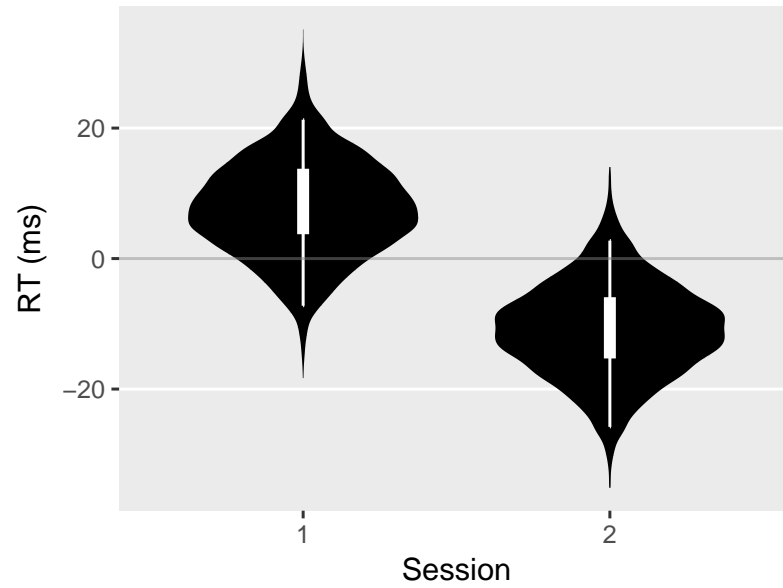


Figure 3.5: Interaction between nS and xE and in each session.

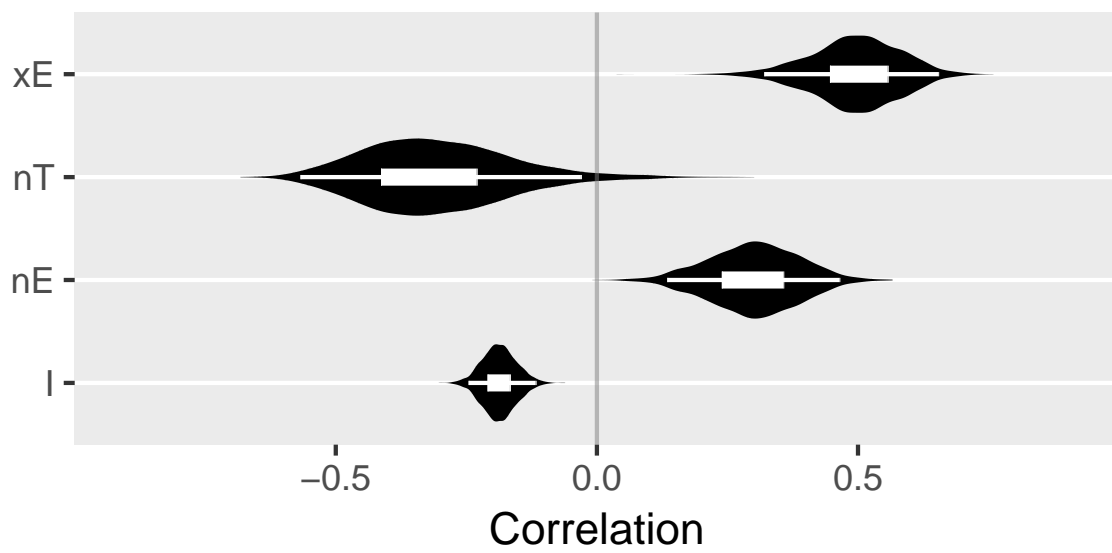


Figure 3.6: Credibly non-zero correlations between outcomes (ER & RT) within a CAST effect.

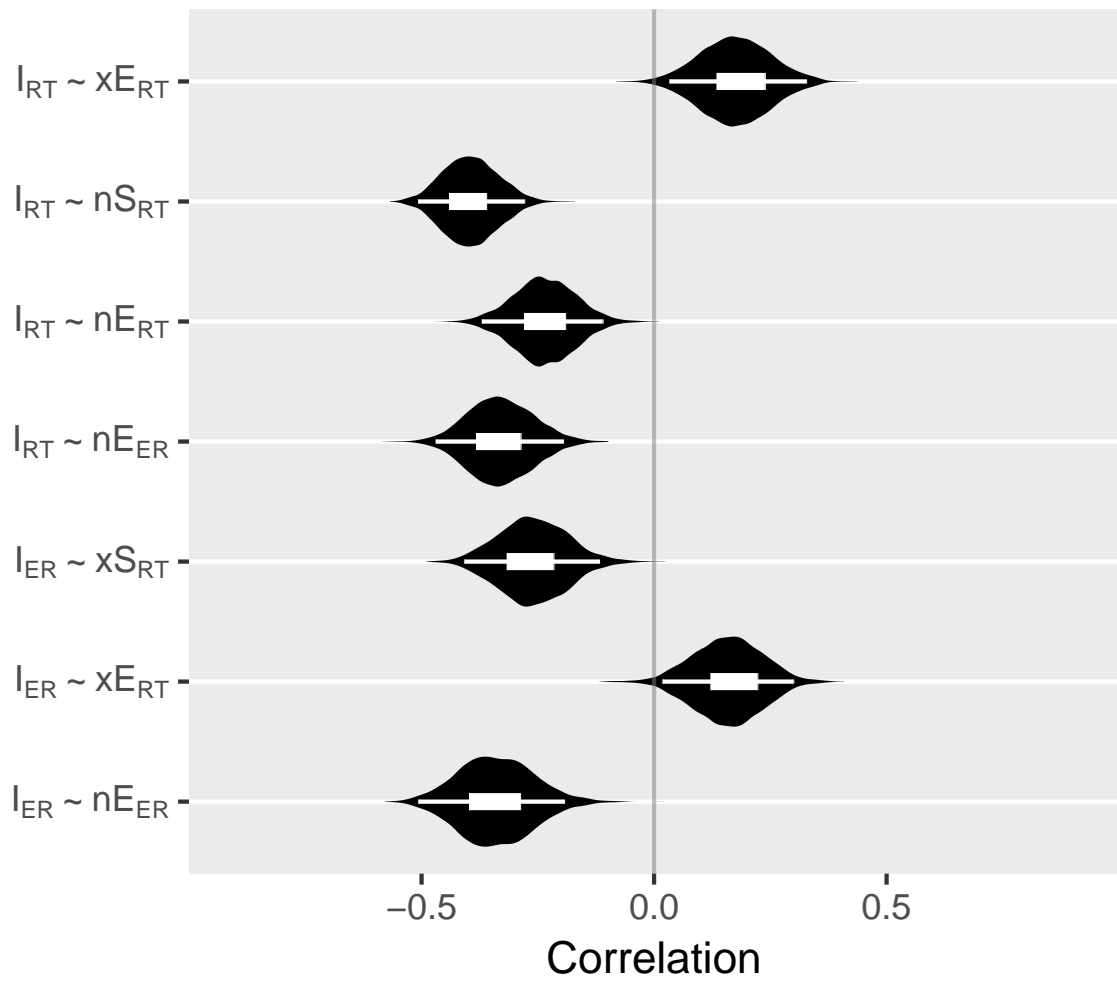


Figure 3.7: Credibly non-zero correlations involving the intercept.

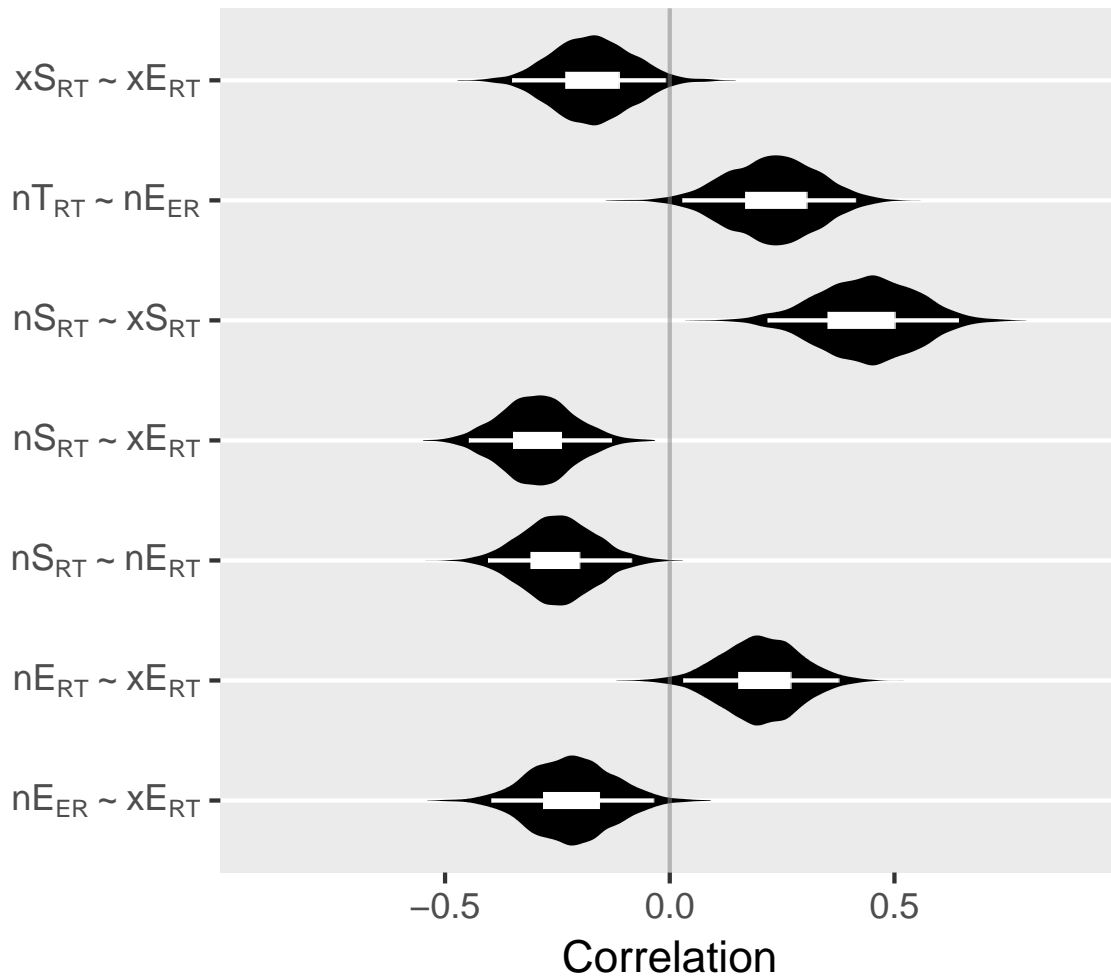


Figure 3.8: Credibly non-zero correlations between different attention systems.

Figure 3.7 shows the credibly non-zero correlations involving the intercept, where we observe that all attention subsystems measured by the CAST involve at least one correlation with overall performance with the exception of nT and xT (c.f. young adults, in whom only xT failed to correlate with overall performance).

Figure 3.8 shows the credibly non-zero correlations among the CAST measures, of which there are more than manifest in the adults, where we observed correlations between nT and nE , as well as between xS and nE . In children, we observe the same correlation between nT and nE , but fail to observe a correlation between xS and nE . Instead, we observe multiple additional correlations, including between: xS and xE ; nS and xS ; nS and nE ; nS and xE ; nE and xE . Of the 60 correlations among the measures, 32 have posterior distributions in which zero falls within the 50% credible

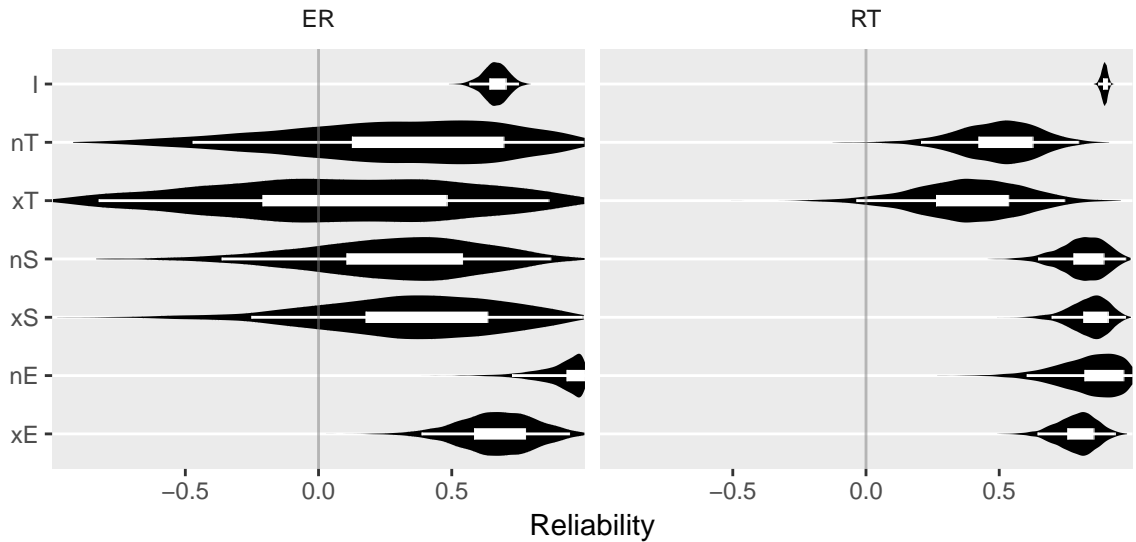


Figure 3.9: Reliability for each of the primary measures in the CAST.

interval, and that the median interval of these ranged from -0.08 to 0.10.

3.2.6 Reliabilities

As in E1, reliability of the primary measures in the CAST were computed and are presented in Figure 3.9, showing relatively reliable measurement for all measures in RT with the exception of both nT and xT , for which values reflecting low reliability (0 to 0.5) maintain reasonable credibility. In ER, the posterior distributions for both nT and xT remain relatively wide and centered at zero, which is expected for effects that are fairly confidently zero in the first place. The reliability of both nS_{ER} and xS_{ER} is relatively middling, with a 50% CrI centered below .5 and 95%CrI extending into the negative domain. Finally, both xE and nE achieve relatively reliable measurement in ER.

3.2.7 Effect sizes

Both between- and within-participant effect sizes were computed as in E1. As shown in Figure 3.10, all measures achieve relatively large effect sizes save those associated with the effect of temporal attention on error rate, which is expected as they did not manifest a credible effect. The only notable difference in these effect sizes obtained in children as compared to those obtained in young adults is that the effect size for

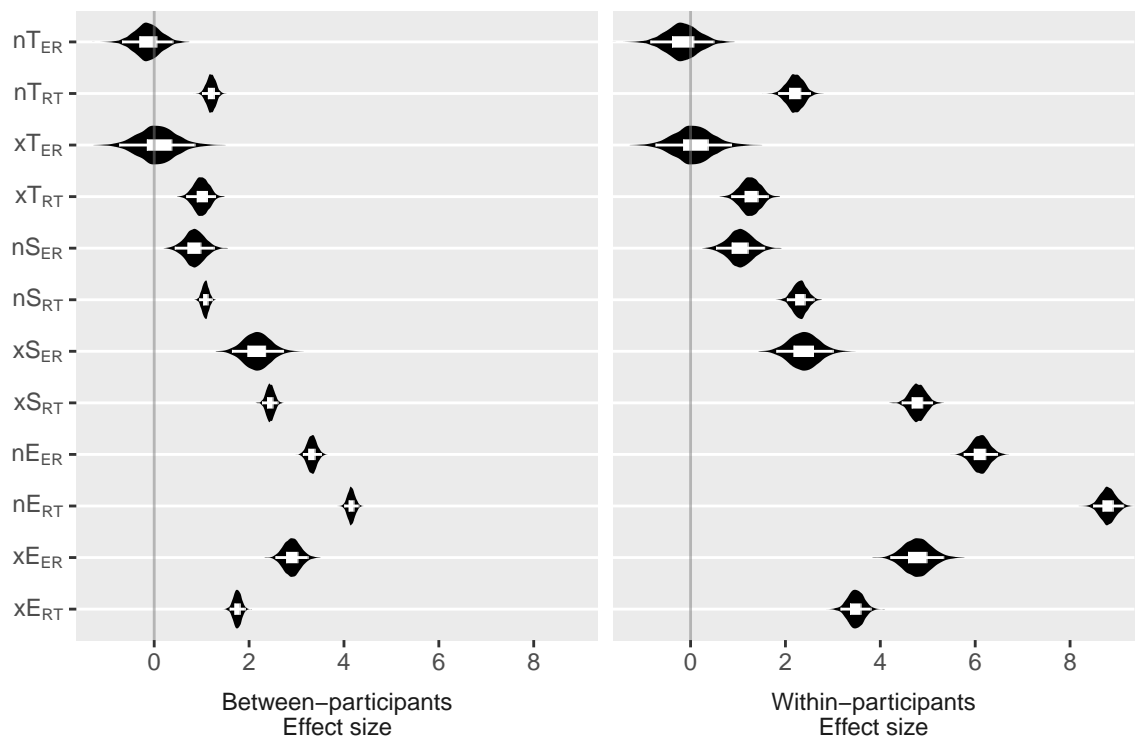


Figure 3.10: Between- and within-participant effect sizes for primary measures of the CAST.

nS_{RT} and nS_{ER} are both substantially lower in children than young adults.

3.3 Discussion

As in adults, we find that the CAST produces reliable and stable measurement of the systems of attention. While there is some reduction of reliability of measurements in children relative to adults, this is as expected from a population with still-developing attention systems and concomitant performance variability. E2 also observed more inter-relatedness among the CAST measures as operationalized by the count of credibly non-zero correlations observed in Figure 17. In contrast to the adults in E1, where a mere 3 of 60 correlations were observed to be credibly non-zero, 7 of 60 were observed as such in the children of E2. While this is still a small degree of interrelatedness, it is nonetheless worth noting that even this amount of relatedness may be accounted for by increased developmental variability in the child sample of E2 relative to that of the adult sample of E1. That is, for a given pair of measures, if each manifests its own independent non-zero developmental trajectory across the observed sample of participants, such trajectories will manifest as a spurious correlation between that pair of variables. Given the broad age range (Ages 8 – 16) included in E2, it is plausible that such developmental trajectories thus account for the observed correlations. This hypothesis can be tested directly by combining the data from both experiments and modelling the effects of age across the combined sample; if no individual trajectories are observed for a pair of correlating variables, then the observed correlations must be attributable to common causal mechanisms other than sheer development.

Chapter 4

Development of attention as observed by the CAST

Given that the samples from E1 and E2, when combined, form a relatively continuous sample of ages from 8 to 24 (see Figure 4.1), we took this opportunity to explicitly evaluate the effect of age on performance in the CAST.

4.1 Results

4.1.1 Data pre-processing

Data were pre-processed as reported in the results sections of E1 and E2, then combined to form a single data set.

4.1.2 Statistical modelling

The same models as used in E1 and E2 were used here with the addition of parameters to achieve inference on the linear effect of age on all measures.

4.1.3 Effects of age

Figure 4.2 displays posterior distributions for the effect of age on all measures, revealing an effect on both I_{ER} and I_{RT} such that older participants were both faster and more accurate. There was also an effect of age on both nS_{ER} and nS_{RT} such that the magnitude of both these effects was larger in older participants. As shown in Figure 4.3 it appears that this increase manifests in ER as an increase from an nearly-zero nS at younger ages to a substantial nS at older ages. The effect of age on nE appears to manifest opposite effects on RT and ER; as shown in Figure 4.4, while both measures of nE are credibly non-zero and positive throughout the age range observed, the magnitude of nE_{RT} reduces among older participants while the magnitude of nE_{ER} increases among older participants. As shown in Figure 4.5, xE_{RT} reduces among older participants, though maintaining a credibly non-zero and positive magnitude in

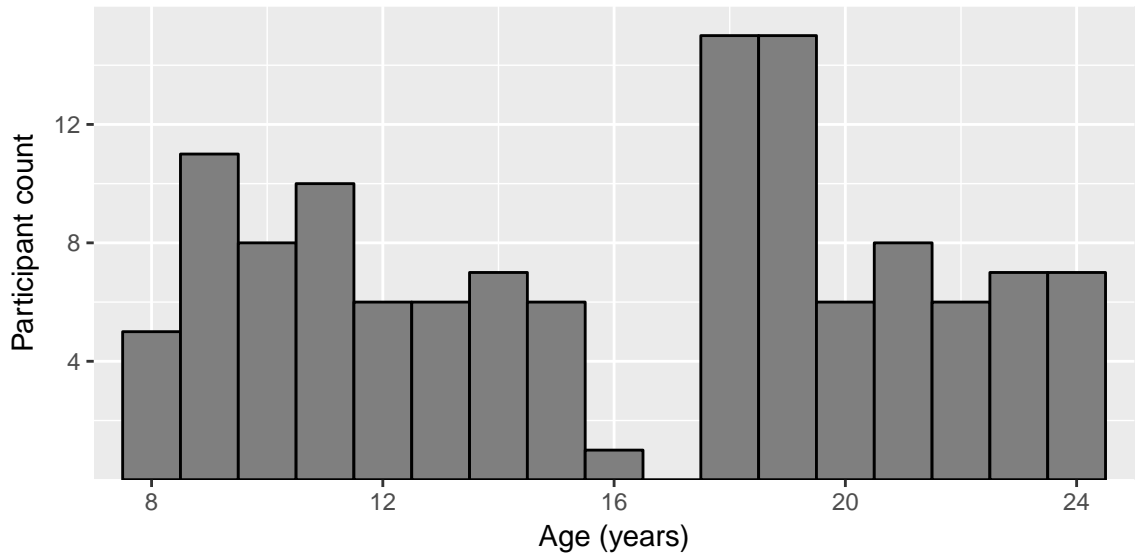


Figure 4.1: Distribution of participant ages from E1 and E2.

the older ages observed. Finally, as shown in Figure 4.6, the interaction between nE and xE as manifest in ER reduces among older participants from a credibly non-zero and negative magnitude at the younger ages observed to a nearly-zero magnitude at the older ages observed.

4.1.4 Correlations

In the separate analyses of the two age groups, credibly non-zero correlations were observed and it was subsequently noted that these might be an artifact of unmodelled developmental effects. Having explicitly modelled and observed several such effects, it is useful to re-evaluate the correlations among CAST measures. Specifically, when effects of age are included in the model, it is possible to evaluate the correlations among CAST measures having taken possible parallel age effects into account.

Figure 4.7 shows credibly non-zero correlations between ER and RT measures within the same effect, where we observe the same set of correlations as observed in E1 such that overall RT and ER are negatively correlated while there are positive correlations in the manifestation of effects in RT and ER for nS , nE and xE , reflecting that for these measures, participants with larger effects in RT also had larger effects in ER.

Figure 4.8 shows credibly non-zero correlations involving the intercept, where

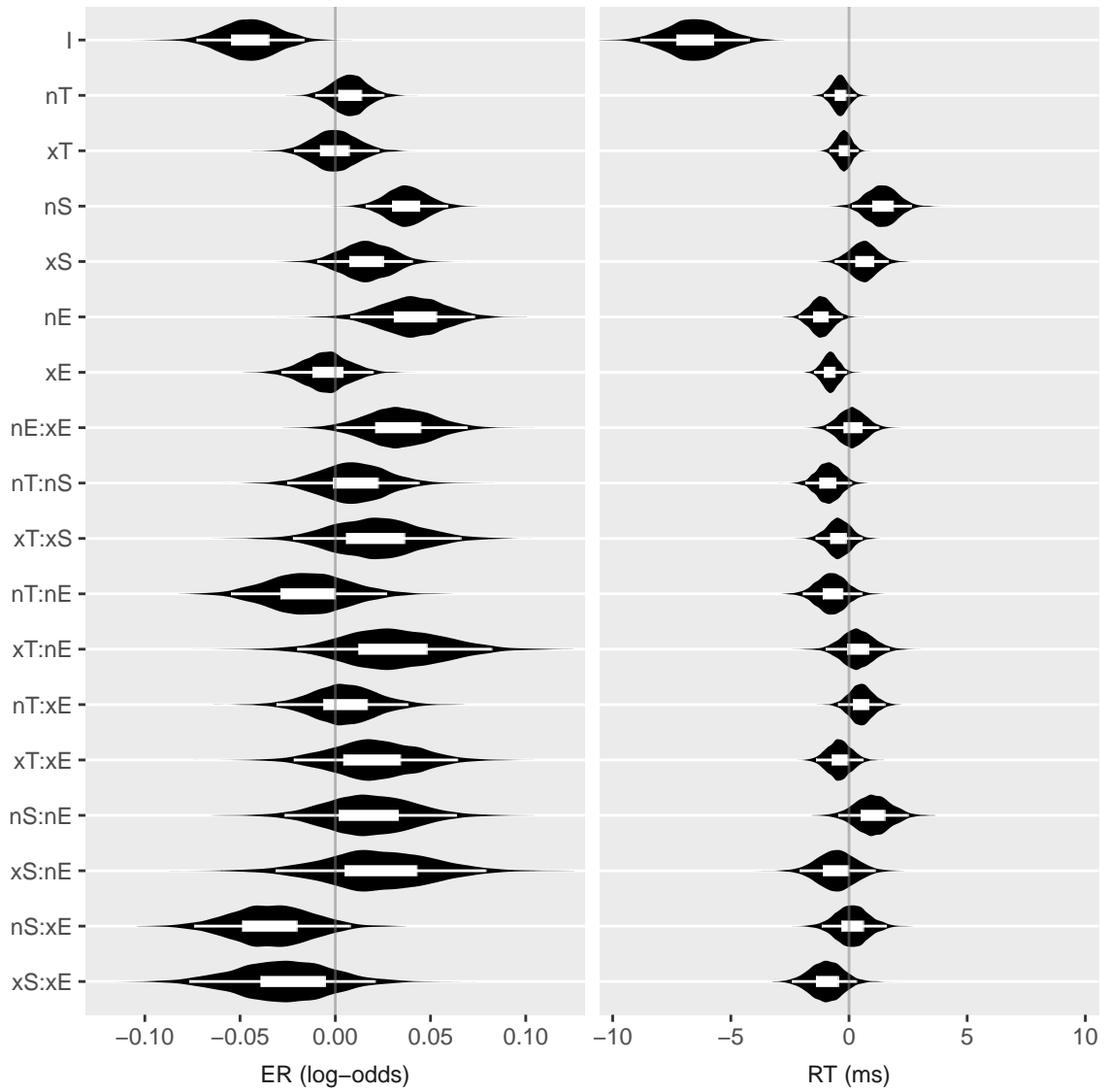


Figure 4.2: Effects of Age (change per year) on primary effects and interactions in the CAST.

