Cognitive Therapy and Research
https://doi.org/10.1007/s10608-023-10409-1

ORIGINAL ARTICLE

Improved Reliability of Dot Probe Measures with Response-Based Computation: An Application with Young Violence-Exposed Children

Alysse M. Loomis[1] · Travis C. Evans[2] · Damion J. Grasso[3] · Margaret Briggs-Gowan[3]

Accepted: 28 June 2023

Abstract
Background Anxiety and posttraumatic stress symptoms have been associated with threat-related attention bias assessed in the widely-used dot-probe task. A novel method of scoring attention bias using response-based computation has shown improved reliability over standard computation approaches with adult samples.
Methods The current study applies this approach using secondary analysis of dot probe data from a sample of preschool-aged children ages 3–7 (n = 280) with varying levels of violence exposure. Attention bias indices were then examined alongside mother-report and observational measures of child anxiety. The standard approach used average reaction times to create a single measure of vigilant *or* avoidant bias per participant, whereas the response-based computation approach used trial-level reaction times to dissociate separate measures of vigilant *and* avoidant bias per participant.
Results The standard computation approach demonstrated unacceptable levels of internal consistency. In contrast, response-based computation of vigilant and avoidant bias demonstrated good and acceptable levels of internal consistency, respectively. Using the standard computation approach, no significant symptom associations were observed. Using the response-based computation approach greater ratio of vigilant bias relative to avoidant bias (i.e., vigilant bias > avoidant bias) was related to more observed anxious behaviors.
Conclusions Similar to research in adults, response-based attention bias indices demonstrated superior psychometric properties and stronger symptom associations compared to the standard computation approach and may offer advantages over the standard computation approach to study attention bias in young children.

Keywords Dot probe · Attentional bias · Early childhood · Anxiety

Introduction

Anxiety and post-traumatic stress disorder (PTSD) are both characterized by threat-related dysregulation/hypersensitivity to threat. One domain in which hypersensitivity to threat manifests is attention to environmental threat cues. The dot probe task is widely used to measure attentional biases toward or away from threat stimuli (Bar-Haim et al., 2007).

Attentional bias towards threat is a response pattern shown to be elevated in both children and adults with anxiety and PTSD (Bar-Haim et al., 2007; Briggs-Gowan et al., 2015, 2016; Dudeney et al., 2015; Pine et al., 2005).

The dot probe task has also been used with preschool aged-children (Cremone et al., 2017; Susa et al., 2012; Ursache & Blair, 2015), and in particular, young children exposed to violence (Briggs-Gowan et al., 2016; Loomis, 2020). In preschool-aged children, children with stronger attentional bias to threat had stronger links between violence exposure and anxiety symptoms which has led researchers to suggest that profiles of attention bias may differentially predict children's risk or resiliency to symptomatology after violence exposure (Briggs-Gowan et al., 2015). These findings suggest that attention bias plays an important role in the development of anxiety and stress-related disorders. Longitudinal work has also highlighted the role that attention bias may play in the developmental unfolding of trauma

✉ Alysse M. Loomis
alysse.loomis@utah.edu

1 College of Social Work, University of Utah, Salt Lake City, UT, USA

2 Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA

3 Department of Psychiatry, University of Connecticut, West Hartford, CT, USA

Published online: 10 July 2023

symptoms in young children, specifically hyperarousal and dissociation (Briggs-Gowan et al., 2016). Additionally, a number of studies with adults have used attention bias modification as a behavioral strategy to reduce anxiety symptoms (Bar-Haim, 2010), demonstrating the potential clinical utility of the dot probe paradigm with children and adults.

Limitations of Dot Probe Measures of Attentional Bias

Despite the promise of the dot probe task to identify early precursors of anxiety symptoms, there are notable limitations and critiques of dot probe-based measures of attentional bias. Among adult populations, there is considerable variability in the relationship and magnitude of the association between anxiety symptoms and attentional bias across studies (Bar-Haim et al., 2007) and some studies fail to detect effects at all (Kruijt et al., 2019). In studies of preschool-aged children, cross-sectional relationships (Briggs-Gowan et al., 2015), but not longitudinal relationships (Briggs-Gowan et al., 2016), have been identified between attention bias toward threat and anxiety. Other work has found that anxiety symptoms at preschool age (4.5 years) was related to attention bias to threat 3 years later (Aktar et al., 2019), demonstrating inconsistent longitudinal patterns between attention bias and anxiety across studies. Several studies have cited poor psychometric properties as a potential contributor to inconsistencies across studies in both youth and adult samples (Evans & Britton, 2018; Kruijt et al., 2019; Schmukle, 2005). Poor psychometric properties of dot probe measures may obfuscate developmental changes in attention bias across time and anxiety-related associations, which limits clinical utility of these measures with early childhood populations.