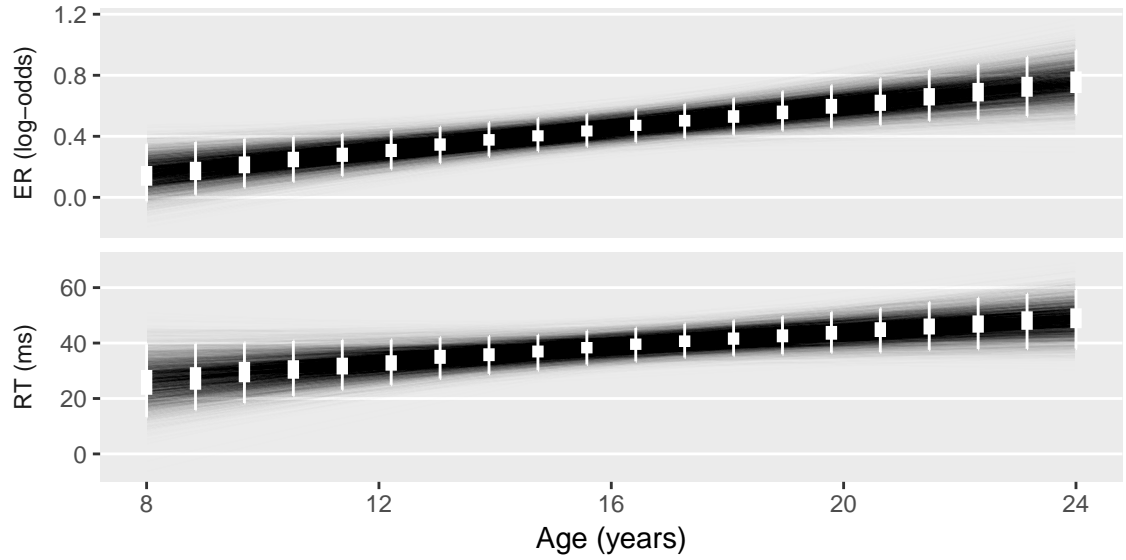


Figure 4.3: Interaction between age and nS .

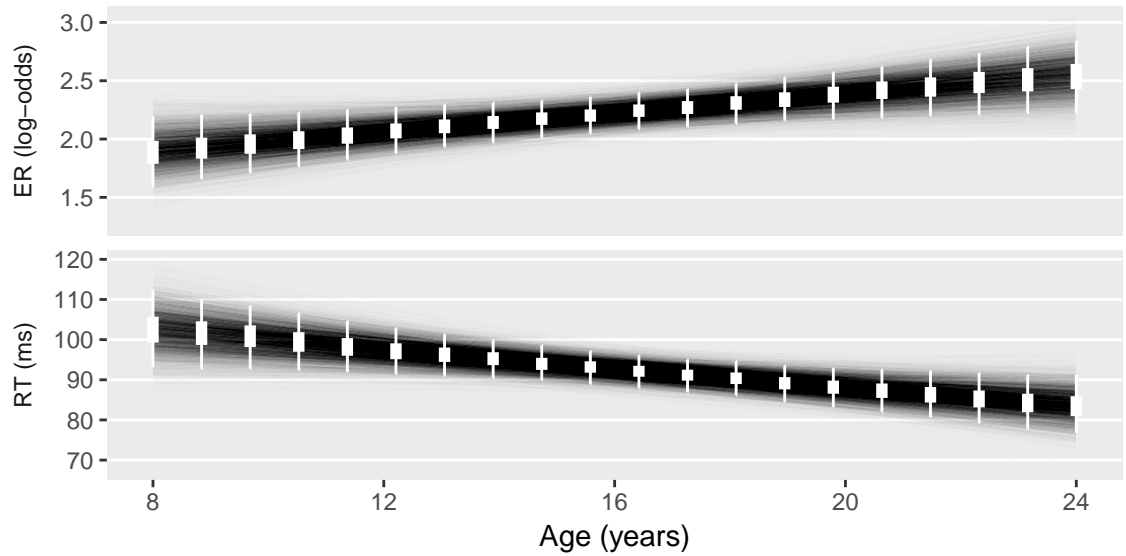


Figure 4.4: Interaction between age and nE .

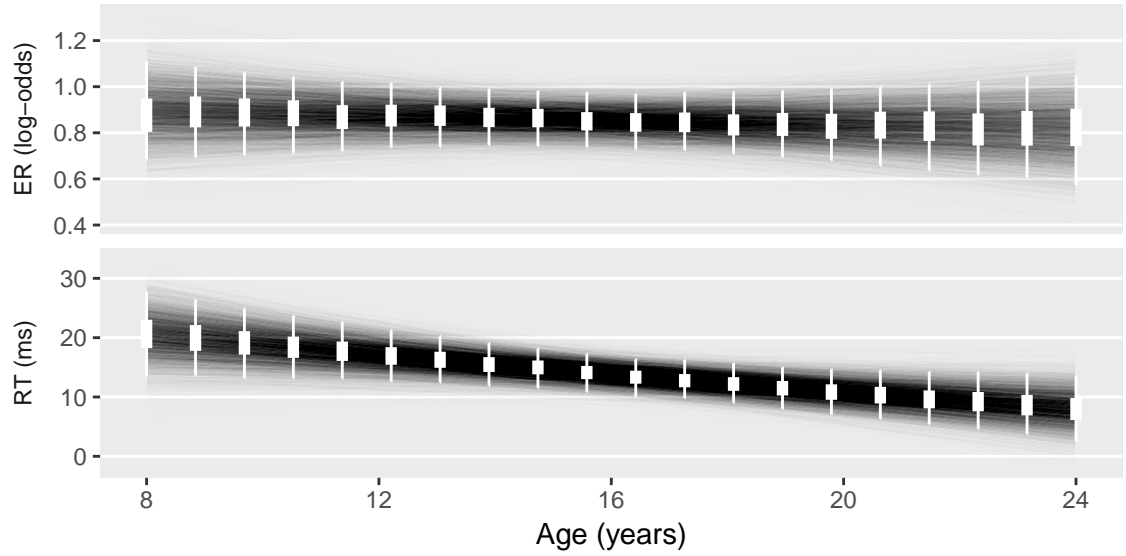


Figure 4.5: Interaction between age and xE .

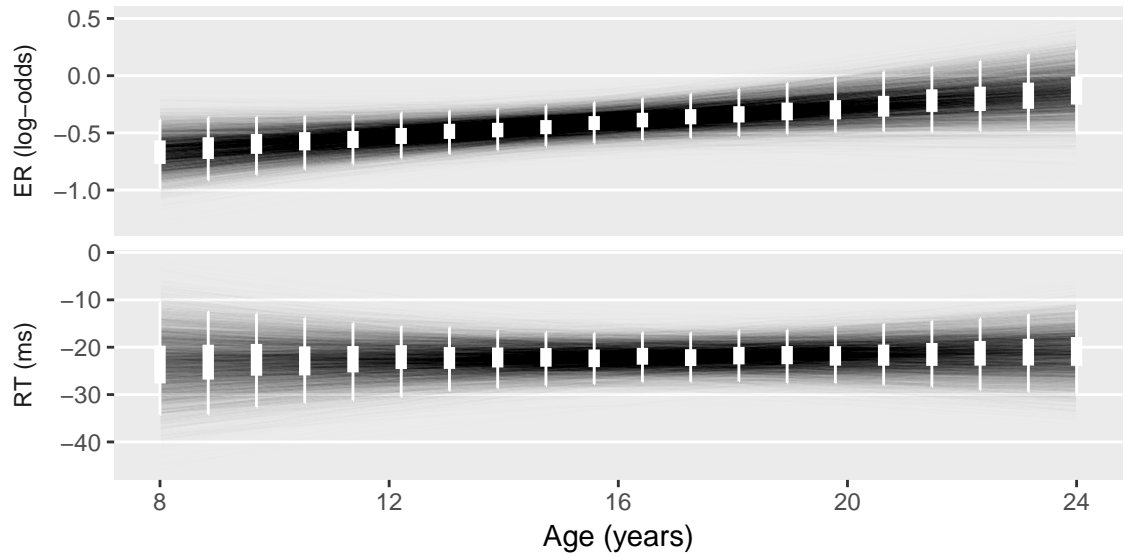


Figure 4.6: Interaction between age, nE and xE .

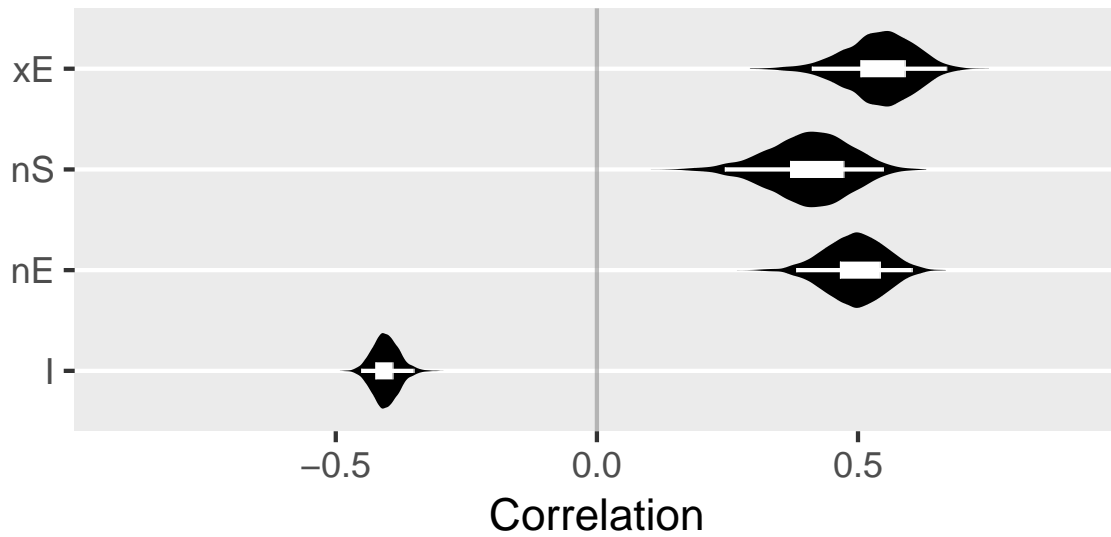


Figure 4.7: Credibly non-zero correlations between outcomes (ER & RT) within a CAST effect.

we observe that all attention subsystems measured by the CAST involve at least one correlation with overall performance, however most of these correlations are relatively small (credibly less than .5).

Finally, Figure 4.9 shows credibly non-zero correlations among the CAST measures, where we observe only one correlation, between nT_{RT} and nE_{ER} , validating that all other correlations observed in the separate analyses of the two age groups were artifacts of parallel age effects. Posterior distributions for the full set of 91 correlations are depicted in Appendix E.

4.2 Discussion

The results from the developmental analysis provide both replication of prior observations as well as new findings. First, replicating the results of Mullane, Lawrence, Corkum, and Klein (2016), we observe no age effects on xS , as the neural mechanisms associated with exogenous spatial attention are likely fully developed by the minimum age in this sample. In contrast, we do observe large age effects on nS , suggesting ongoing development of endogenous spatial attention across the age range of this sample.

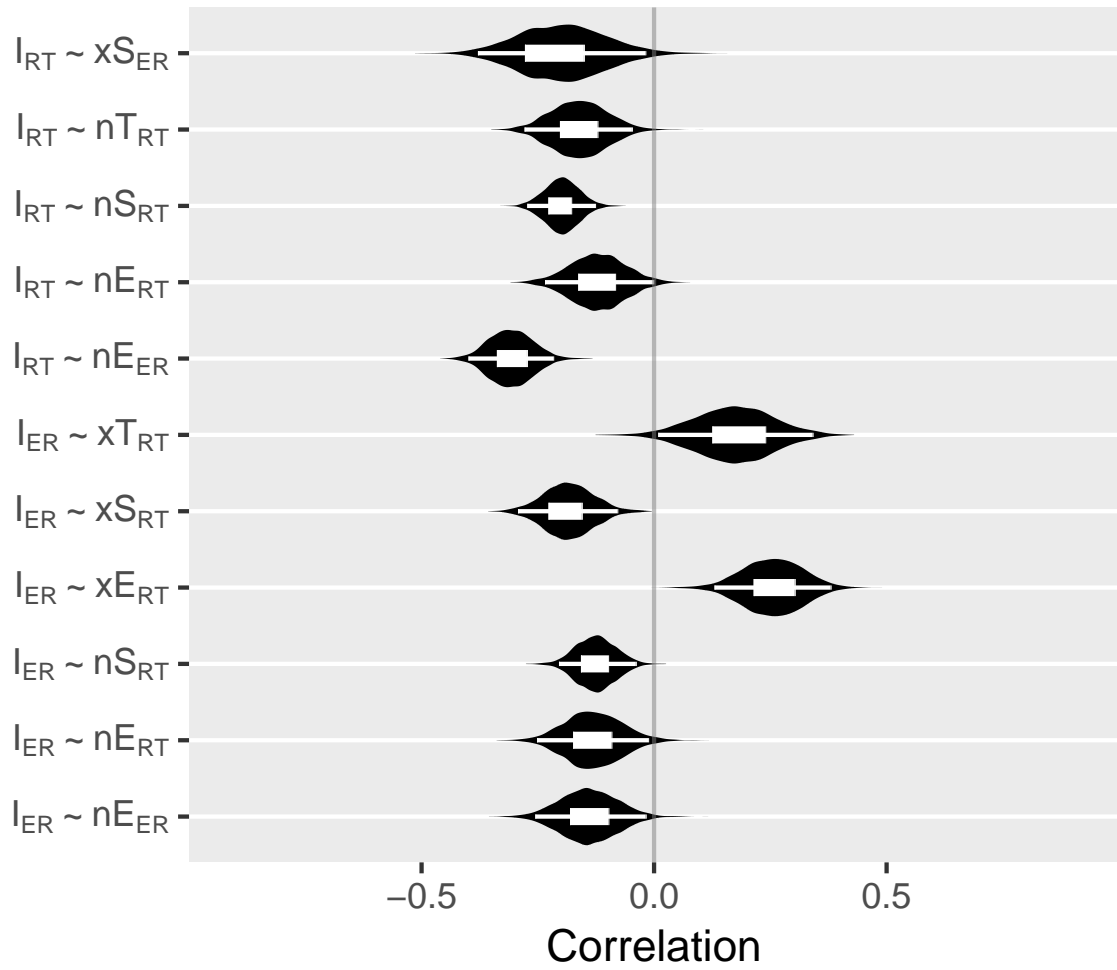


Figure 4.8: Credibly non-zero correlations involving the intercept.

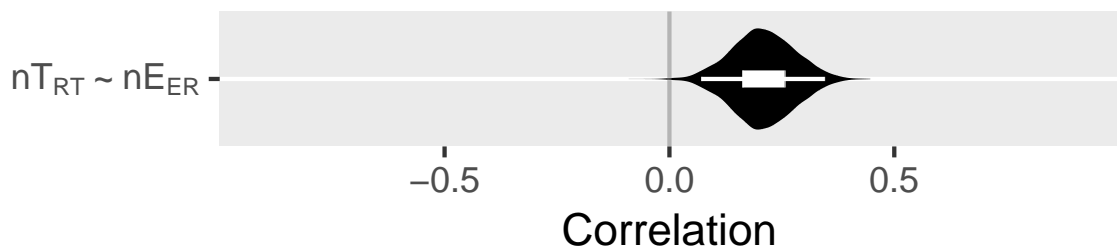


Figure 4.9: Credibly non-zero correlations between different attention systems.

Furthermore, no effects of age are apparent among the measures of temporal attention, contrasting with results from both the ANT (which used visual stimuli to measure alertness; Rueda et al., 2004) and the ANT-I (which used auditory stimuli to measure alertness; Mullane et al., 2016), where decreases in the magnitude of the alertness network score at older ages have been observed. The source of these differences is unclear, as both the ANT and ANT-I employ stimuli to manipulate alertness that should be expected to engage both endogenous and exogenous temporal attention, while the CAST seeks their independent manipulation and measurement. It may be that only when simultaneously engaging both nT and xT does a developmental trajectory manifest.

While we do observe the effect of age on nE_{RT} previously observed in the ANT and ANT-I such that the nE_{RT} effect gets smaller at older ages, we also observe an age effect in the opposite direction on nE_{ER} whereby the nE_{ER} effect gets larger at older ages. Since error rates are rarely reported and even more rarely treated in the statistically proper manner (as a binomial outcome) as they are in the above modelling, it comes as no great surprise that the effect of age on nE_{ER} has been missed by previous work. Specifically, it has been demonstrated (Dixon, 2008; Jaeger, 2008) that analysis of binomial data collapsed to a proportion can lead to an increase in both false-positive and false-negative errors for the evaluation of interactions, while explicit modelling of the raw data as a binomial outcome avoids these issues. So the failure of prior research to observe the interaction between Age and nE_{ER} may be a consequence of inflated false-negative error associated with analysis of binomial data collapsed to a proportion. Given the observation of an effect of age on nE_{ER} , it is worth noting that the pattern whereby age has an opposite effect on the manifestation of an nE in RT and ER is consistent with an explanation whereby there are no developmental effects on nE at all, but an overall developmental trend whereby participants shift their speed-accuracy criterion as they age. That is, compared to participants of middling age in the present sample, younger participants may prioritize accuracy over speed, leading to larger effects of nE on RT and smaller effects of nE on ER, while older participants may prioritize speed over accuracy, leading to larger effects of nE on ER and smaller effects of nE on RT. While there may additionally exist true changes in information processing in addition to the observed speed-accuracy trade-off, more

advanced modelling than presented above would be required to discern them (see General Discussion).

Previous reports observing the decrease in nE_{RT} with age have interpreted this to signal improved capacity for conflict resolution at older ages, and while the results above for nE do not support this account, the results for xE do; xE_{RT} decreases with age, replicating previous observations of the same trend (Williams, Strauss, Hultsch & Hunter, 2007) and likely reflecting a process whereby older participants have learned how to better manage the conflict manifest by stimuli located in space that require responses also located in space, but orthogonally so. The final developmental effect observed here also involves xE whereby the magnitude of the interaction between xE and nE appears to diminish for older ages, at least as measured in ER. This result is somewhat unexpected insofar as the operational differentiation of spatial Stroop (xE) and the more general Simon effect is the use of a target stimulus with a response property that has a well-learned spatial association, and a characteristic behavioural differentiation of spatial Stroop and the more general Simon effect is that spatial Stroop interacts with the flanker effect (nE) while the Simon effect does not. It might have been expected, then, that as participants age and gain more experience with arrows as spatial stimuli in their everyday lives, the magnitude of the interaction between xE and nE would increase (or at least remain constant). It is possible, then, that by the minimum age of this sample participants have already achieved as strong a spatial association as is possible for arrows, and that the observed decrease in the xE and nE interaction reflects a change in the mechanism by which this interaction manifests in the first place.

Finally, it was noted in the discussion of E2 that the multiple correlations observed in that sample might be spurious associations attributable to independent non-zero developmental trajectories, and that exploration of such trajectories would arbitrate this hypothesis. Evaluation of correlations in a model that explicitly includes effects of age reveals that only a correlation between nT and nE persist, supporting the view that the measures of the CAST provide relatively independent information about the operation of attention.

Chapter 5

Conclusion

This work described the theoretical motivation for and exploration of a new experimental task for the study of attention. The results from two experiments, spanning children and young adults, provide evidence that this task can be effectively deployed by researchers to simultaneously study a variety of phenomena of attention, including both endogenous and exogenous modes of temporal, spatial and executive attention.

In addition to providing reliable and stable measurement of expected phenomena in both young adults and children, the unique methodology employed provides the opportunity to observe a number of novel phenomena not yet reported in the attention literature nor addressed by theory thereof. In adults, we observe that previously reported interactions between temporal and spatial attention manifest here between their endogenous modes only. Similarly, we observe that previously reported interactions between temporal and executive attention manifest such that endogenous temporal attention interacts with both endogenous and exogenous executive attention, but exogenous temporal attention interacts with neither form of executive attention. When it comes to interactions between spatial and executive attention, both exogenous and endogenous spatial attention interact with endogenous executive attention, but neither mode of spatial attention interacts with exogenous executive attention.

By extending exploration of the CAST across a sample that includes both children and adults, this work replicates previously reported developmental effects on endogenous spatial attention and exogenous executive attention, but also shows that previously reported developmental effects on endogenous executive attention may be attributable to a mere speed-accuracy criterion shift with age. This work also makes the (to our knowledge) novel observation of a developmental effect on the interaction between endogenous and exogenous modes of executive attention.

Finally, while not presented in detail above, the project from which the data from

E2 were re-purposed also highlights the utility of the CAST for providing a more nuanced view of attention and effects thereon. In that study of the influence of exposure to nature-vs-urban environments, Johnson, Snow, Lawrence, and Rainham (submitted) deployed the CAST to evaluate the prediction from Attention Restoration Theory (ART; Kaplan, 1995) that exposure to nature would selectively improve endogenous, but not exogenous, modes of attention. Indeed, both endogenous temporal attention and endogenous spatial attention showed benefits from exposure to nature, and no such benefits were observed for their exogenous counterparts. Research seeking to isolate the influence of variables on attention with greater specificity than provided by previous tests of attention might similarly benefit from use of the CAST.

5.1 Other ANT-based tests of attention

While the introduction situates the CAST as an evolution from the ANT and ANT-I as methodological ancestors, it is worth noting similar work building from these seminal tests.

After Rueda et al. (2004), who provided the first modification to the ANT whereby arrow target stimuli were replaced by child-friendly cartoon fish, Roberts, Summerfield, and Hall (2006) constructed a version of the ANT that presented all stimuli in the auditory modality. Although they did not observe a significant auditory orienting effect, auditory scores for both Alerting and Executive were of similar magnitude to, and correlated with, their visual counterparts as measured by the standard ANT. These results support a supramodal mechanism of attention for at least the Alerting and Executive systems, and suggest that in cases where visual targets are not possible (for example, among blind participants), auditory analogs may serve in their stead. In such cases, it should be possible to similarly implement an auditory measure of xE by lateralized presentation of the auditory stimuli.

Greene et al. (2008) modified the ANT by rotating the orientation of all stimuli such that cues and targets appeared to the left and right of fixation rather than above and below. By also presenting targets very briefly (170ms, too briefly for participants to move their eyes to the target) Greene et al. argued that the resulting performance on left- vs right-located targets could provide insight into the operation of attention in each cortical hemisphere separately. With this manipulation, they found that

scores for both hemispheres were largely similar and correlated with their standard ANT counterparts. Green et al. also innovated a modification of the ANT whereby spatial cues were only 75% accurate in their prediction of the subsequent target's location, thereby providing both valid and invalid trials that, by comparison to the standard double-cue condition, provide measures of Orienting benefits and Orienting costs, respectively. Similar metrics of the costs and benefits of both nS and xS may be possible in the CAST by addition of appropriate neutral cue conditions in both subtests against which to compare its existing valid and invalid cue conditions.

Fuentes and Campoy (2008) modified the ANT-I to explore multiple cue-target intervals, permitting the exploration of how these intervals affected the exogenous spatial attention and its interaction with Alerting. With this design they indeed observed an interaction between xS , Alerting and cue-target interval such that xS remained relatively constant across intervals on trials with no Alerting stimulus, on trials with an Alerting stimulus there was a larger xS effect on trials with shorter cue-target intervals. While the CAST as described here employs intervals designed to achieve maximum system scores, for those more interested in measuring a similar timecourse of effects and interactions it should be relatively trivial to modify the CAST to achieve this. Specifically, in the X subtest, the interval between the simultaneous xT/xS stimulus and the target could be varied across multiple blocks of trials (mixed trials would induce differential temporal uncertainty across intervals that would then confound their comparison). Similarly, in the N subtest the interval between the nT stimulus and the target could be varied across multiple blocks of trials. Finally, while not done here for brevity, it would be possible to explore the current data for the influence of the specific nS -to- nT interval experienced on each trial in the N subtest, where these intervals were randomly varied by design.

Combining the innovation by Greene et al. (2008) of the inclusion of invalid trials and the innovation of Fuentes and Campoy (2008) of the inclusion of multiple cue-target intervals, Fan et al. (2009) created a revision of the ANT called the ANT-R that additionally lateralized the cue and target stimuli. Unlike Green et al. (2008), who presented arrow targets pointing either upwards or downwards, Fan et al. presented arrow targets pointing left or right, thus permitting measurement of xE as achieved in the CAST. While providing a novel (if obvious, as demonstrated by

parallel implementation by the CAST) methodological contribution to the evolution of ANT-like tasks, the empirical contribution of Fan et al.'s deployment of the ANT-R is hampered by a combination of a small sample size (30 participants), haphazard statistical procedures, and the addition of multiple new network scores that are poorly explained/reasoned. It is clear, however, that the ANT-R is able to measure all the networks measured by the original ANT, with the addition of xE as a new measure. As observed in the CAST, Fan et al. also observed a strong $xE:nE$ interaction such that a reversed xE effect occurs on trials with incongruent flankers.

Roca et al. (2011) developed the ANTI-V, an extension of the ANT-I that added a measure of Vigilance as measured by a secondary simple-detection task whereby on some trials (25%) the central target stimulus in an array of flankers was displaced slightly, in which case participants were instructed to ignore the standard direction response of standard ANT-I trials and simply detect the displaced target by pressing a central response key. Later, Luna et al. (2018) presented the ANTI-Vea, which included an easier simple-detection task that made it more appropriate for use among impaired, developing, or aging populations. The ANTI-Vea also included additional Vigilance-related measures of fatigue indexed by changes in the mean and variance of response times on the vigilance task across the course of the testing session. It should be relatively trivial to similarly add a target-displacement-detection vigilance task to the CAST, though it would be interesting to evaluate the degree to which the measures thereby obtained dissociate from those that might be obtained from the CAST as-is; that is, it should be possible to obtain a measure of vigilance from the CAST in the change in overall RT and the variability of RT across the testing session.

A last noteworthy entry in the ANT/ANTI-I lineage is Klein et al. (2017)'s AttentionTrip, which extends the ANT-I to a more game-like task (using a steering wheel to fly a space ship through a worm-hole, shooting targets that appear) and including measures of Simon conflict. Klein et al. (2017) also explored both purely exogenous orienting cues as employed by the ANT-I, as well as cues with both exogenous and endogenous properties, as employed by the ANT. While they obtained reliable scores for both Alerting and Executive networks across all variants, only in the variant with cues that had both exogenous and endogenous properties did they achieve reasonably sized orienting scores, suggesting that the visual complexity of the more engaging

game environment may be eliminating contributions from exogenous spatial attention. These results demonstrate a risk to the otherwise understandable impulse to game-ify these tasks; while increased game-like tasks may increase engagement and thereby yield psychometric benefits, task designers need to be wary that the changes made do not inadvertently interfere with the constructs to be measured.

5.2 Limitations and future directions

Chapter 1 elaborates the theoretical justification for use of the CAST as a more rigorous tool for the exploration of attention than provided by similar tests in wide use, and Chapters 2–4 bolster this proscription with evidence of the CAST’s reliability and power to discern nuanced phenomena of attention. While it may be rhetorically convenient to end discussion here, in the spirit of encouraging further improvements in either the CAST or its successors, we feel it important to note several areas of limitation and uncertainty remaining in the design of the CAST and its measurement of the phenomena of attention.

One design choice made in the development of the CAST was to separate the manipulation of endogenous and exogenous modes of both spatial and temporal attention into separate tasks. This choice was forced by the difficulty of any alternative strategy; certainly it was not feasible to explore the full interaction of exogenous and endogenous modes of temporal attention as achieved by the involved experimental design of Lawrence and Klein (2013), and pilot testing of a design exploring the full interaction of exogenous and endogenous modes of spatial attention revealed that the endogenous mode completely dominated, eliminating the ability to measure the exogenous mode.¹ Similarly, it was difficult to conceive of a method of exploring alternative cross-domain interactions; for example, exploration of the $nT:xS$ interaction requires a short interval between the spatial-cue and target for the manipulation of exogenous spatial attention, providing insufficient time to properly manipulate endogenous temporal attention. It may be possible to explore the $xT:nS$ interaction using a design similar to the N task used here but exchanging the presence/absence

¹Note that Müller and Rabbitt (1989) reported of seemingly successful efforts in the exploration of the interactions between the modes of spatial attention. However, where substantial methodological differences exist between their work and that presented here, further exploration of the parameter space is required to evaluate whether it is possible to add similar measurement to the CAST.

of the auditory signal with a manipulation of the signal's volume as in the X task. Indeed, it may even be possible to avoid the addition of an entirely separate additional block for exploration of this interaction and instead modify the existing N task to include not only absent-signal and low-volume signal trials but also high-volume signal trials as well. This addition would increase the task duration by 50%, but it would be worth at least exploring the consequences of this modification in a healthy young adult sample.

Another design decision made in the development of the CAST was to employ arrows as endogenous spatial cues, motivated by a desire to provide an easy stimulus-contingency mapping for participants to remember. However, it may be argued that precisely because this mapping is easy to remember, arrows might not provide a strictly pure measure of endogenous spatial attention. That is, because arrows have well-learned extra-experimental spatial associations, they may engage somewhat different mechanisms of spatial attention than would stimuli without such associations. Indeed, it has been demonstrated that the behavioural (Ristic, Wright & Kingstone, 2007) and electroencephalographic time-course (Brignani, Guzzon, Marzi & Miniussi, 2008) of arrow-induced cuing differs from that of stimuli with novel spatial associations, with arrow-induced cuing manifesting a time-course similar (though not identical) to that of exogenous spatial cues. Furthermore, while the spatial association manifest by arrows might be expected to be well-learned by adults, there is surely a time in development when they are not well-learned and we thus might expect to observe a developmental trajectory (like the one observed in this report) that is associated not with a changing capacity for endogenous spatial attention, but instead associated with increased exposure to arrows as extra-experimental spatial stimuli. Indeed, it is possible to argue that the developmental effects observed across the age range of this report reflect, at least in part, increased exposure to arrows as spatial stimuli: the nS effect not only grows with age, rising to nearly the magnitude of the xS effect within the adult sample, but also changes from having an ambiguous (i.e. not credibly zero, but also not credibly non-zero) interaction with nE in the child sample to a credibly non-zero interaction with nE in the adult sample.² Where xS interacts with nE in both samples, this development of an $nS:nE$ interaction

²Note, however, that the age: $nS:nE$ interaction itself is ambiguous when assessed formally in the developmental analysis.

arguably signals a transition of the arrow from engaging purely endogenous modes of spatial attention to also engaging exogenous-like modes as well. In addition to these developmental complications, it would also be problematic to deploy the existing arrow stimuli to compare the function of endogenous spatial attention across cultures that differ in their use of the arrow shape as a spatial symbol. Given the relatively trivial motivation for the use of arrows in the first place and the ease with which they could be replaced by spatial cuing stimuli without such extra-experimental spatial associations, we advise future research to explore this alternative design choice. This change may induce greater cognitive demands on participants to interpret the cue stimuli and remember their instructed spatial association, but the use of such stimuli in a test like the CAST should at least be explored.

A similar limitation manifests in the use of spatial Stroop as a measure of endogenous executive attention. That is, the spatial Stroop effect is distinguished from more standard Simon effects in its manifestation in experimental designs that use a spatial property of the target stimulus to identify the required response by the participant. But the degree to which a fish looking left or right could be considered “spatial” (versus a simple difference in shape) is likely affected by the participant’s history of extra-experimental experience with similar stimuli having consequent spatial associations (ex. fish pointing in a direction and subsequently swimming that way). Thus, the above developmental and cultural complications apply. As noted in the discussion of the developmental analysis, one feature of the current data is that it is arguably able to arbitrate whether differential experience with arrows across the age ranges sampled affected the manifestation of the spatial Stroop effect as measured by the interaction with the flanker effect. Contrary to expectation, the developmental analysis observed a decrease in this interaction, suggesting that concern for developmental effects might reasonably be limited to younger samples than that observed here. It may also be worthwhile for future research to explore the use of alternative stimuli without extra-experimental spatial association (ex. “Square vs diamond”) to avoid these issues entirely, though such stimuli involve more difficult-to-remember stimulus-response mappings, making the task more difficult for participants with not-yet-mature or diminished cognitive systems.

Another limitation of the CAST lies in the measurement of endogenous temporal

attention. As illustrated by Lawrence and Klein (2013), there exists a procedure by which this can be achieved optimally, wherein comparison is made in performance amidst blocks of trials where cues are informative with regards to the time of subsequent target presentation and blocks of time where cues and targets are presented randomly and independently. However, this procedure requires considerable time for data collection and in designing the CAST we felt it would be better to employ a “next-best” method whereby the timing of individual trials sought to minimize the temporal informativeness of stimuli other than the warning signal, thence comparing warning-present against warning absent trials. We advise future research to investigate the consequence of this choice, starting with a replication of Lawrence and Klein with the addition of blocks with the non-aging foreperiod and both warning present and warning absent trials, to ascertain the degree to which this measure of endogenous temporal attention differs from the pure measure proposed by Lawrence and Klein.

A limitation, not of the CAST, but of the approach to analysis employed above is that while the modelling improves on common practice by simultaneous modelling of both response times and errors, allowing each to inform the other, there is room for further improvement. Specifically, it has become increasingly common for data from speeded choice experiments to employ what might be called a “process model” whereby response times and errors are simultaneously modelled as outputs of a latent model of information processing. The most popular of such models include the drift-diffusion model (Ratcliff, 1978; Ratcliff & Rouder, 1998) and the linear ballistic accumulator (Brown & Heathcote, 2008), both of which are well-validated as effective descriptions of the information processing involved in speeded choice experiments (Donkin, Brown, Heathcote, & Wagenmakers, 2011). The chief benefit of these models, in addition to contributing to the general scientific aim of not simply describing phenomena but also quantitatively expressing an underlying causal structure, is that they are able to combine the two facets of performance, speed and accuracy, to achieve direct inference on information processing efficiency independent of participants’ speed-accuracy criterion. The task of the researcher in interpreting the consequent results is thereby greatly aided, reducing the number of outcome variables from two to one and permitting inference on information processing efficiency

amidst changes in speed-accuracy trade-offs. This latter feature would be particularly useful with the current data where the effect of age on nE appears to involve a speed-accuracy trade-off that prevents further inference using only the tools employed here. Additionally, it is increasingly common to combine such process models with an explicit model of data contamination, providing more robust inference amidst outliers (ex. very fast or very slow RTs) than can be achieved by more traditional outlier trimming procedures (such as the absolute 200ms lower-bound employed above). While, primarily due to time constraints, we employed a more traditional “descriptive” approach to modelling, we advise that contaminated process models be explored.

Another limitation to the analytic approach employed above is the use of a strictly linear model for effects of age on the CAST measures. While linear models are faster/simpler to construct and compute, they will fail to capture more nuanced developmental trajectories that may be of theoretical and applied interest. Tools exist for the flexible analysis of possibly-non-linear effects of a continuous variable like age, including Generalized Additive models (Hastie & Tibshirani, 1986) and Gaussian Process models (Williams & Rasmussen, 1996), and while for the sake of time this report did not explore application of these models, we advise future exploration of this data using these tools.

5.3 Final remarks

While there is substantial work ahead for the ongoing development of effective tools for the exploration of attention, the CAST provides a substantial update to existing and popular tools. This work demonstrates that the CAST provides a more nuanced picture of attention than possible with prior tools, and does so through reliable and stable measurement that promises great utility for the elaboration of both typical and disordered states of attention.

Bibliography

- Baayen, H.R., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3(2).
- Baddeley, A. D., & Hitch, G. J. (1974) Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 8). New York: Academic Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 10.1016/j.jml.2012.11.001. doi:10.1016/j.jml.2012.11.001
- Baumeister, A., & Joubert, C. (1969). Interactive effects on reaction time of preparatory interval length and preparatory interval frequency. *Journal of Experimental Psychology*, 82, 393–395. doi:10.1037/h0028119
- Bertelson, P. (1967). The time course of preparation. *Quarterly Journal of Experimental Psychology*, 19(3), 272-279. doi:10.1080/14640746708400102
- Bertelson, P., & Tisseyre, F. (1969). The time-course of preparation: Confirmatory results with visual and auditory warning signals. *Acta Psychologica*, 30, 145-154. doi:10.1016/0001-6918(69)90047-X
- Boehnke, S. E., Hall, S. E., & Marquardt, T. (2002). Detection of static and dynamic changes in interaural correlation. *The Journal of the Acoustical Society of America*, 112(4), 1617-1626. doi:10.1121/1.1504857
- Brignani, D., Guzzon, D., Marzi, C., & Miniussi, C. (2008). Attentional orienting induced by arrows and eye-gaze compared with an endogenous cue. *Neuropsychologia*. 47. 370-81. doi:10.1016/j.neuropsychologia.2008.09.011.
- Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57(3), 153-178.
- Callejas, A., Lupiáñez, J., & Tudela, P. (2004). The three attentional networks: On their independence and interactions. *Brain and Cognition*, 54, 225-227. doi:10.1016/j.bandc.2004.02.012

- Callejas, A., Lupiáñez, J., Funes, M. J., & Tudela, P. (2005). Modulations among the alerting, orienting and executive control networks. *Experimental Brain Research*, 167(1), 27-37. doi:10.1007/s00221-005-2365-z
- Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76. doi:10.18637/jss.v076.i01
- Christie, J., Chun, A., Wylie, G., & Klein, R. M. (under revision) Spatial Correlation is Sufficient to Generate One Form and Nearly Extinguish Another Form of Automatic Attentional Orienting. *Psychonomic Bulletin & Review*.
- Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *Journal of Neuroscience*, 18(18), 7426-7435. doi:10.1523/JNEUROSCI.18-18-07426.1998
- Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language*, 59(4), 447-456. doi:10.1016/j.jml.2007.11.004
- Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E. J. (2011). Diffusion versus linear ballistic accumulation: different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review*, 18(1), 61-69.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143-149. doi:10.3758/BF03203267
- Fan, J., Gu, X., Guise, K. G., Liu, X., Fossella, J., Wang, H., & Posner, M. I. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain and Cognition*, 70(2), 209-220. doi:10.1016/j.bandc.2009.02.002
- Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, 14(3), 340-347. doi:10.1162/089892902317361886
- Fuentes, G., Campoy, G.(2008). The time course of alerting over orienting in the attention network test. *Experimental Brain Research*. 185: 667–672.
- Goldberg, M. C., Maurer, D., & Lewis, T. L. (2001). Developmental changes in attention: The effects of endogenous cueing and of distractors. *Developmental Science*, 4(2), 209-219.

- Greene, D. J., Barnea, A., Herzberg, K., Rassis, A., Neta, M., Raz, A., & Zaidel, E. (2008). Measuring attention in the hemispheres: The lateralized attention network test (LANT). *Brain and Cognition*, 66(1), 21-31. DOI: 10.1016/j.bandc.2007.05.003
- Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science*, 1(3), 297-310.
- Hyndman, R. (1996). Computing and Graphing Highest Density Regions. *The American Statistician*, 50(2), 120-126. doi:10.2307/2684423
- Jaeger, F. (2008) Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 434-446.
- James, W. (1890). *The Principles of Psychology*. New York: Henry Holt & Company.
- Johnson, S.A., Snow, S.L., Lawrence, M.A. & Rainham, D.G. (submitted). Quasi-Randomized Trial of Contact with Nature and Effects on Attention in Children. Submitted for review at PLOS One.
- Jonides, J., Irwin, D. E., & Yantis, S. (1982). Integrating visual information from successive fixations. *Science*, 215(4529), 192-194. doi:10.1126/science.7053571
- Karlin, S., & McGregor, J. (1959). Coincidence probabilities. *Pacific Journal of Mathematics*, 9(4), 1141-1164. doi:10.2140/pjm.1959.9.1141
- Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology*, 44(1), 69-104. doi:10.1080/14640749208401284
- Klein, R. M., & Lawrence, M. A. (2011). On the modes and domains of attention. In M.I. Posner (Ed.) "Cognitive Neuroscience of Attention", 11-28. New York: Guilford Press.
- Klemmer, E. T. (1956). Time uncertainty in simple reaction time. *Journal of experimental psychology*, 51(3), 179. doi:10.1037/h0042317
- Lawrence, M. A., & Klein, R. M. (2013). Isolating exogenous and endogenous modes of temporal attention. *Journal of Experimental Psychology: General*, 142(2), 560. doi:10.1037/a0029023
- Lu, C. H., & Proctor, R. W. (1995). The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. *Psychonomic Bulletin & Review*, 2(2), 174-207. doi:10.3758/BF03210959

- Luna FG, Marino J, Roca J, Lupiáñez J (2018). Executive and arousal vigilance decrement in the context of the attentional networks: The ANTI-Vea task. *Journal of Neuroscience Methods*; 306:77-87. doi: 10.1016/j.jneumeth.2018.05.011.
- Luce, R. (1986). *Response times*. New York: Oxford University Press.
- MacLeod, J. W., Lawrence, M. A., McConnell, M. M., Eskes, G. A., Klein, R. M., & Shore, D. I. (2010). Appraising the ANT: Psychometric and theoretical considerations of the Attention Network Test. *Neuropsychology*, 24(5), 637. doi:10.1037/a0019803
- Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception & Psychophysics*, 49(3), 270-288.
- Mullane, J. C., Lawrence, M. A., Corkum, P. V., Klein, R. M., & McLaughlin, E. N. (2016). The development of and interaction among alerting, orienting, and executive attention in children. *Child Neuropsychology*, 22(2), 155-176.
- Müller, H. J., & Rabbitt, P. M. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, 15(2), 315-330. doi:10.1037/0096-1523.15.2.315
- Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, 89(1), 133. doi:10.1037/0033-2909.89.1.133
- Norman, D., Shallice, T. (1980). Attention to Action: Willed and Automatic Control of Behaviour. In M. Lansman, E. Hunt (Eds.) *Proceedings of the Lake Wilderness Attention Conference: Interim Technical Report*.
- Pillsbury, W. B. (1908). *Attention*. London: L S. Sonnenschein & Company.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32(1), 3-25. doi:10.1080/00335558008248231
- Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, 78(5), 391. doi:10.1037/h0031333
- Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, 13(1), 25-42. doi:10.1037/h0042317
- R Core Team (2017). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–532.
- Ratcliff, R., Rouder, J.N. (1998) Modeling response times for two-choice decisions. *Psychological Science*, 9, 347–356.
- Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental psychology: Human Perception and Performance*, 18(3), 849. doi:10.1037//0096-1523.18.3.849
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, 72(2), 77. doi:10.1037/h0027760
- Ribot, T. (1898). *The psychology of attention*. Open Court Pub. Co.
- Ristic, J., Wright, A. & Kingstone, A. (2007) Attentional control and reflexive orienting to gaze and arrow cues. *Psychonomic Bulletin & Review*, 14: 964. doi:10.3758/BF03194129
- Roberts KL, Summerfield AQ, Hall DA. (2006) Presentation modality influences behavioral measures of alerting, orienting, and executive control. *Journal of the International Neuropsychological Society*; 12(4): 485-92.
- Roca, J, Castro, C , Fernanda López-Ramón, M, Lupiáñez, J. (2011). Measuring vigilance while assessing the functioning of the three attentional networks: The ANTI-Vigilance task. *Journal of neuroscience methods*. 198. 312-24. 10.1016/j.jneumeth.2011.04.014.
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, 42(8), 1029-1040. doi:10.1016/j.neuropsychologia.2003.12.012
- Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, 81(1), 174. doi:10.1037/h0027448
- Stan Development Team. 2017. *Stan Modeling Language Users Guide and Reference Manual, Version 2.17.3*. <http://mc-stan.org>
- Titchener, Edward Bradford. (1908). *Lectures on the elementary psychology of feeling and attention*. New York, NY, US: MacMillan Co. doi:10.1037/10867-000

- Williams, B.R., Strauss, E.H., Hultsch D.F., & Hunter, M.A. (2007) Reaction Time Inconsistency in a Spatial Stroop Task: Age-Related Differences Through Childhood and Adulthood. *Aging, Neuropsychology, and Cognition*, 14:4, 417-439, DOI: 10.1080/13825580600584590
- Williams, C.K., & Rasmussen, C.E. (1996). Gaussian processes for regression. In *Advances in Neural Information Processing Systems* (pp. 514-520).
- Wundt, W. M. (1874). *Grundzüge de physiologischen Psychologie* (Vol. 1). W. Engelman.

Appendix A

Task instructions

A.1 Initial task instructions

[Instructions to the experimenter are presented inside square braces]

In this experiment you'll be using this gamepad to respond to things on the screen.

You only have to use these two trigger buttons here. [Indicate trigger buttons]

Whenever you see a fish on the screen, press the button corresponding to the direction that the fish is pointing. So if a fish comes up on the screen pointing left like you see here, I'd press the left button like this [press the left button].

After you press the button, you'll see a number appear briefly at the center of the screen. This number shows how long it took you to press the button, in milliseconds. You want this number to be as small as possible, usually between 300 and 700, so try to press the buttons as fast as you can.

[press "1" key] Now, sometimes you may press the wrong button by mistake. For example, if a right pointing fish appeared like you see here I might mistakenly press the left button [press the left button]. Don't worry too much if this happens, it turns out that we learn almost as much from your mistakes as we do from your accurate answers, so going fast is what really matters. However, if you do make a mistake, don't bother trying to correct your response by pressing the other button. You can only press the button once per fish, and if you press any other buttons [press "1" key] a message like this will appear telling you to respond only once.

You only have about one second to press a button once the fish appears, and if you don't respond in time [press "1" key] a message like this will appear telling you that you missed the fish. However, try to wait until the fish actually appears before pressing the button, because if you press a button before the fish appears [press "1" key] a message like this will appear telling you that you pressed the button too soon. [press "1" key] So the fish will appear pointing either left or [press "1" key] right and they can appear on the left side of the screen as you've been seeing so far, but

they can also appear on the right side of the screen, again pointing either left [press “1” key] or right [press “1” key]. Where each fish appears and what direction it’s pointing are both completely random, so there won’t be any pattern to where the fish will appear, and there also won’t be any pattern to the direction the fish is pointing. Remember your job is to indicate what direction the fish is pointing no matter of where it appears.

So far I’ve been showing you screens with only one fish, but during the experiment there will sometimes be a whole school of fish like this [press “1” key]. When you see a school of fish, your job is to press the button corresponding to the direction of the center fish and ignore the buddy fish on either side. Sometimes the buddy fish will be pointing in the same direction as the center fish, as you see here, and sometimes [press “1” key] the buddy fish will be pointing in the opposite direction as the center fish.

A.2 Subtest instructions for “N-first” group

A.2.1 N subtest instructions

During this part of the experiment, before the fish appears you’ll see an arrow that is there to help you guess where the fish are going to appear. Most of the time, the fish will appear where the arrow is pointing. Sometimes the arrow will get it wrong and the fish won’t appear where the arrow is pointing, but the arrow is correct most of the time, so it makes sense to always try to pay attention to the place where the arrow is pointing.

[press “1” key] So the arrow tells you where the fish is likely to appear, but you still don’t know when the fish is going to appear. To help you know when the fish are going to appear, the computer will play a special sound through these headphones that you’ll be wearing. The headphones will actually play a constant fuzz sound, and it’s this fuzz sound that will change to let you know that the fish will appear soon. To give you an example of what this sounds like, when I tell you to put on your headphones the computer will play the fuzz for two seconds, then the fuzz will change for just a short period, then it will change back. One second after changing, a fish will appear. [Have participant put on headphones and press “1” key. Repeat (by

pressing the “2” key then the “1” key again) until they report hearing the change] Now, this sound won’t always happen. To let you know if the sound is going to happen, a shape will appear around the arrow. [press “1” key] If you see an arrow with a circle around it like this, that means that the sound will happen. [press “1” key] If you see an arrow with a square around it, that means that the sound won’t happen.

[press “1” key then “q” to quit the demo]

So the first few minutes of this part of the experiment is practice where you can try to get the feel for how it works.

Do you have any further questions before you start practice? [Answer any questions]

When you’re ready to start practice, take the gamepad and press one of the buttons to begin. [Make sure they are pressing the correct trigger buttons. Watch practice to make sure they’re performing reasonably fast & accurately.]

Practice is over and it looks like you have the hang of it. Do you have any further questions before the experiment starts? [Answer any questions]

A.2.2 X subtest instructions

During this part of the experiment, before the fish appears you’ll see a dot flicker on the left like this [press “1” key TWICE quickly] or on the left like this [press “1” key TWICE quickly]. The location of this flicker is completely random and doesn’t have anything to do with where the fish is going to appear.

Just like in the previous part of the experiment, during this part of the experiment you will be wearing headphones through which you’ll hear a fuzz sound, and this fuzz sound will change right before the fish appears. Sometimes the change will be quiet, like this [have participant put on headphones and press “1” key], but other times the change will be loud like this [have participant put on headphones and press “1” key]. [press “1” key then “q” to quit the demo]

So the first few minutes of this part of the experiment is practice where you can try to get the feel for how it works.

Do you have any further questions before you start practice? [Answer any questions]

When you’re ready to start practice, take the gamepad and press one of the buttons to begin. [Make sure they are pressing the correct trigger buttons. Watch practice to

make sure they're performing reasonably fast & accurately.]

Practice is over and it looks like you have the hang of it. Do you have any further questions before the experiment starts? [Answer any questions]

A.3 Subtest Instructions for “X-first” group

A.3.1 X subtest instructions

During this part of the experiment, before the fish appears you'll see a dot flicker on the left like this [press “1” key TWICE quickly] or on the left like this [press “1” key TWICE quickly]. The location of this flicker is completely random and doesn't have anything to do with where the fish is going to appear.

Before the fish appears, the computer will also play a special sound through these headphones that you'll be wearing. The headphones will actually play a constant fuzz sound, and it's this fuzz sound that will change before the fish appears. To give you an example of what this sounds like, when I tell you to put on your headphones the computer will play the fuzz for two seconds, then the fuzz will change for just a short period, then it will change back. Almost immediately after the fuzz changes back, a fish will appear. [Have participant put on headphones and press “1” key. Repeat (by pressing the “2” key then the “1” key again) until they report hearing the change]

Now sometimes the sound change will be quiet like you just heard, but other times the sound change will be loud like this [have participant put on headphones and press “1” key].

[press “1” key then “q” to quit the demo]

So the first few minutes of this part of the experiment is practice where you can try to get the feel for how it works.

Do you have any further questions before you start practice? [Answer any questions]

When you're ready to start practice, take the gamepad and press one of the buttons to begin. [Make sure they are pressing the correct trigger buttons. Watch practice to make sure they're performing reasonably fast & accurately.]

Practice is over and it looks like you have the hang of it. Do you have any further questions before the experiment starts? [Answer any questions]

A.3.2 N subtest instructions

During this part of the experiment, before the fish appears you'll see an arrow that is there to help you guess where the fish are going to appear. Most of the time, the fish will appear where the arrow is pointing. Sometimes the arrow will get it wrong and the fish won't appear where the arrow is pointing, but the arrow is correct most of the time, so it makes sense to always try to pay attention to the place where the arrow is pointing.

[press "1" key] So the arrow tells you where the fish is likely to appear, but you still don't know when the fish is going to appear. To help you know when the fish are going to appear, you'll hear the same fuzz change that you heard in the first part of the experiment, but this time it will always be quiet, and the fish will always appear one whole second after the sound changes, like this [Have participant put on headphones and press "1" key. Repeat (by pressing the "2" key then the "1" key again) until they report hearing the change]

Now, this sound won't always happen. To let you know if the sound is going to happen, a shape will appear around the arrow. [press "1" key] If you see an arrow with a circle around it like this, that means that the sound will happen. [press "1" key] If you see an arrow with a square around it, that means that the sound won't happen.

[press "1" key then "q" to quit the demo]

So the first few minutes of this part of the experiment is practice where you can try to get the feel for how it works.

Do you have any further questions before you start practice? [Answer any questions]

When you're ready to start practice, take the gamepad and press one of the buttons to begin. [Make sure they are pressing the correct trigger buttons. Watch practice to make sure they're performing reasonably fast & accurately.]