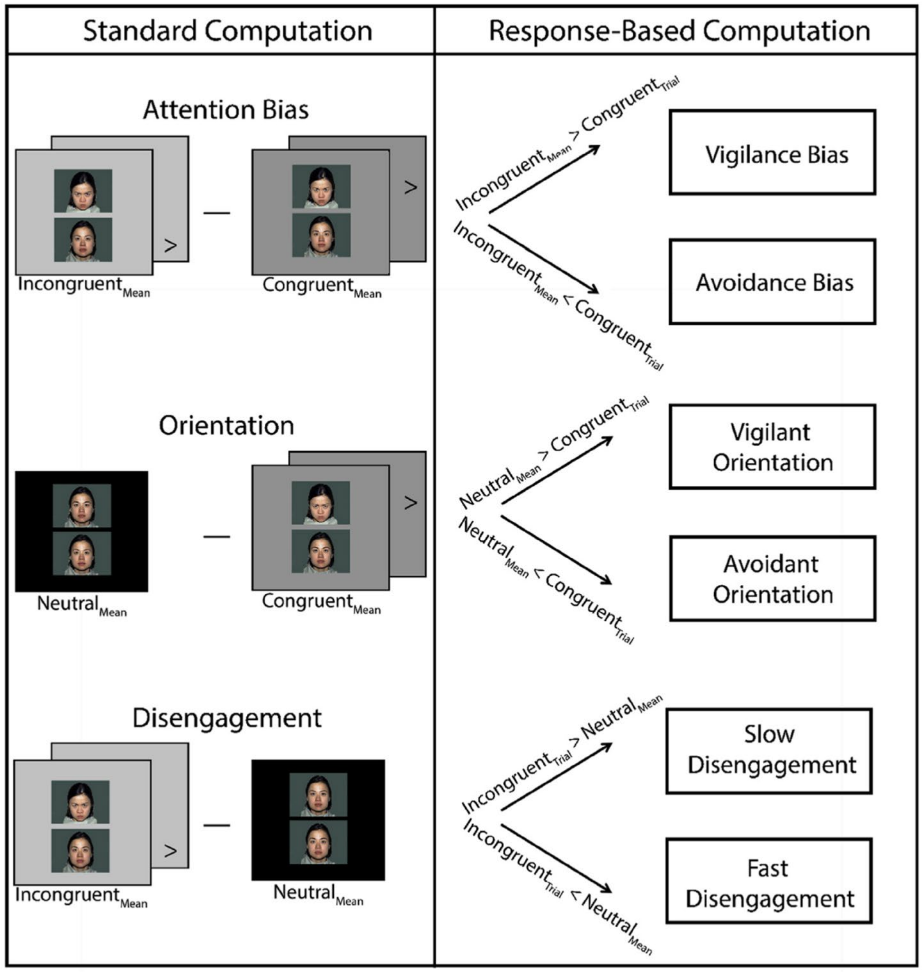
Many of the psychometric limitations and inconsistency of findings across studies using the dot probe task have been attributed to how attention bias indices from the dot probe are computed. Standard computation approaches calculate attention bias scores by subtracting average reaction times (RTs) on trials in which participants respond to a probe (e.g., an arrow or a coin) following the location of a threatening stimulus from average RTs on trials in which response probes follow the location of a neutral stimulus (mean neutral-probe RTs – mean threat-probe RTs). Positive scores indicate that a response probe was detected more rapidly following threat-related stimuli relative to neutral stimuli, which indicates attentional bias *toward* threat. Negative scores indicate that a response probe was detected more slowly following threat-related stimuli compared to neutral stimuli, which indicates attentional bias *away from* threat. Using average RTs in this manner, attentional bias scores can be categorized as "bias to threat" "bias away from threat" or "no attentional bias" (Briggs-Gowan et al., 2016).