Practice is over and it looks like you have the hang of it. Do you have any further questions before the experiment starts? [Answer any questions]

Appendix B

Figures of fixation period distributions

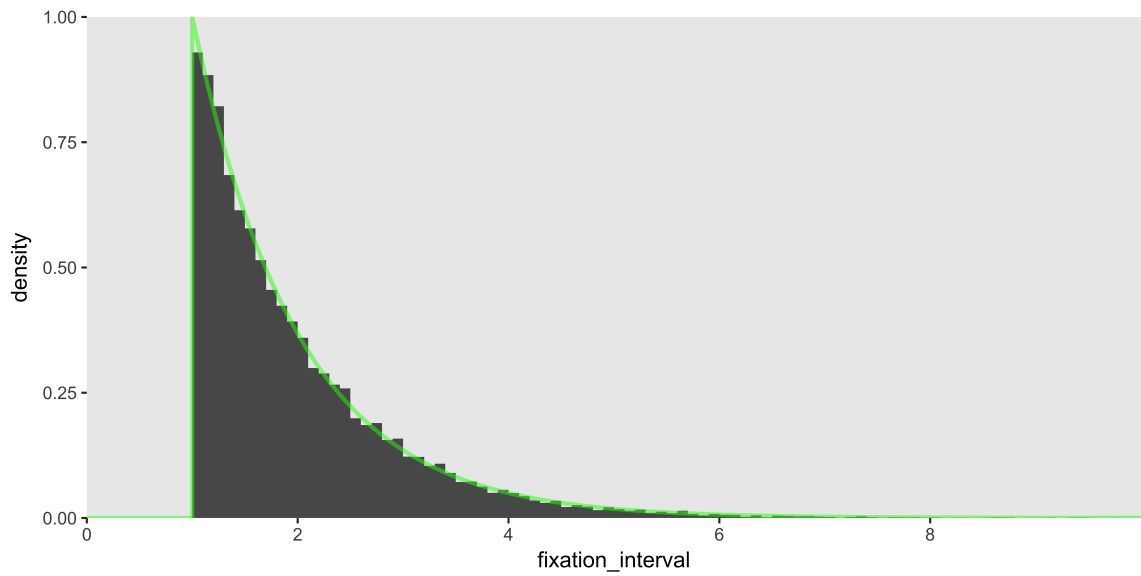


Figure B.1: Histogram of all fixation periods from all participants in E1, using bin-widths of 50ms. Green curve conveys theoretical density function.

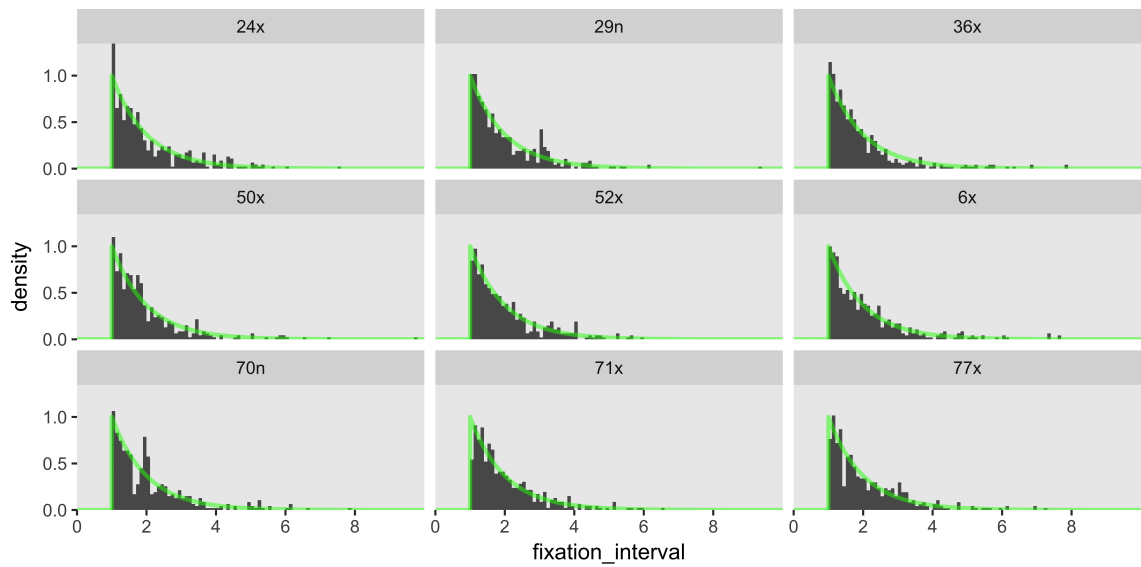


Figure B.2: Histogram of all fixation periods from each of a random selection of participants in E1, using bin widths of 100ms. Green curve conveys theoretical density function.

Appendix C

E1 Plots

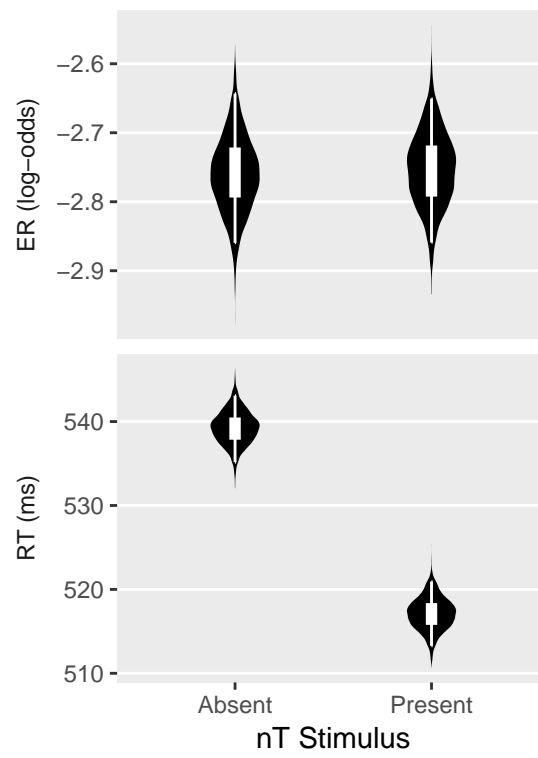


Figure C.1: Main effect of the nT manipulation on RT and ER

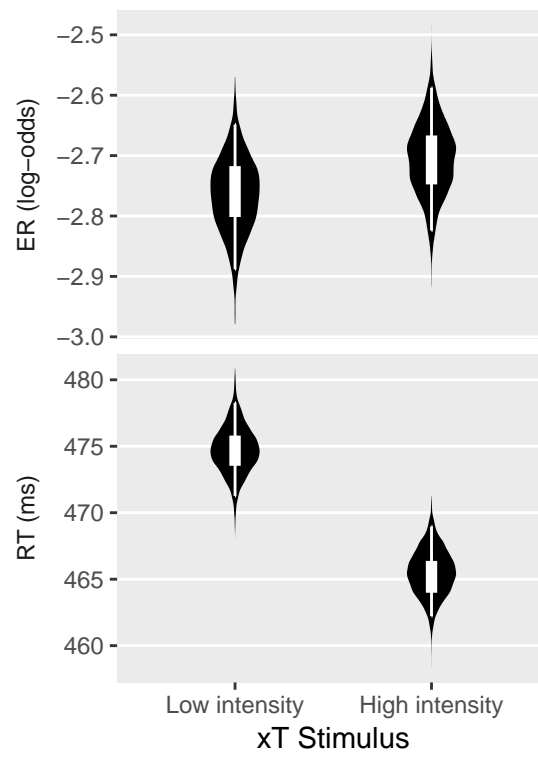


Figure C.2: Main effect of the xT manipulation on RT and ER

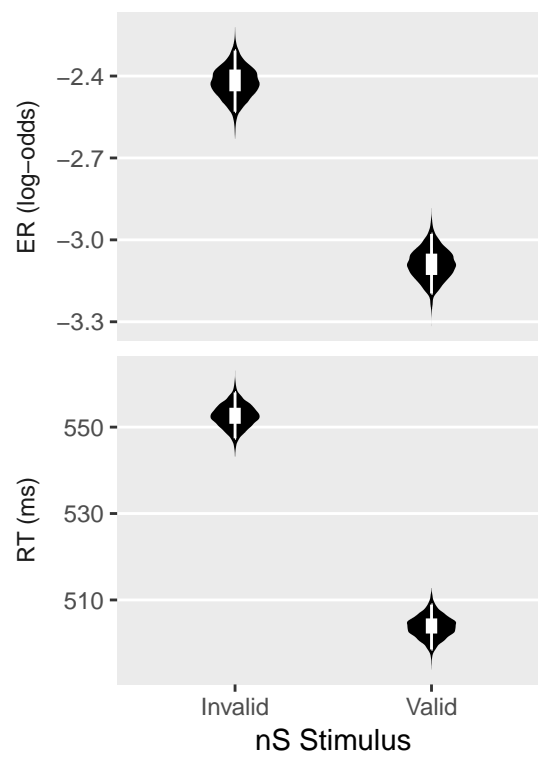


Figure C.3: Main effect of the nS manipulation on RT and ER

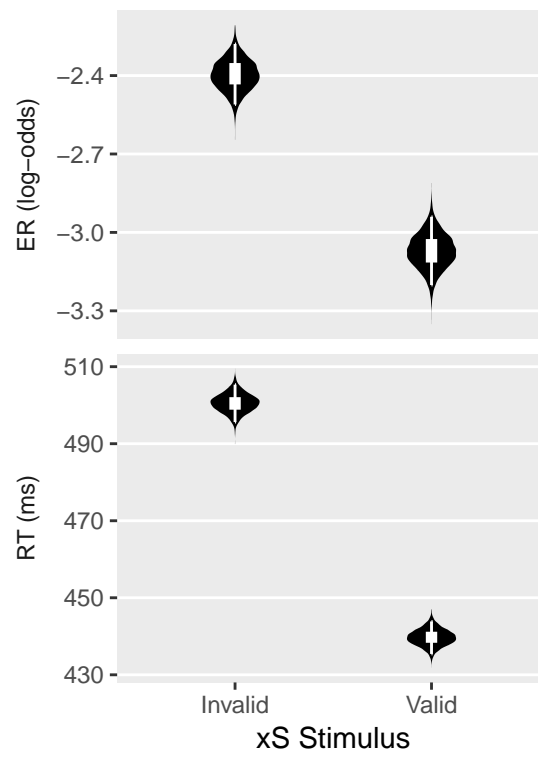


Figure C.4: Main effect of the xS manipulation on RT and ER

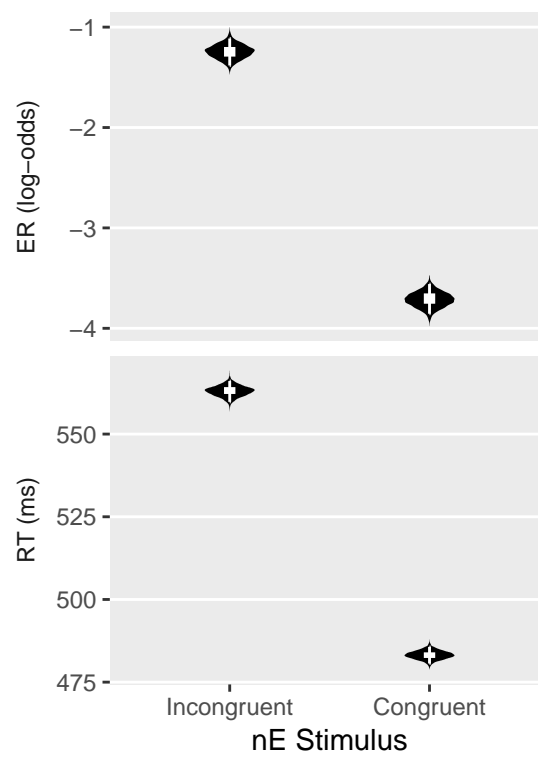


Figure C.5: Main effect of the nE manipulation on RT and ER

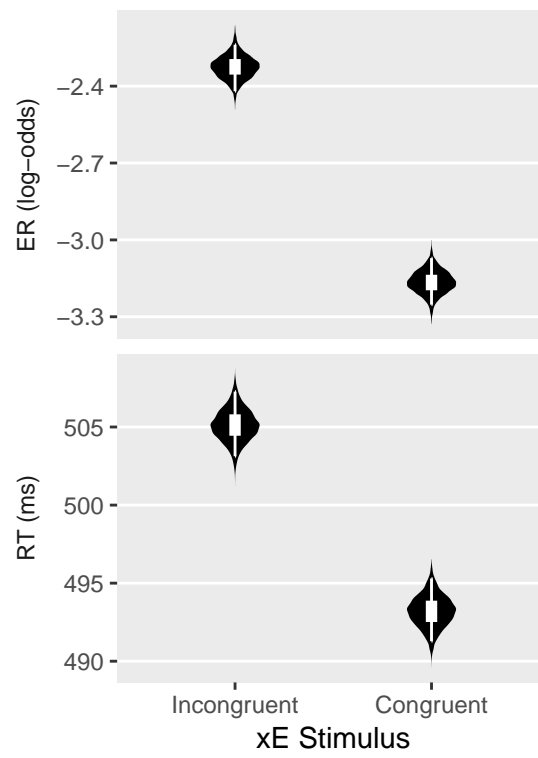


Figure C.6: Main effect of the xE manipulation on RT and ER

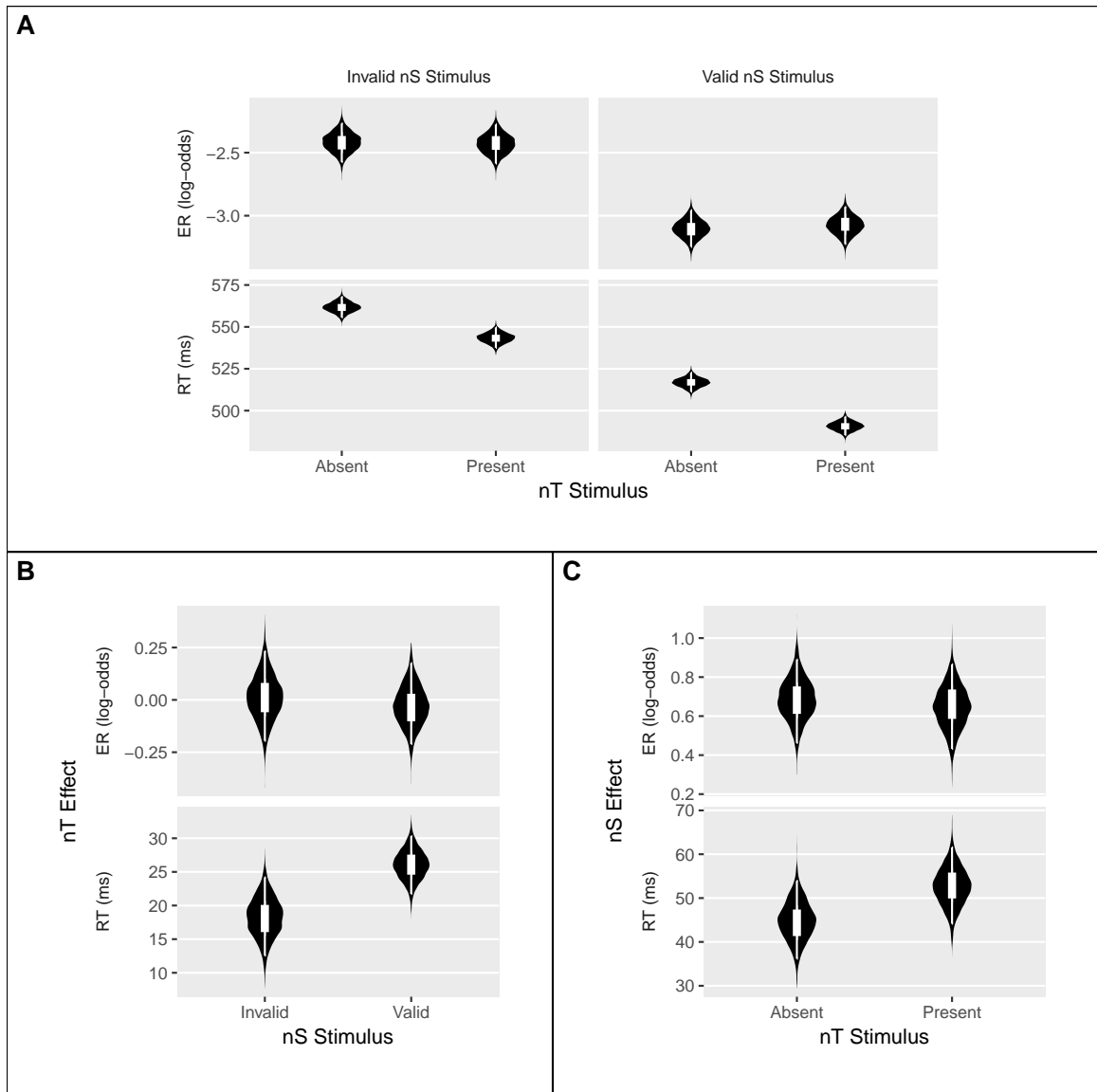


Figure C.7: Interaction of effects of nT and nS manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

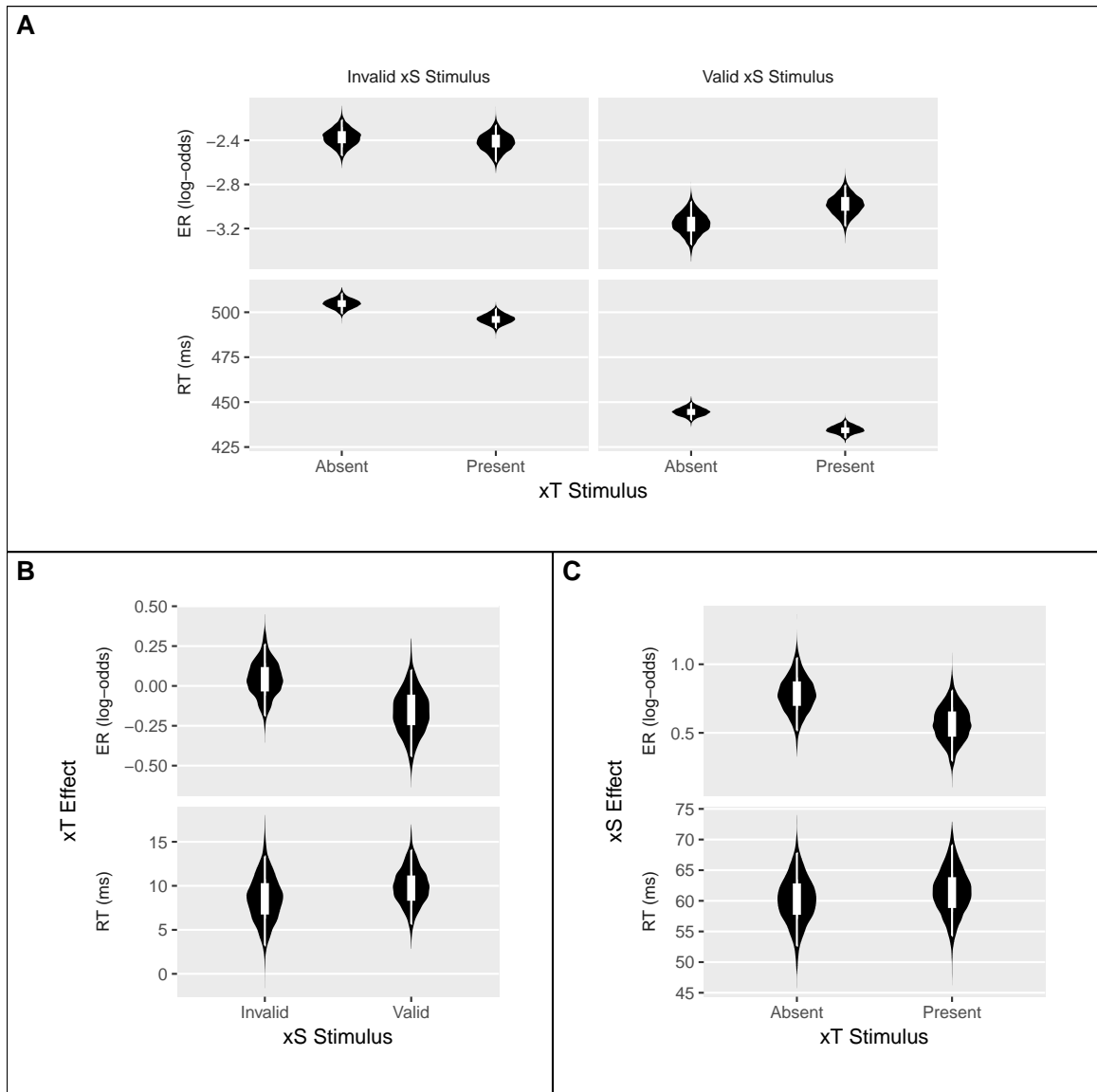


Figure C.8: Interaction of effects of xT and xS manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

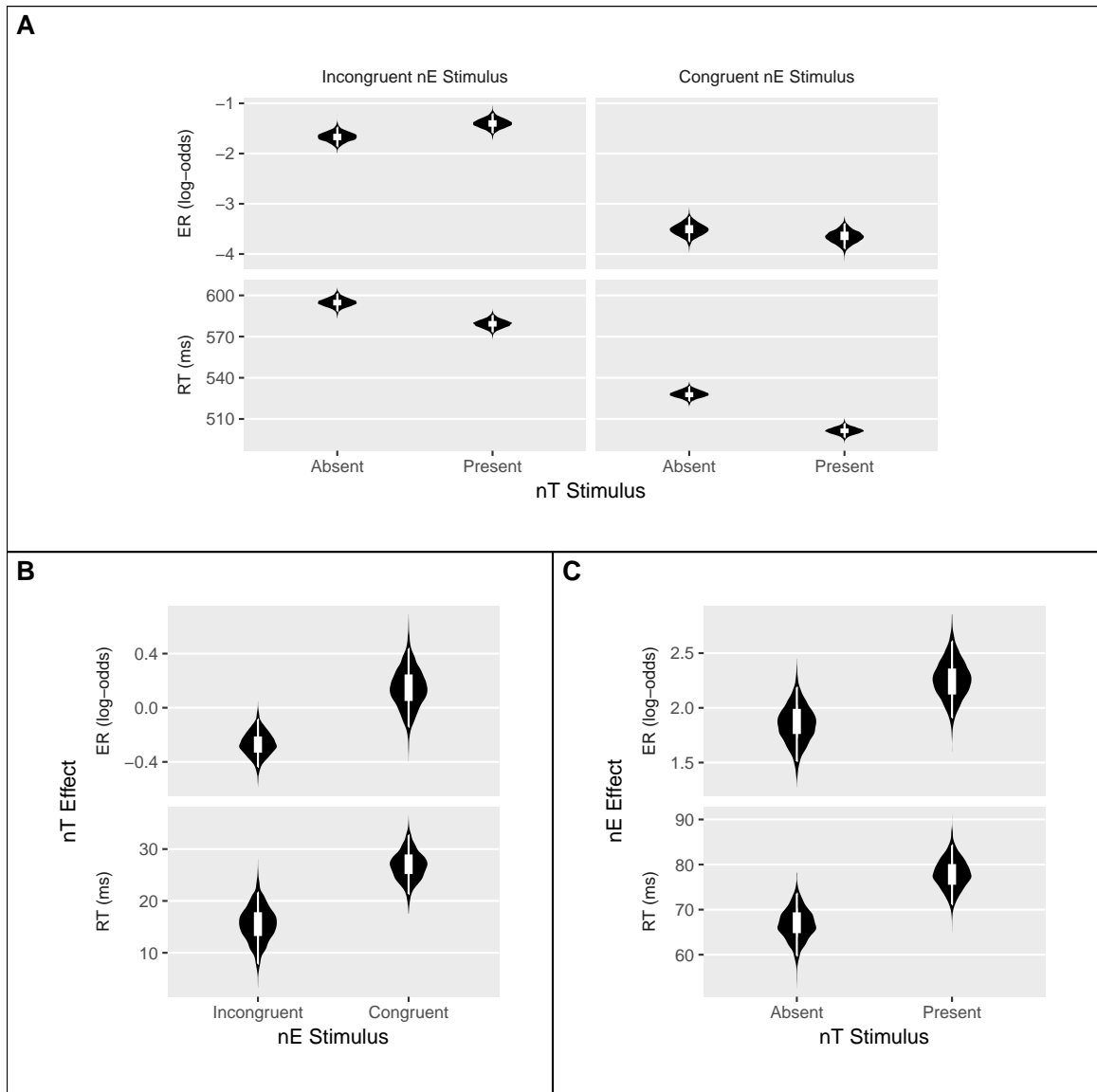


Figure C.9: Interaction of effects of nT and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

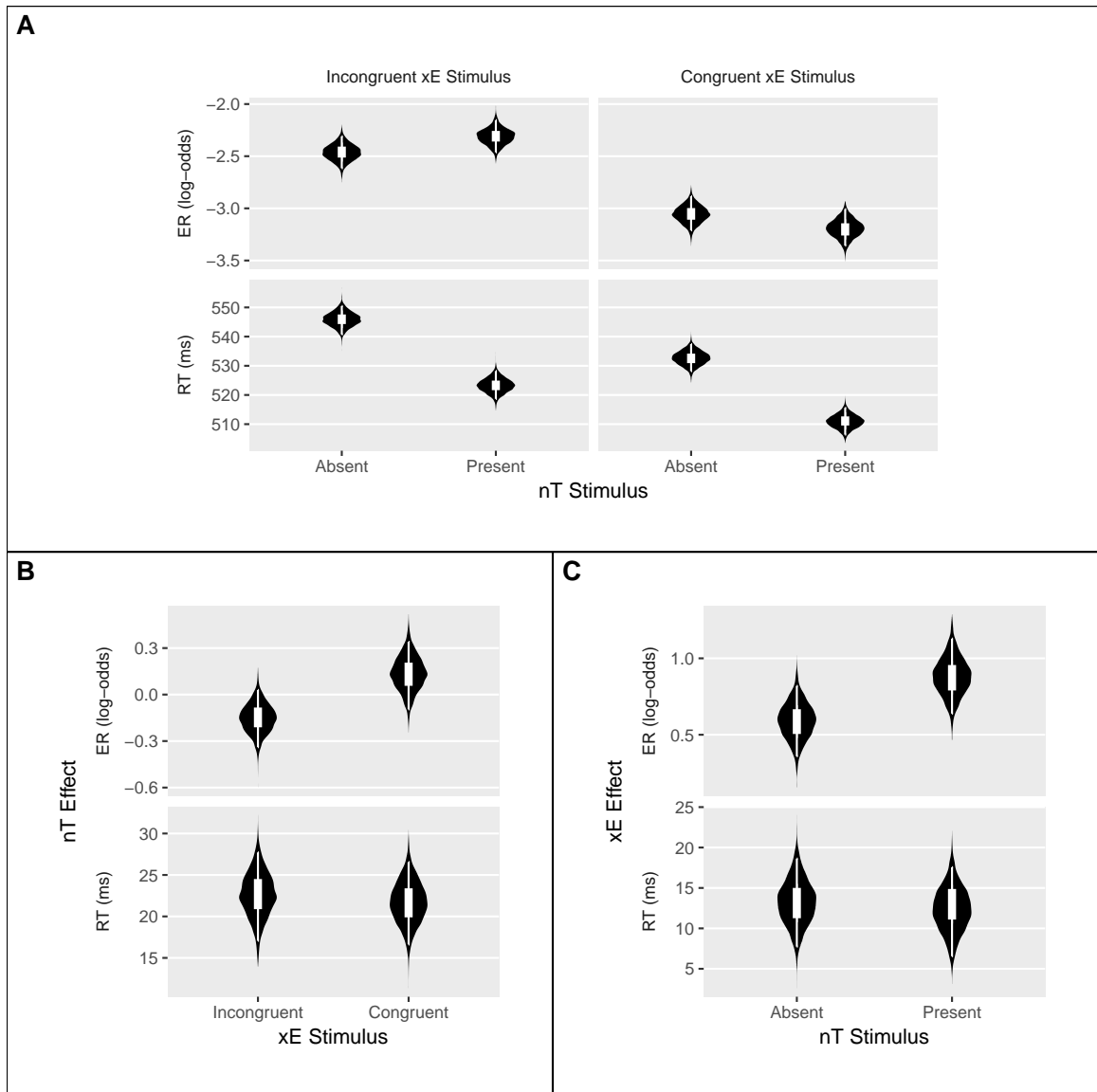


Figure C.10: Interaction of effects of nT and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

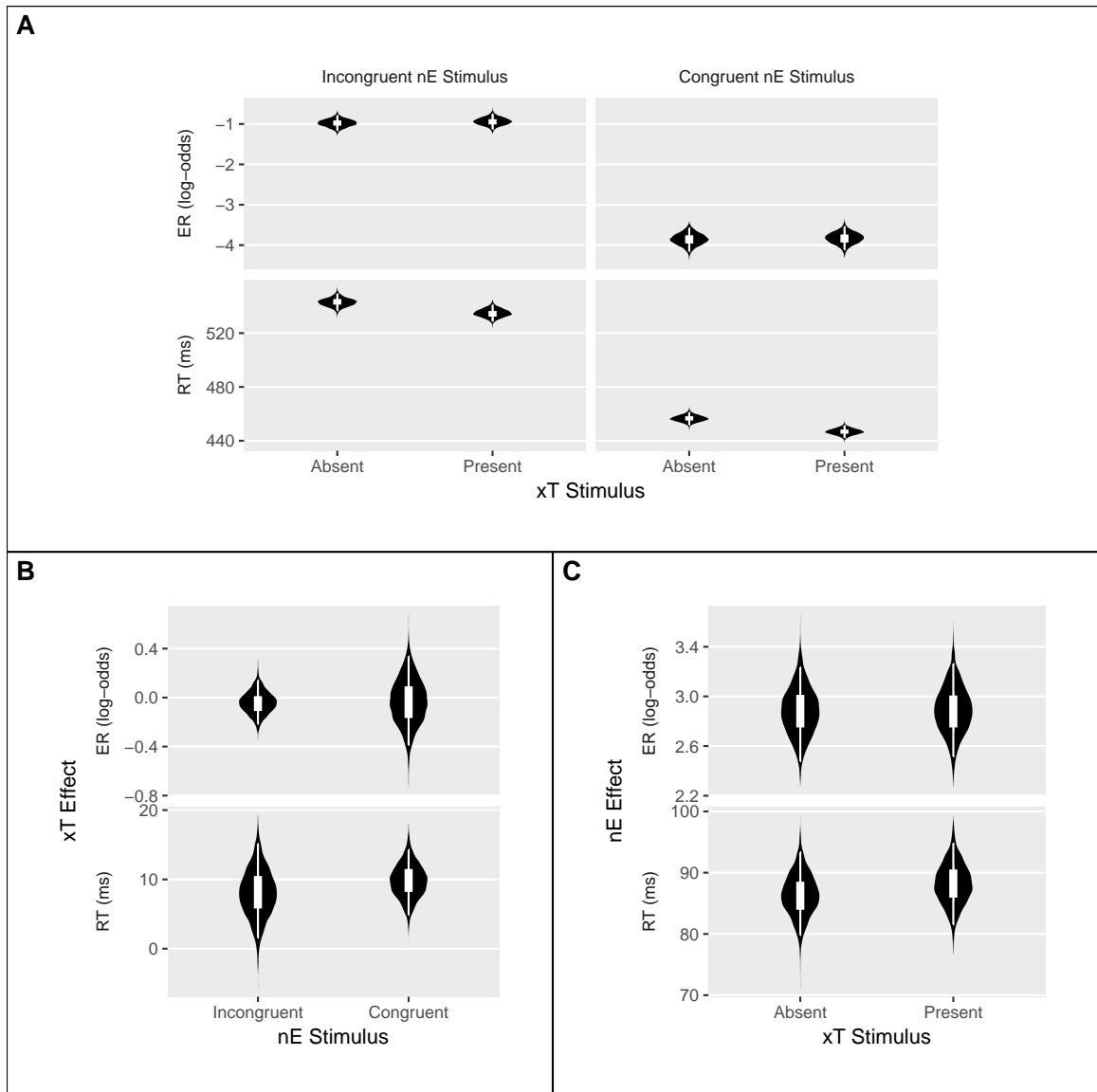


Figure C.11: Interaction of effects of xT and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

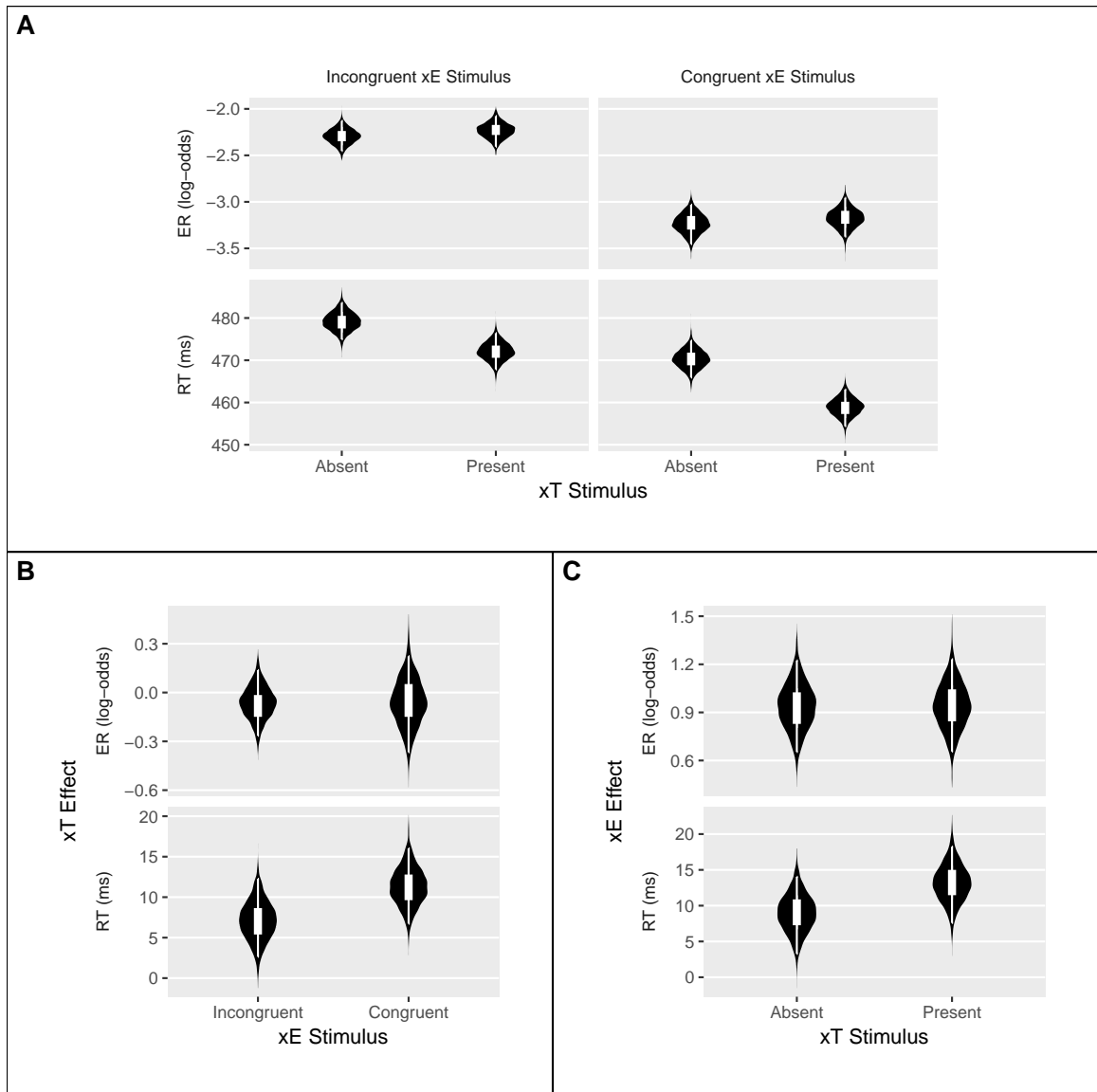


Figure C.12: Interaction of effects of xT and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

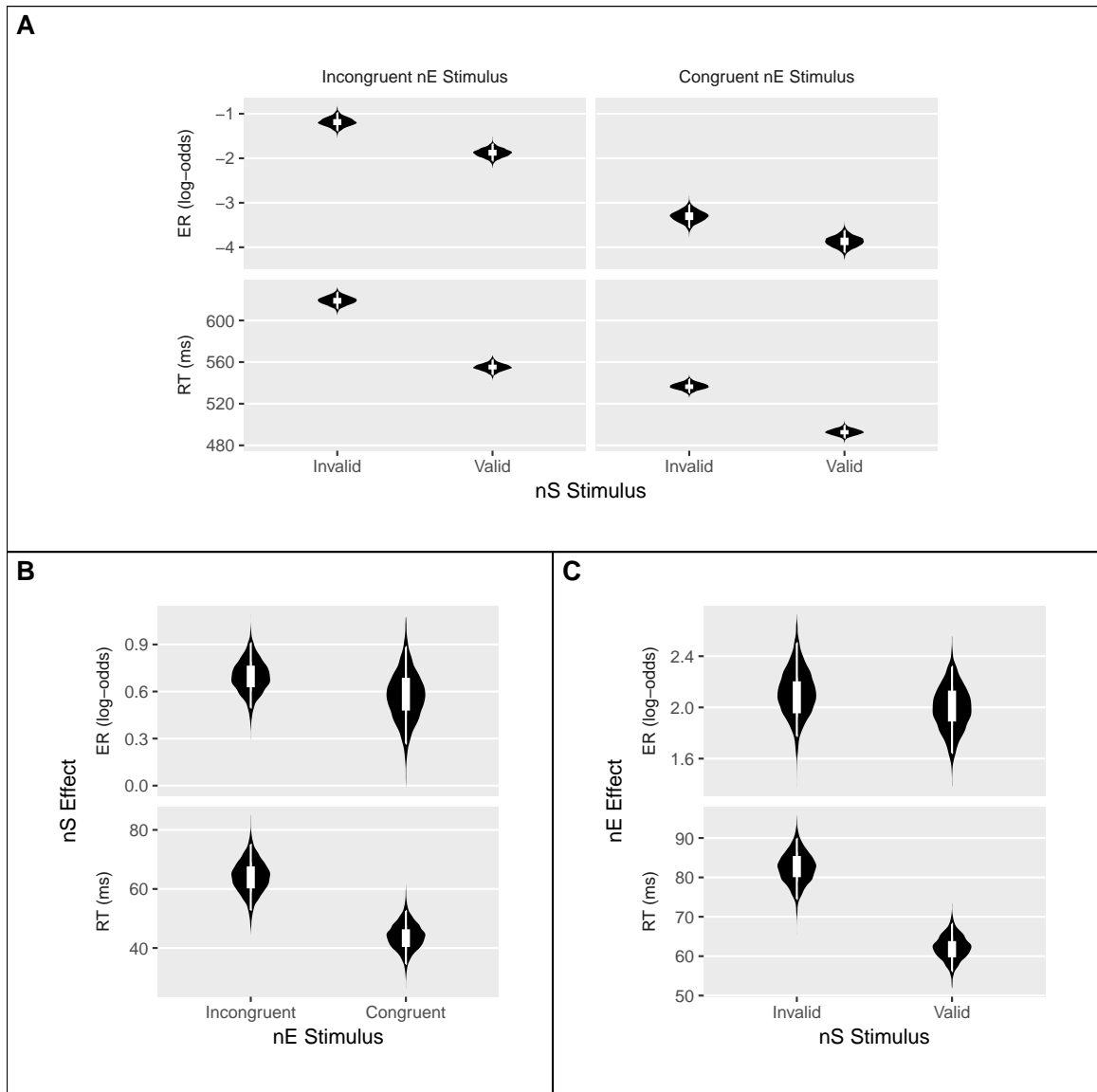


Figure C.13: Interaction of effects of nS and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

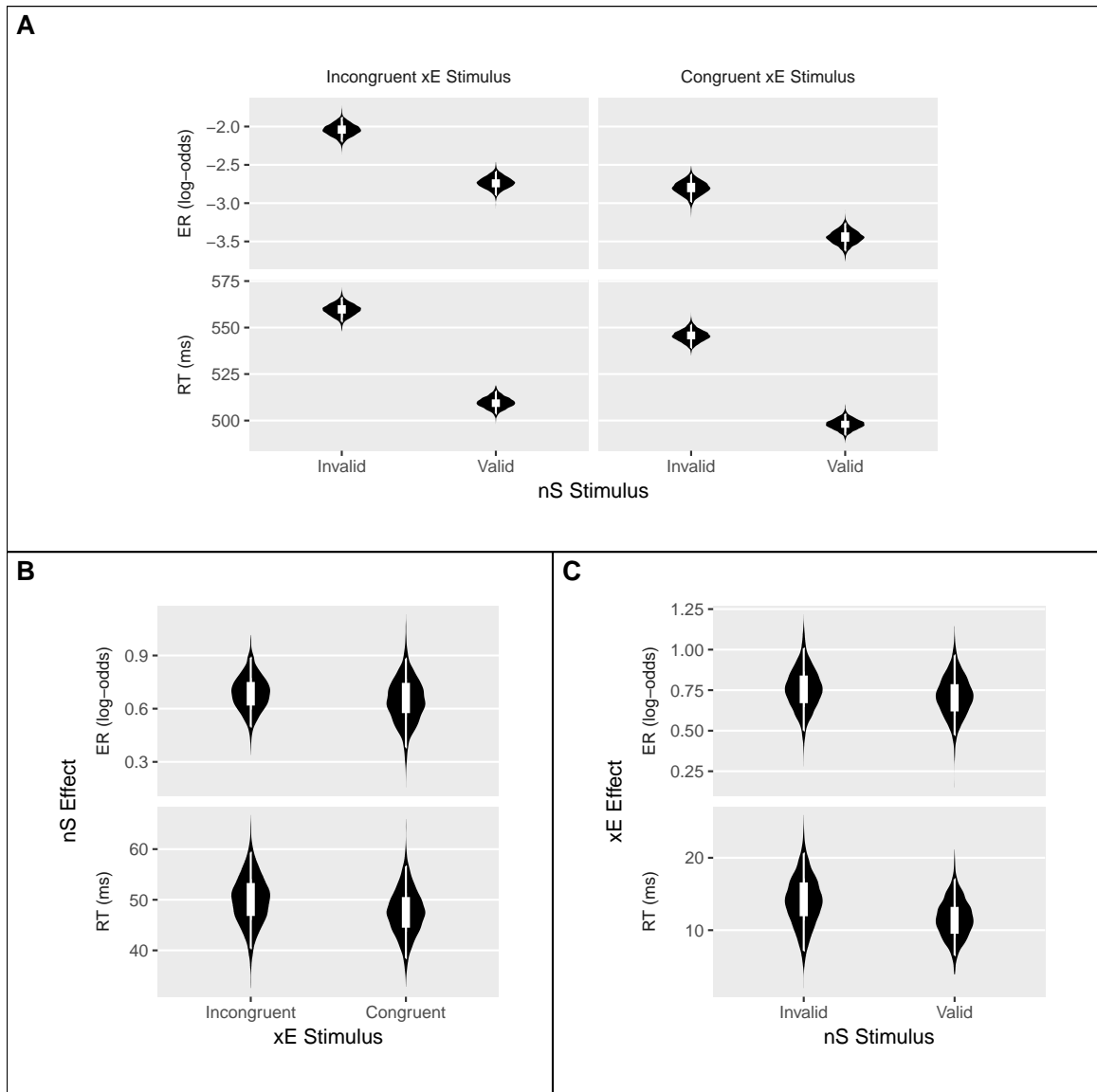


Figure C.14: Interaction of effects of nS and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

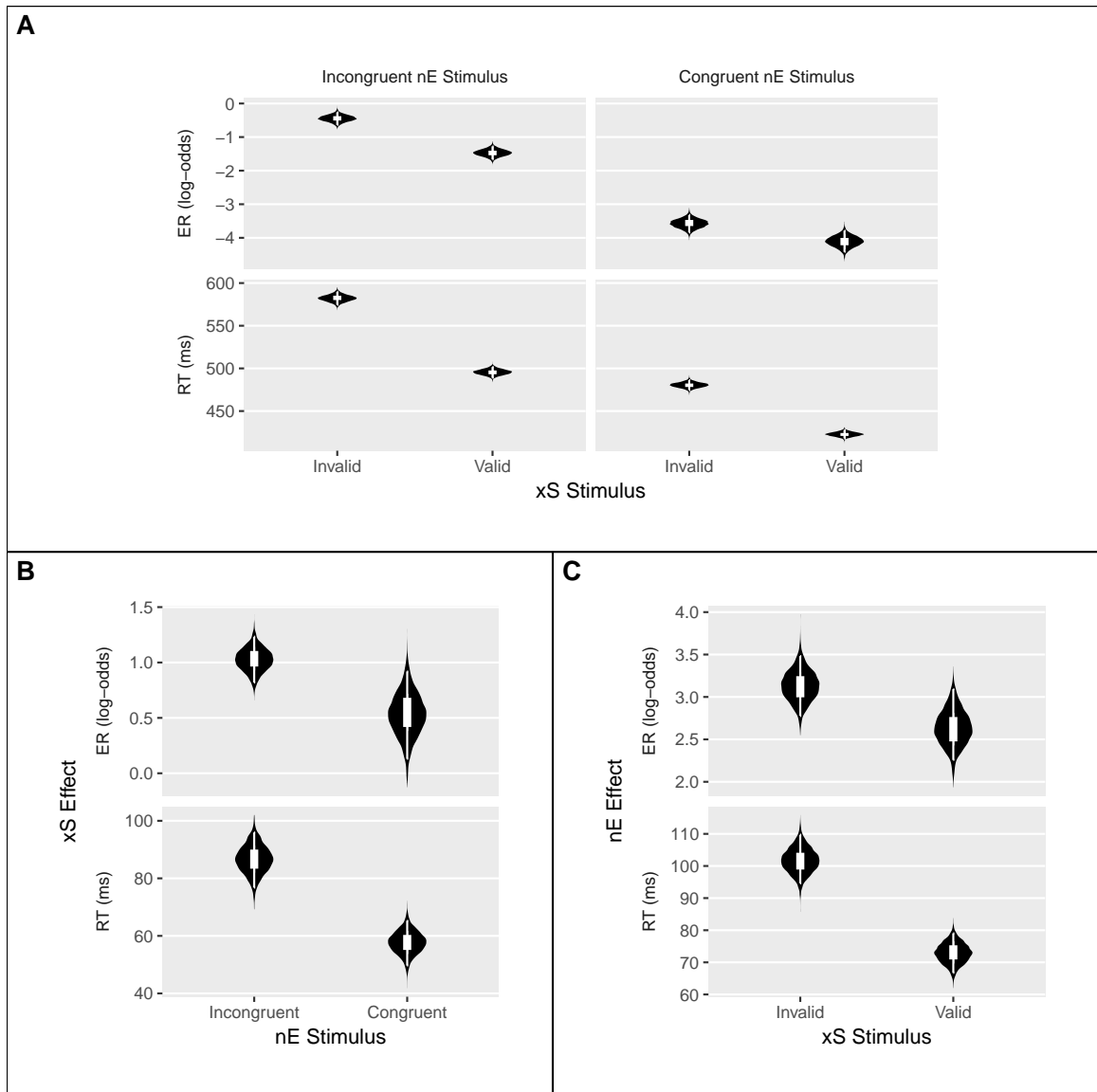


Figure C.15: Interaction of effects of xS and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

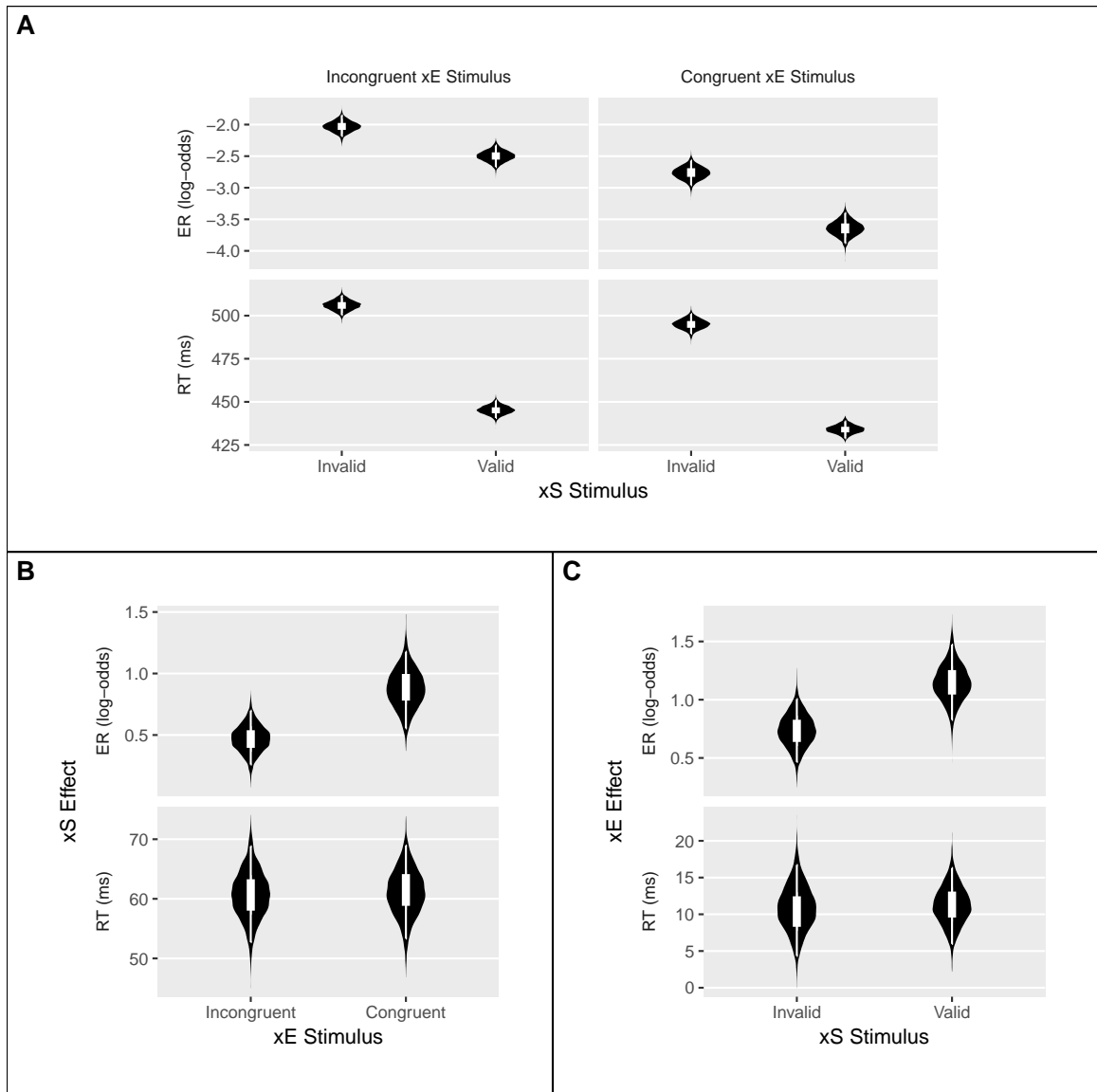


Figure C.16: Interaction of effects of xS and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

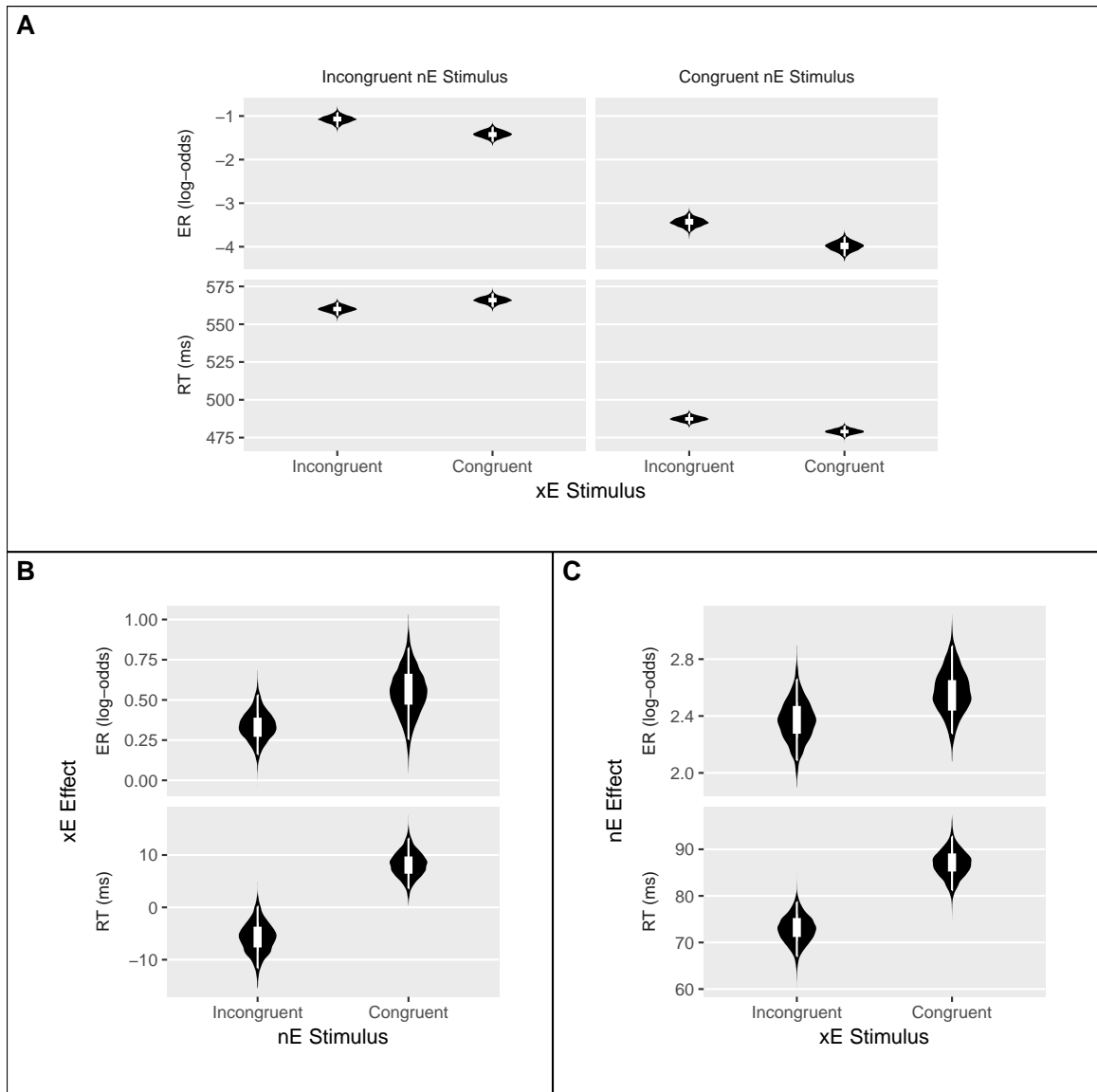


Figure C.17: Interaction of effects of xE and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