However, using average RTs to compute attentional bias ignores variability across trials and the possibility that individuals with anxiety may exhibit both vigilance and avoidance at different time points in the task. Thus, standard computation approaches characterize a participant as *generally* exhibiting either attentional bias toward or away from threatening images, and ignores the intra-individual variability (e.g., variability across the timing of the task) viof threat-related attention (Evans & Britton, 2018; Zvielli et al., 2015). This limitation is particularly salient to consider within a developmental framework given that children exhibit greater variability in behavior over time (Wakschlag et al., 2015) which may correspond with variable attentional biases as well. Additionally, research suggests that computing a difference score between two highly correlated reaction time measures results in poor psychometric properties that may also obfuscate anxiety-related associations (Miller & Ulrich, 2013). However, difference scores in other attention paradigms such as the emotional Stroop task and Flanker task exhibit acceptable psychometric properties, which suggests that dot probe measures may be particularly susceptible to this issue (Evans & Britton, 2018).

New Approaches for Computing Attention Bias from the Dot Probe

To address the critiques to standard computation approaches, novel trial-based measures have been introduced. For example, Zvielli et al. (2015) introduced a "nearest neighbor" computation approach, which calculates attention bias by comparing reaction times between consecutive trials. However, this approach may reduce the number of trials available for analysis, which is of particular concern with samples of young children who may not be able to complete longer dot probe tasks due to limited attention and distractibility (Loomis, 2020). Studies that have used simulation data to examine these trial-level measures have also demonstrated that these measures do not capture threat-related attention specifically, but rather more general intra-individual variability in reaction times (Kruijt et al., 2016). To address limitations associated with standard and nearest-neighbor computation approaches, Evans and Britton (2018) developed and validated a response-based computation approach that compares each individual trial to a mean reference reaction time. For example, trial-level RTs on *congruent* trials are compared against a participant's mean RT on neutral trials to compute separate measures of vigilance towards and avoidance away from threat based on intraindividual variability (see Fig. 1). Similarly, trial-level RTs on *incongruent* trials are compared against a participant's mean RT on neutral trials to compute separate measures of slow disengagement

Fig. 1 Standard and response-based computation. Figure used from Evans and Britton (2018) with permission from Elsevier



Note: Figure used from Evans & Britton, 2018 with permission from Elsevier

from threat and faster disengagement from threat based on intraindividual variability. Moreover, these measures can be used to compute an intra-individual ratio of vigilance/avoidance and slow disengagement/faster disengagement (for a full description, see Evans & Britton, 2018), which is still limited by the time frame of the assessment, offers expanded information about intra-individual variability.

Compared to standard computation approaches, trial-based measures exhibit superior psychometric properties, such as better internal consistency (Evans & Britton, 2018; Meissel et al., 2021) and overall improved split-half reliability (Evans & Britton, 2018; Meissel et al., 2021; Zvielli et al., 2015). Trial-based measures also exhibit anxiety-related associations in adults that are not detected by standard computation approaches, possibly due to these psychometric improvements. For example, Zvielli et al. (2015) found that temporal variability of attention bias, or degree of fluctuating bias toward or away from threatening images across time, significantly discriminated between phobic and healthy control groups. Gade et al. (2021) demonstrated a link between higher mean bias toward threat and social anxiety when using trial-level bias scores, but not when using standard computation approaches. Evans and Britton (2018) identified that greater vigilant orientation relative to avoidant orientation was significantly related to greater state anxiety, which was not observed using the standard computation approach.

Current Study

Previous studies have been conducted entirely in adult samples. Attentional bias may be more "set" in adults due to fully developed cognitive control capacities, but more "fluid" in youth due to still developing cognitive control capacities. As a result, using a trial-based computation approach may be especially important for child samples in which attention bias is even more variable across trials.

The current study aims to address this gap by applying the novel response-based computation approach introduced by Evans and Britton (2018) to a large preschool sample

($n=280$) using parent-report and observed measures of anxiety. Compared to standard attention bias measures, we hypothesized that response-based measures would exhibit superior psychometric properties (e.g., internal consistency and split-half reliability) and improved detection of anxiety-related associations.