Appendix D

E2 Plots

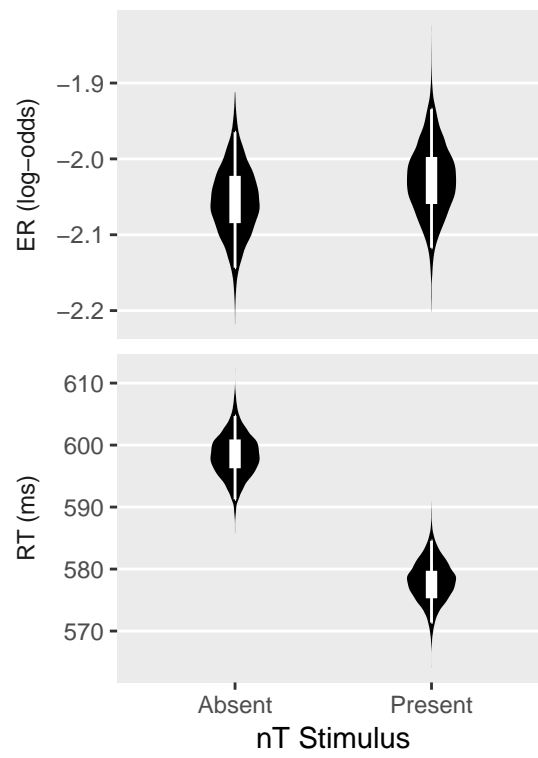


Figure D.1: Main effect of the nT manipulation on RT and ER

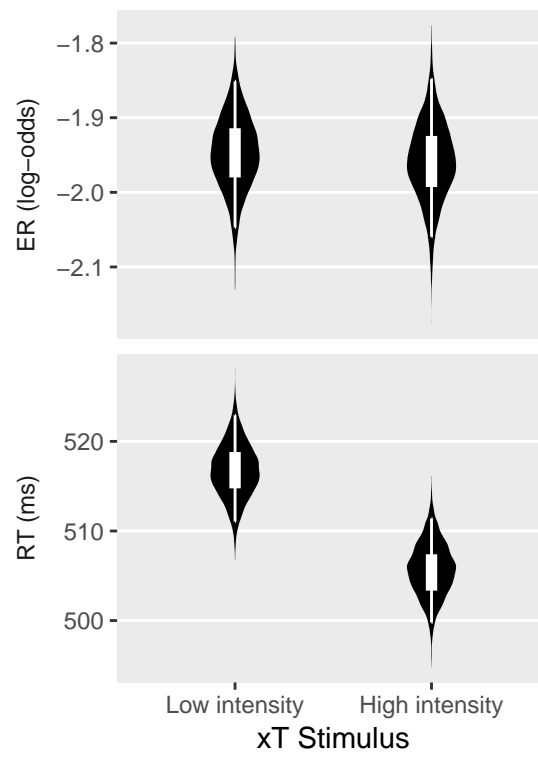


Figure D.2: Main effect of the xT manipulation on RT and ER

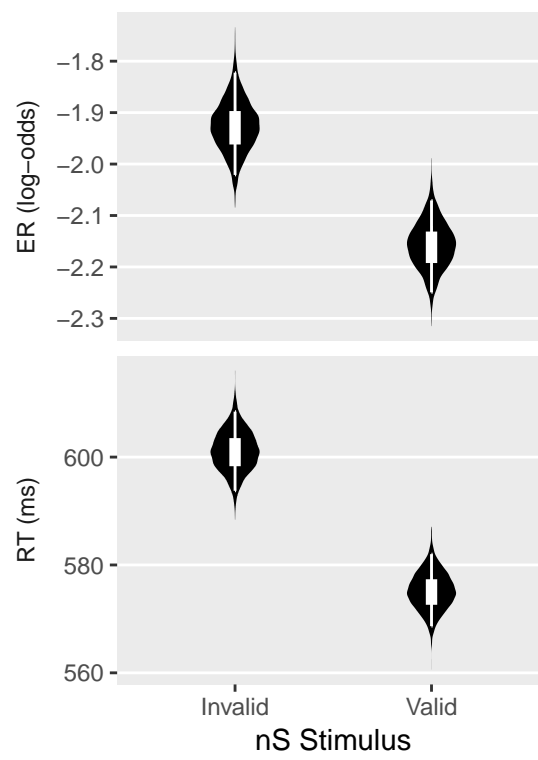


Figure D.3: Main effect of the nS manipulation on RT and ER

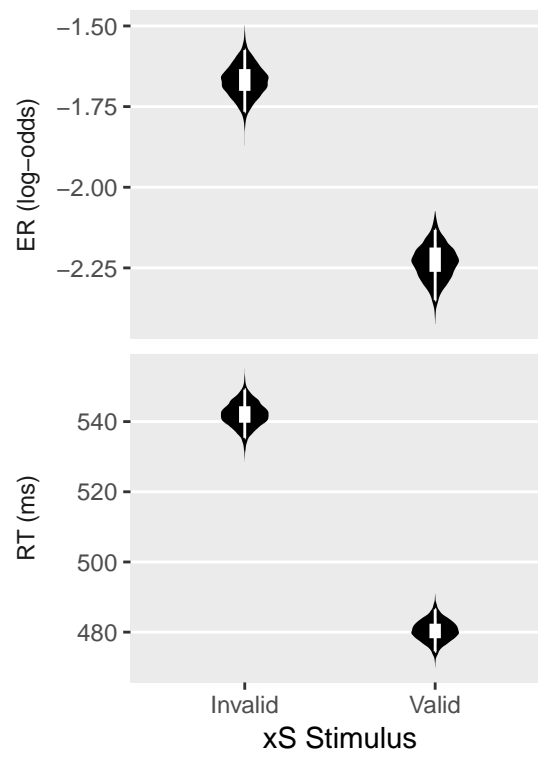


Figure D.4: Main effect of the xS manipulation on RT and ER

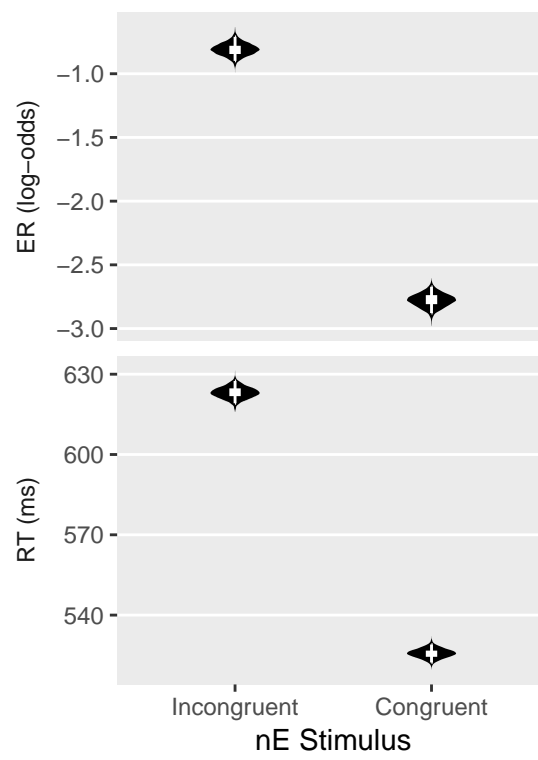


Figure D.5: Main effect of the nE manipulation on RT and ER

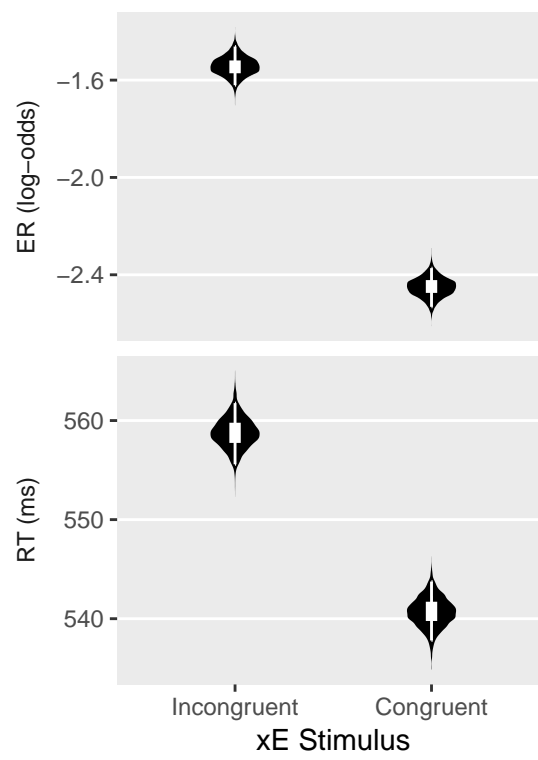


Figure D.6: Main effect of the xE manipulation on RT and ER

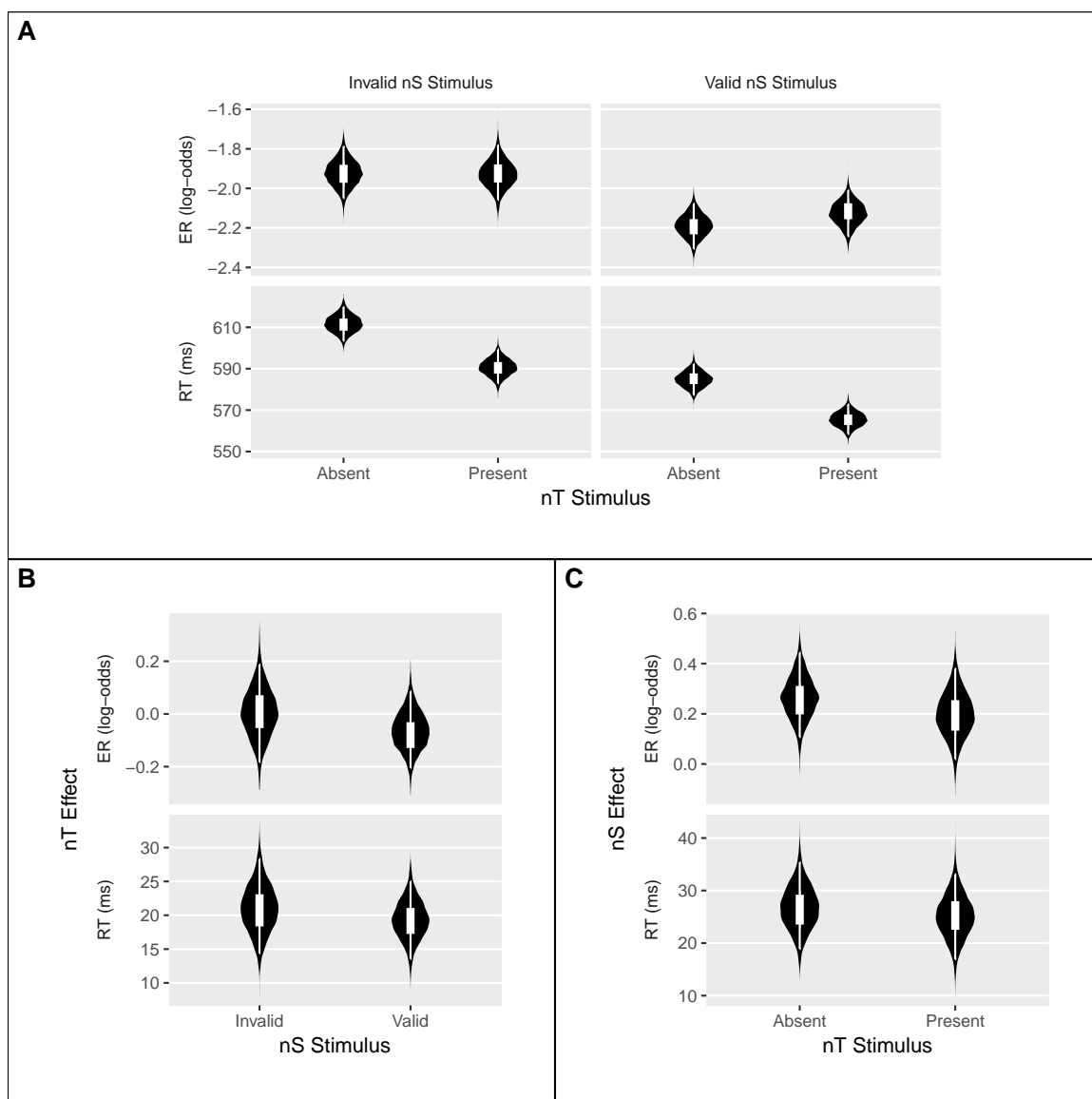


Figure D.7: Interaction of effects of nT and nS manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

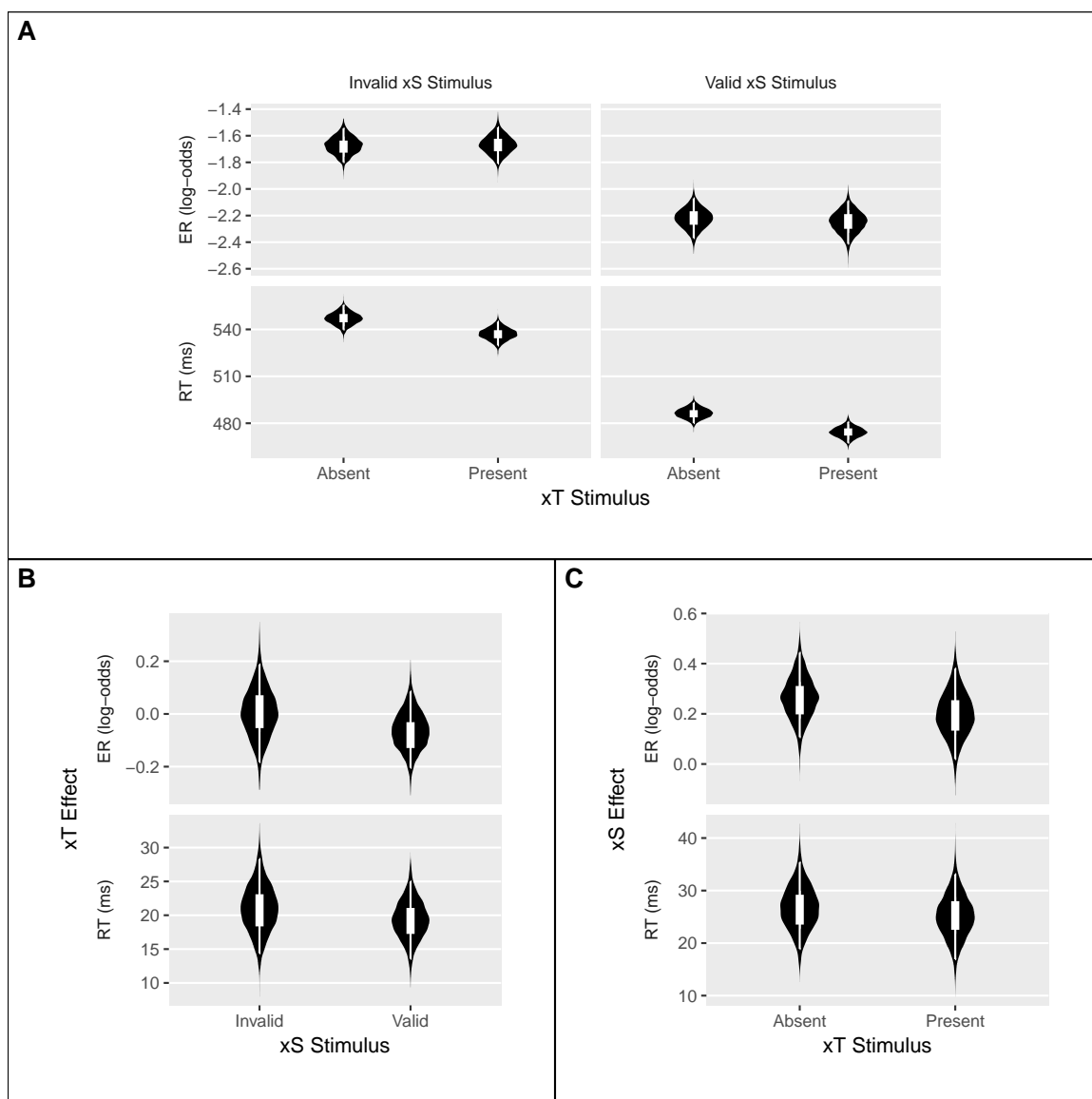


Figure D.8: Interaction of effects of xT and xS manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

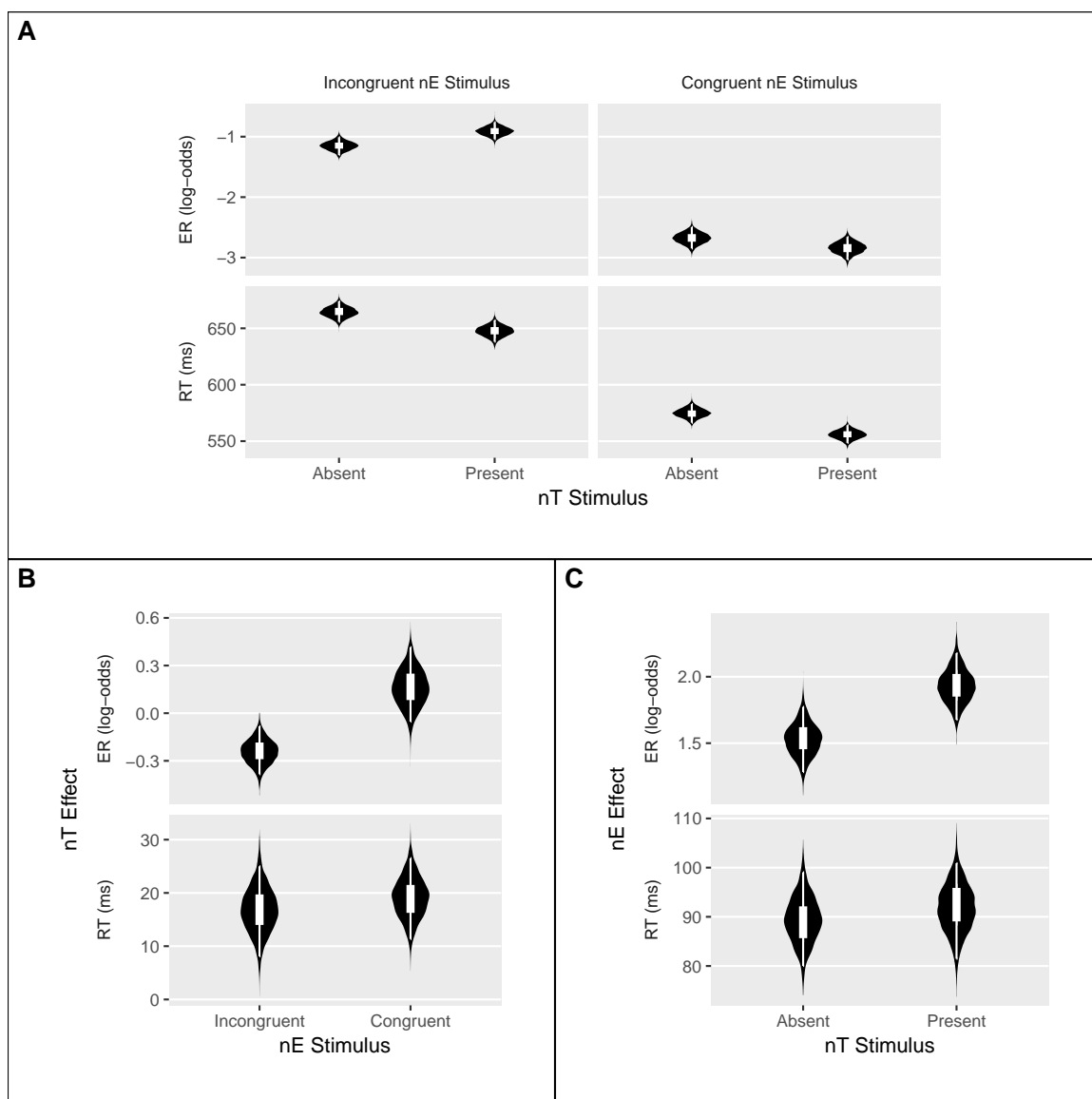


Figure D.9: Interaction of effects of nT and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

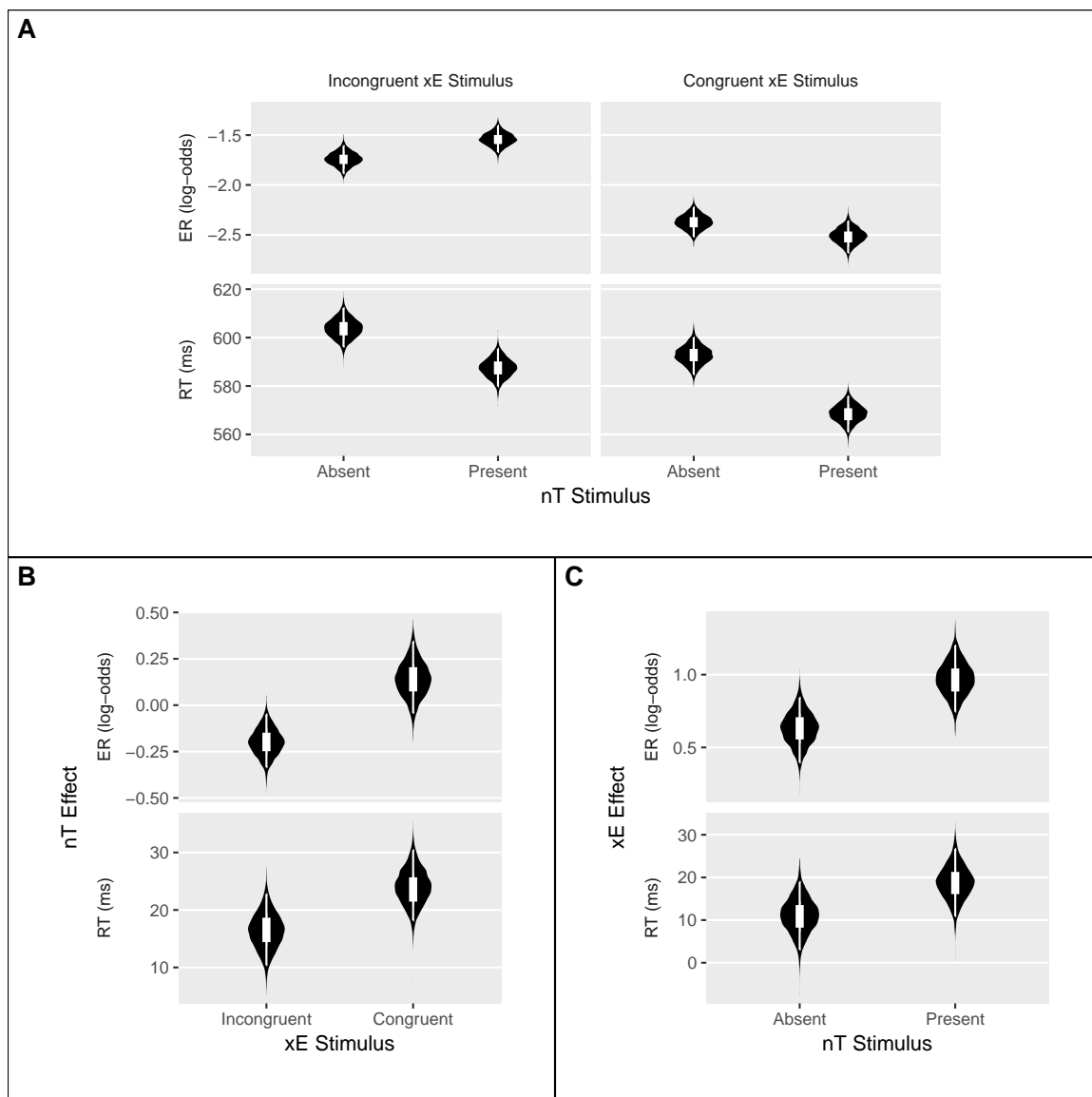


Figure D.10: Interaction of effects of nT and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

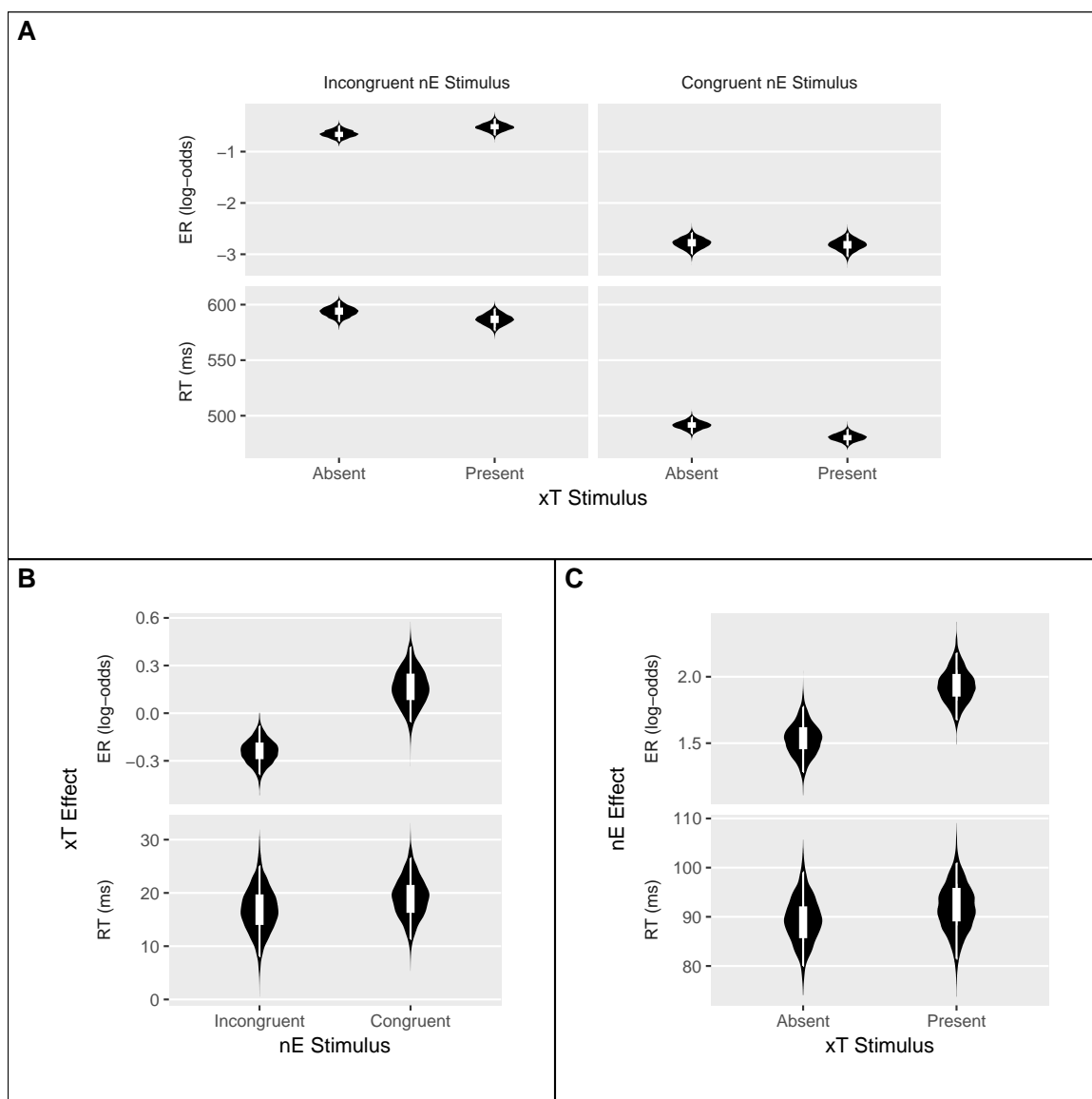


Figure D.11: Interaction of effects of xT and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

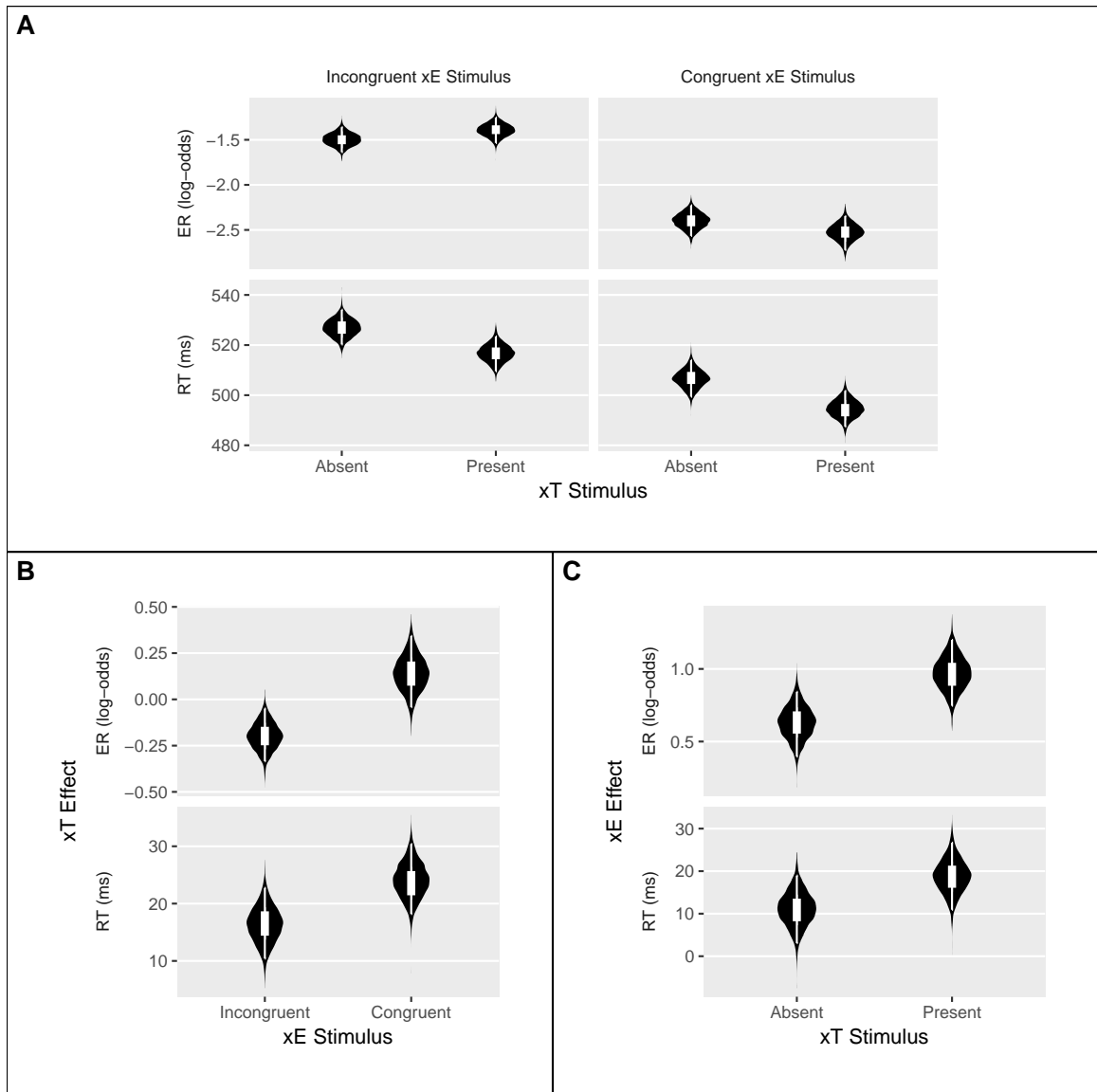


Figure D.12: Interaction of effects of xT and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

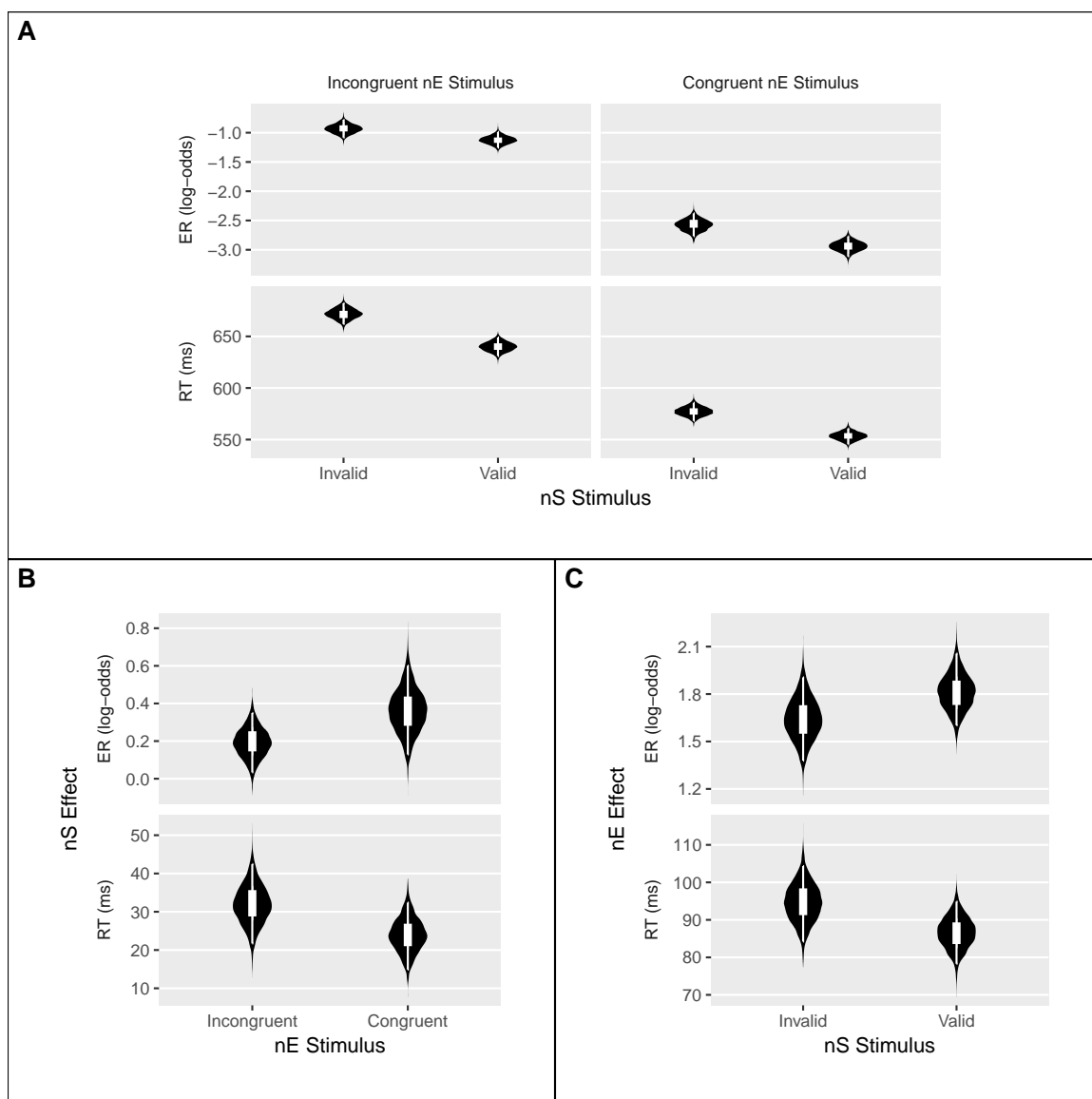


Figure D.13: Interaction of effects of nS and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

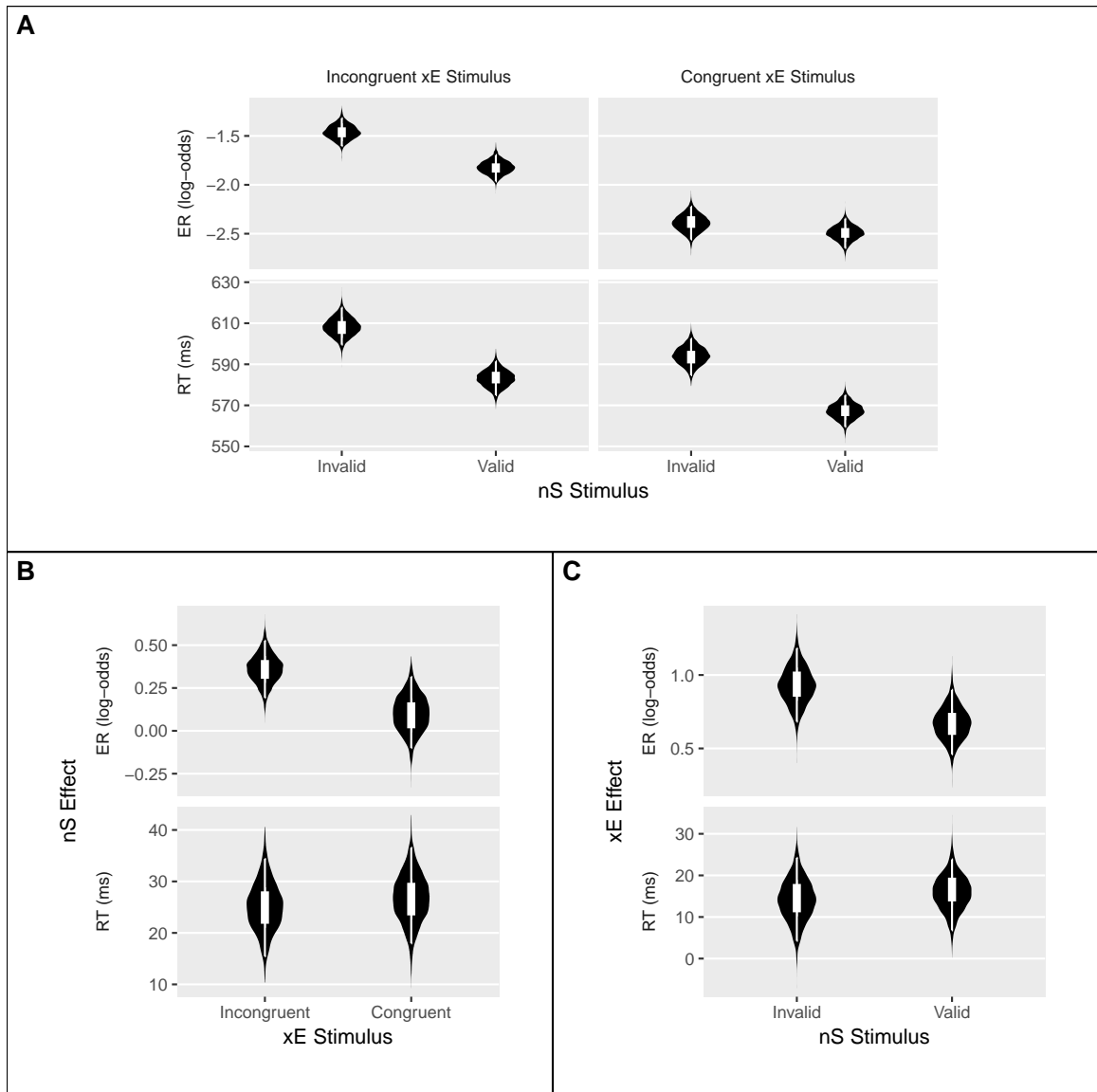


Figure D.14: Interaction of effects of nS and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

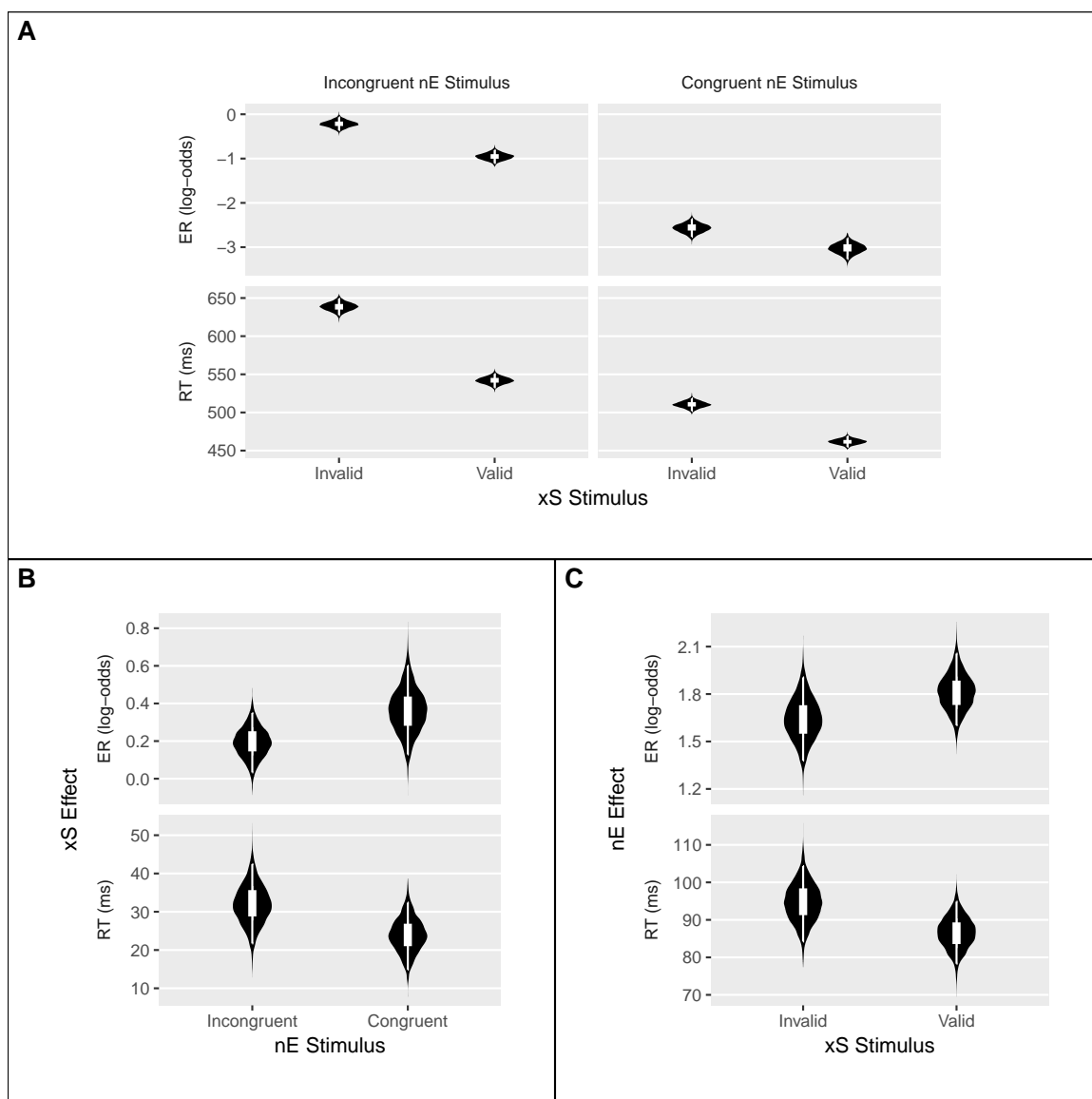


Figure D.15: Interaction of effects of xS and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

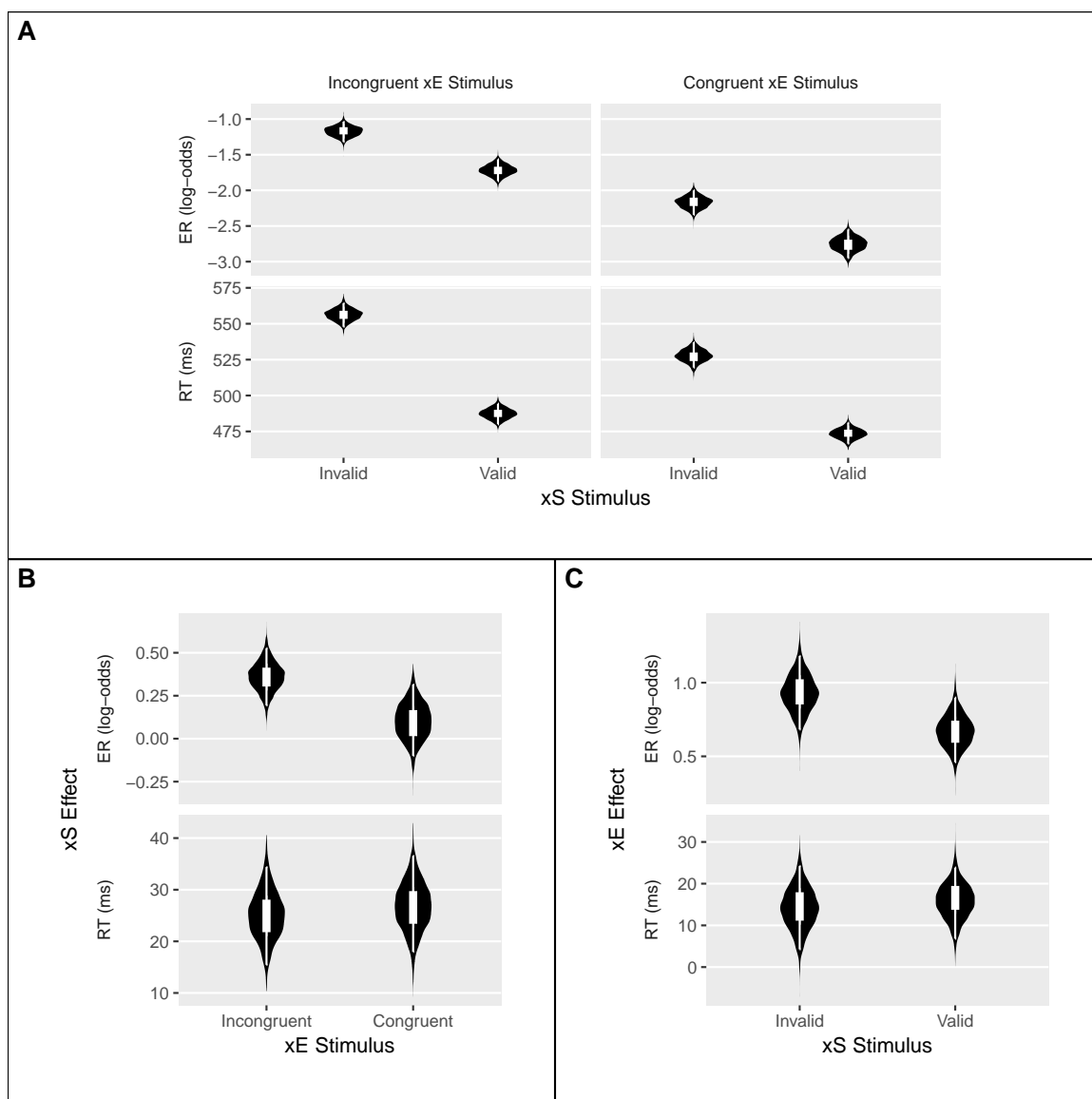


Figure D.16: Interaction of effects of xS and xE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

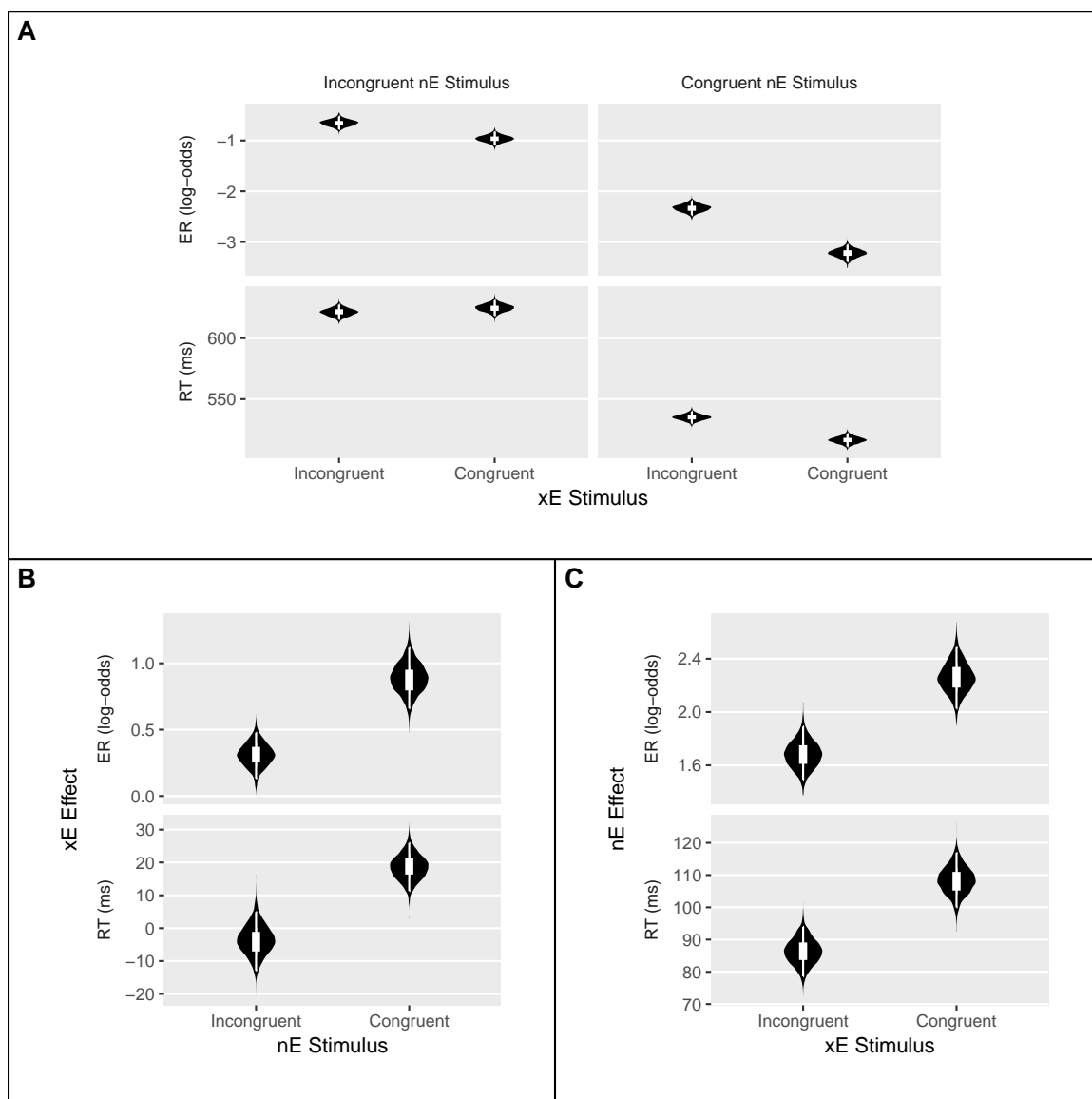


Figure D.17: Interaction of effects of xE and nE manipulations on RT and ER. Panel A shows posterior for each condition combination. Panel B and C collapse one variable to a difference score, plotted as a function of the uncollapsed variable.

Appendix E

Correlations from full developmental analysis

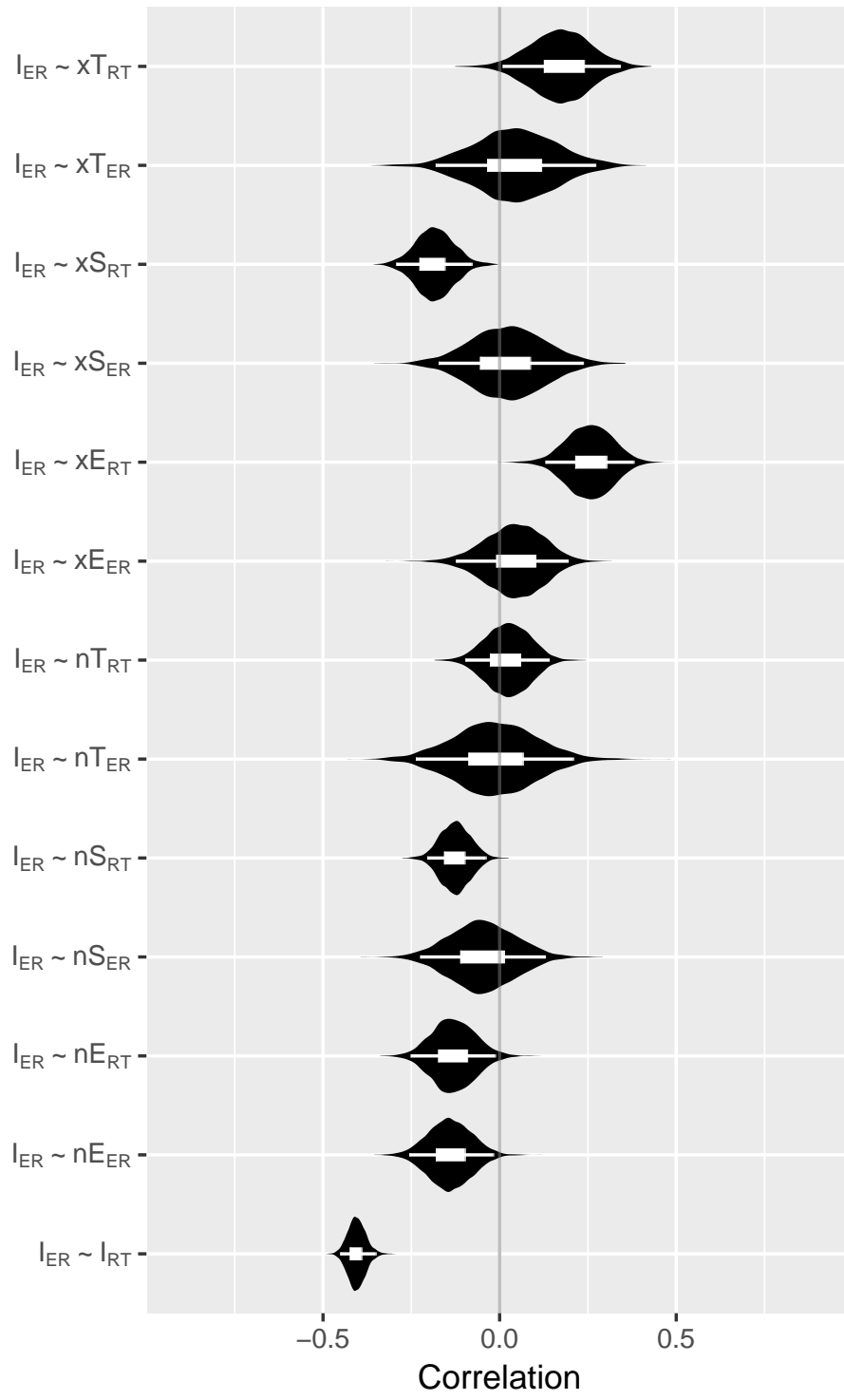


Figure E.1: Correlations subfigure 1.

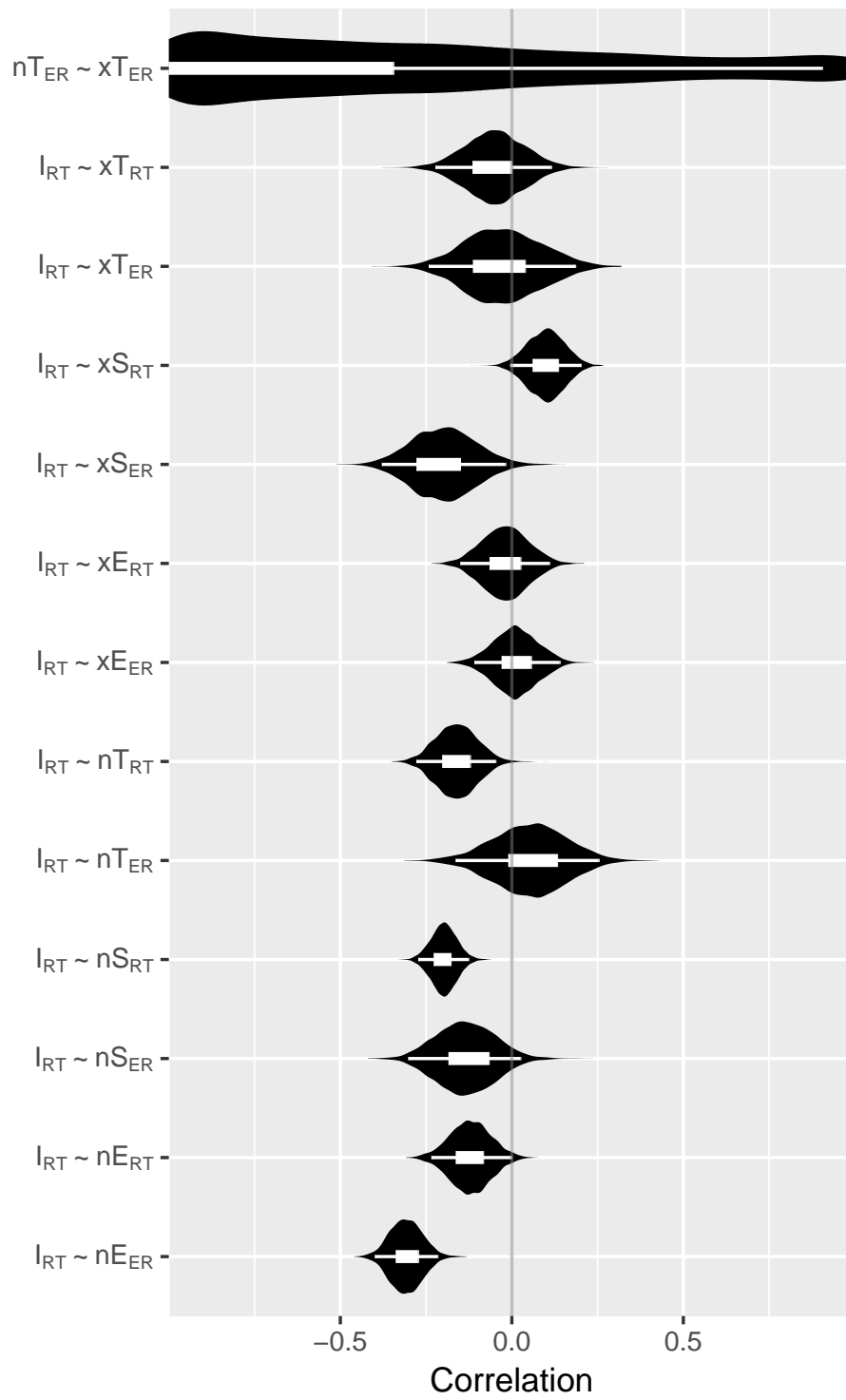


Figure E.2: Correlations subfigure 2.

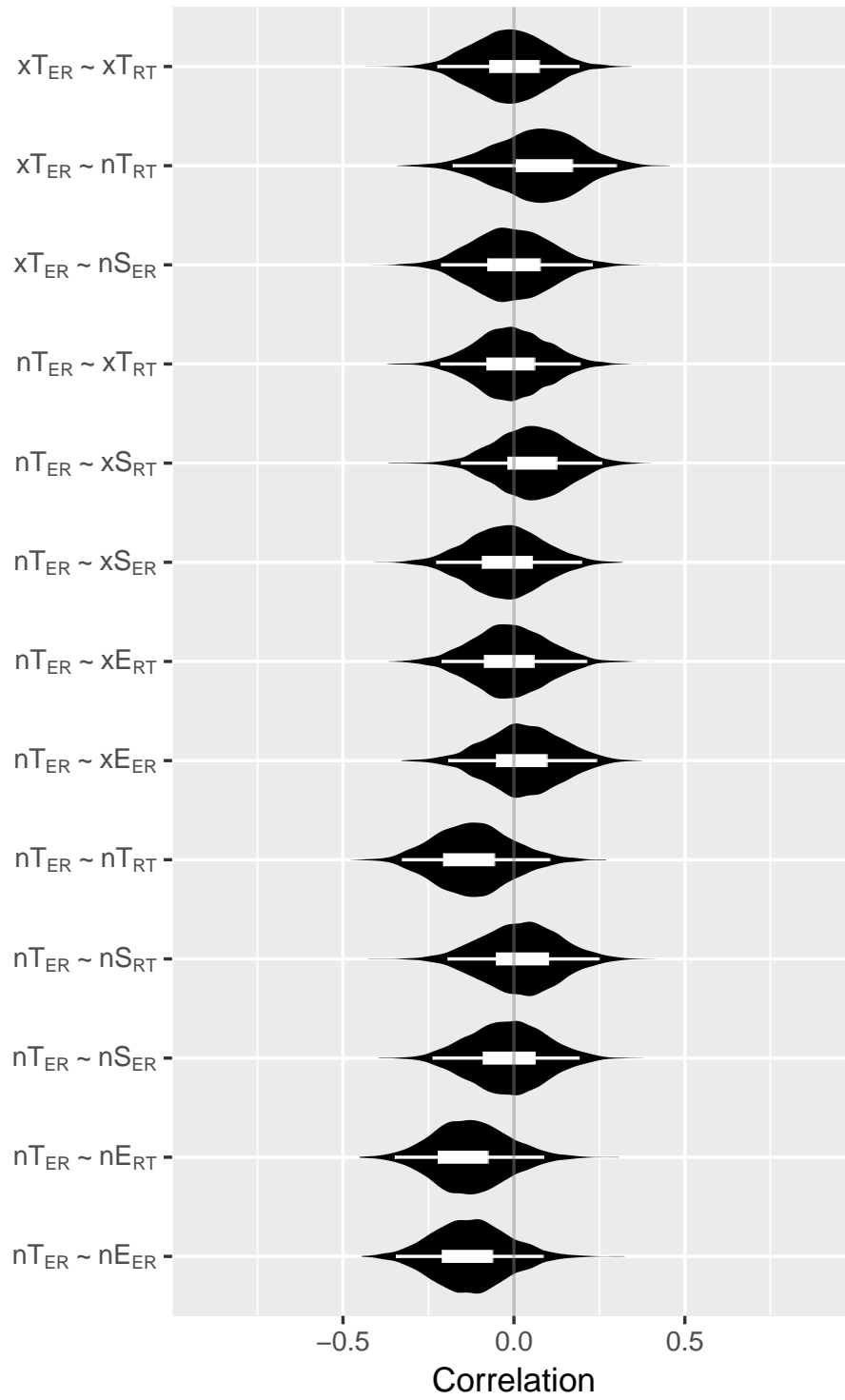


Figure E.3: Correlations subfigure 3.

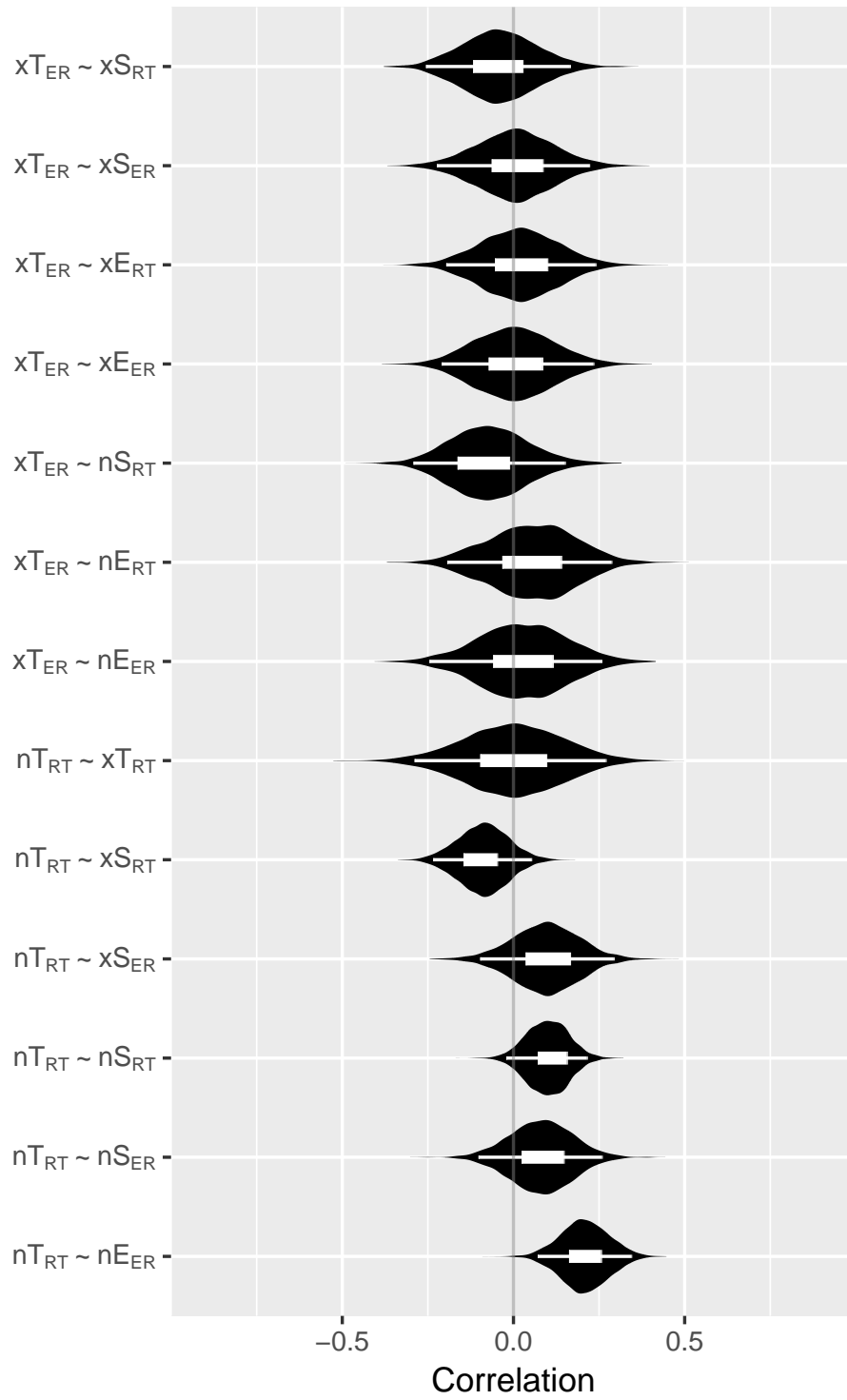


Figure E.4: Correlations subfigure 4.

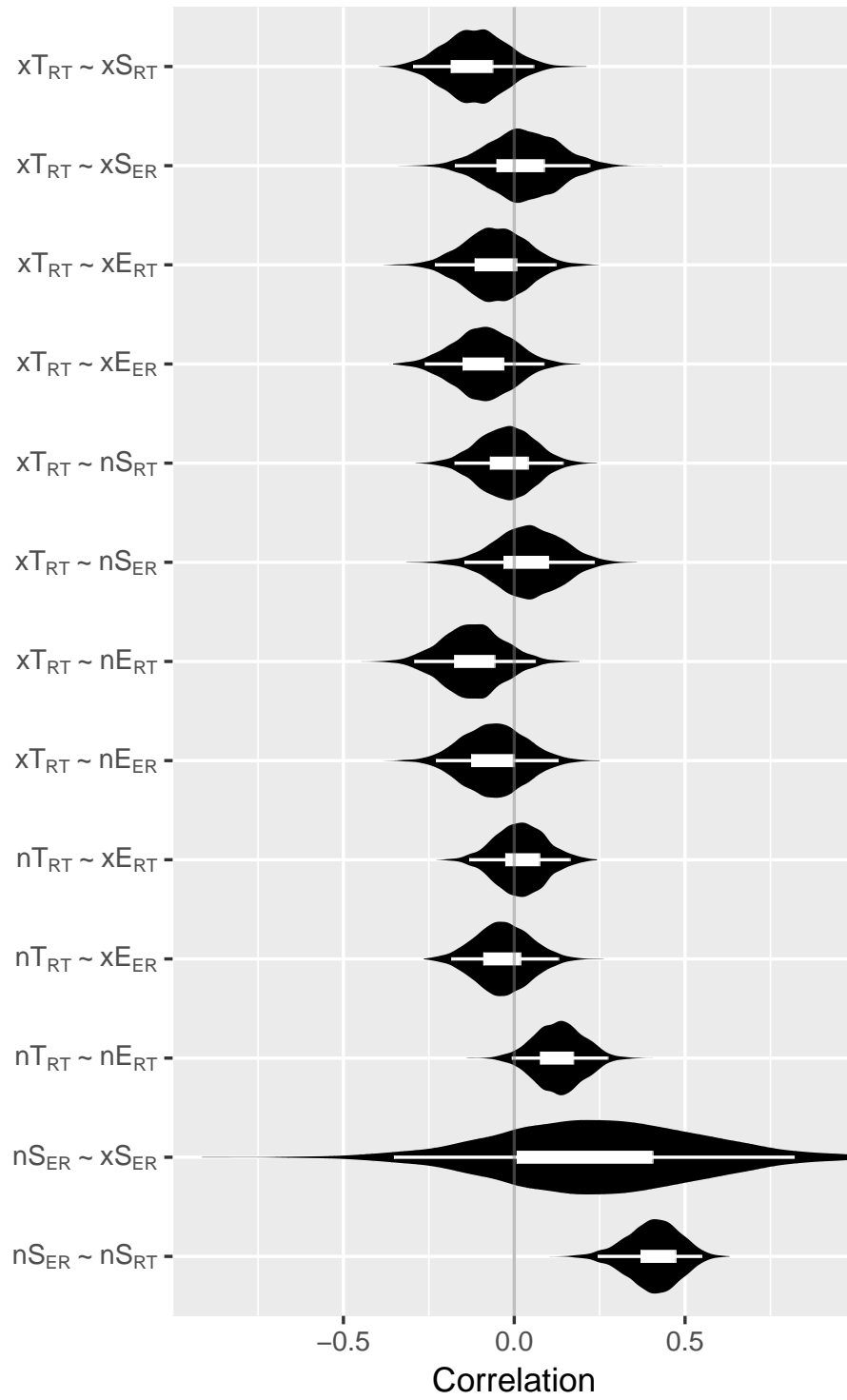


Figure E.5: Correlations subfigure 5.

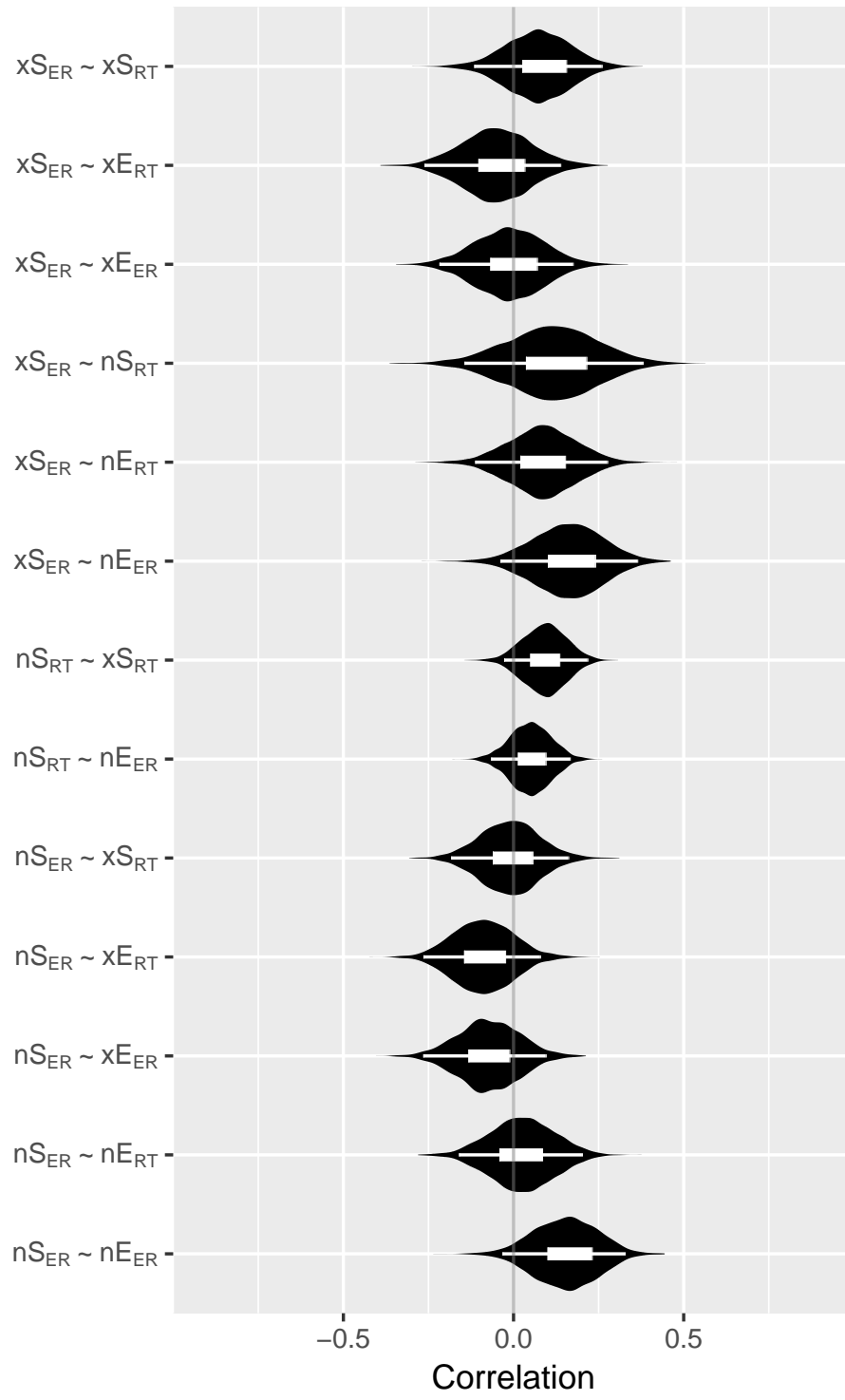


Figure E.6: Correlations subfigure 6.

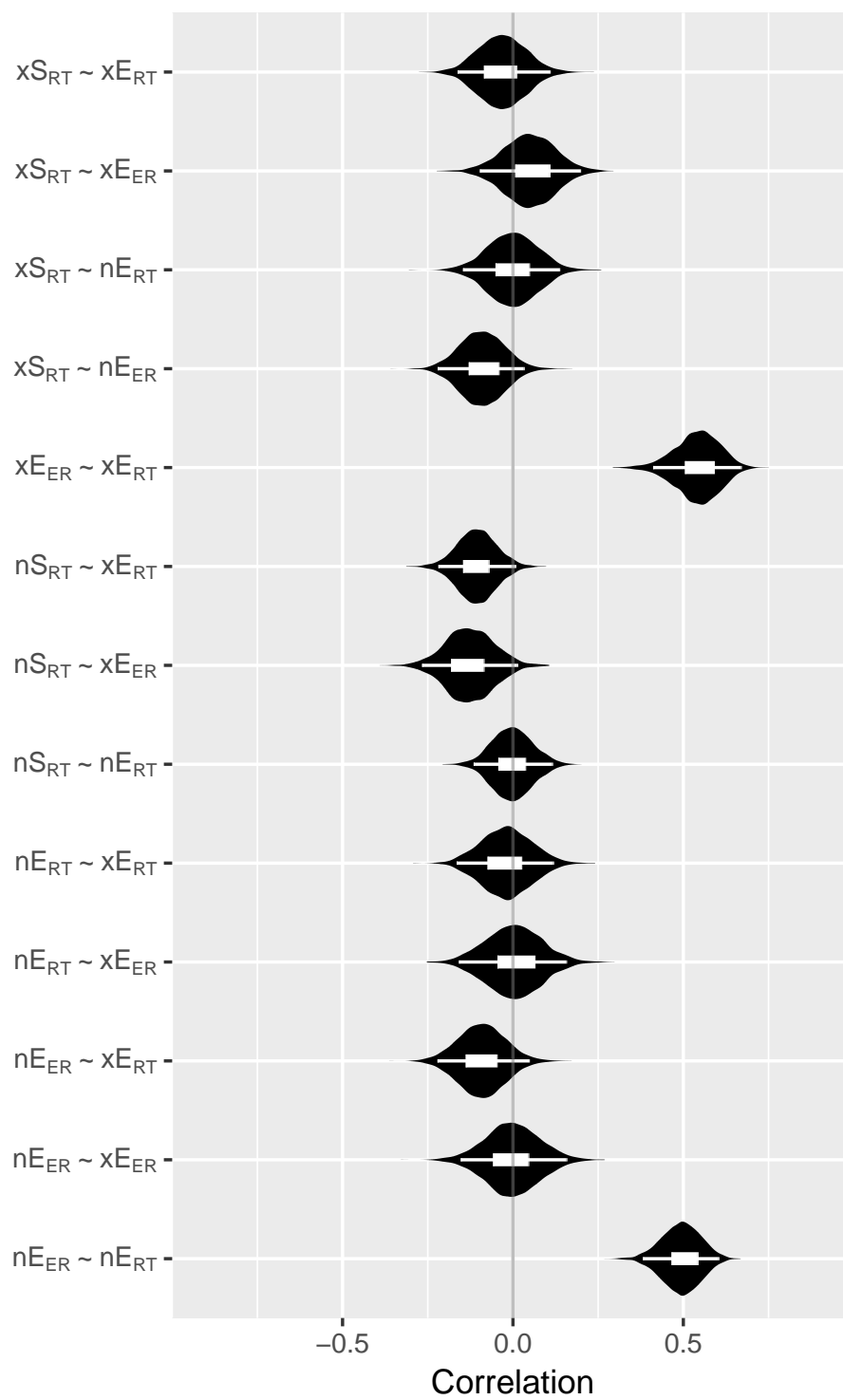


Figure E.7: Correlations subfigure 7.