Method

Participants

The current study is a secondary analysis of data from a larger study. Participants in the current study represent a subsample of a survey cohort of 1857 children ages 3- to 5-years-old who were receiving well-child care at one of 5 pediatric practices (Wakschlag et al., 2014). The subsample was selected for a substudy about developmental psychology and violence exposure (Mian et al., 2015). A stratified random sample of 497 children were drawn for the substudy, in which children with disruptive behaviors (i.e., who scored above the 80th percentile on a multidimensional assessment of disruptive behavior) or children whose mother reported past-year interpersonal violence were oversampled. Children with significant developmental delays or neurodevelopmental conditions were excluded from the substudy. Prior work has demonstrated that approximately a third of the children in the substudy (35.2%) are classified as "low exposure" to family violence (i.e., violence directed toward the child and/or between caregivers), 15% as being polyvictimized, and half (49.5%) as exposure primarily to harsh parenting (Grasso et al., 2016). Of the original 497 children in the substudy, 335 had some dot probe data. During data cleaning 55 children were removed (as noted below) and the 280 children from Wave 1 of the study with usable dot-probe data after cleaning were included in the current study.

Children included in the current sample ranged from 3 to 7 at the preschool-aged visit ($M=4.98$, $SD=0.79$). Out of the sample of 280 children in the current sample, 47% ($n=131$) were male and 53% ($n=149$) female. Almost one third of children in the sample, (29.3%, $n=82$) were Hispanic/Latine. Approximately half of the sample (48.9%, $n=137$) were Black or African American, 32.1% ($n=90$) White, 3.6% ($n=10$) Asian, Native Hawaiian, or Native American, and 3.2% ($n=9$) another race.

Procedures

In the intensive sub study, children and their English-speaking, biological mothers and children attended two 3-h laboratory visits at the first wave of data collection. In the first laboratory visit, parents completed measures of children's symptoms and functioning as well as a parent–child interaction experiment to observationally measure anxiety. In the second laboratory session, the dot-probe task was administered to children and parents completed an in-depth interview about family stress and violence. Children completed either the standard 180-trial version (44.5%) or an extended 360-trial version required for a different sub study on even-related potential (55.5%). There were no differences in attention bias scores, accuracy, or RT between the two trial versions. To standardize the number of trials available for this secondary analysis, only the first 180 trials were used to ensure that all participants had the same amount of data. Mothers provided informed consent and were compensated for participation and transportation. Mandated reporting procedures were followed and staff monitored children for fatigue or distress, providing breaks and discontinuing assessments when appropriate. All study protocols were approved by the relevant institutional review boards.

Measures

Attention Bias

Attention bias to threat was measured using the preschool dot-probe task (see Briggs-Gowan et al., 2015 for more information). Each trial began with the presentation of a 500-ms central fixation cross, followed by the 500-ms presentation of a side-by-side pair of faces from the Nim-Stim stimulus set. Face pairs were comprised of emotional facial expressions that were either angry, happy, or neutral (angry–neutral, happy–neutral, neutral–neutral). Following 500 ms of stimulus presentation, faces were removed from the screen and a response cue (a coin) was randomly presented on the left or right side of the screen. The coin remained on the screen until a button was pressed using a button box. Every 90 trials, children were provided a break during which plastic gold coins were placed in a transparent piggy back by the research assistant. All children received prizes at the end of the task regardless of task performance. See Table 1 for descriptives of key study measures.

Data cleaning followed previously established protocols for the preschool dot probe task (see Briggs-Gowan et al., 2015) to exclude trials. Trials were excluded if they were inaccurate, < 200 ms (702 trials, 1.2% of all trials excluded), > 7000 ms (1530 trials, 2.6% of trials excluded), or > 2.5 SD from an individual child's mean reaction time (1820, 3.2% of trials excluded). Children's scores were excluded if their accuracy (correct trials/overall trials) was ≥ 65%, a threshold established in prior work with young children (Briggs-Gowan et al., 2015). This resulted in 55 children being excluded, with accuracies ranging from 11 to 64%. Following data cleaning, usable scores were obtained for 84% of children (n = 280), who had a mean accuracy score of 90% ($SD=8\%$). For the standard scoring, data were

Table 1 Attention bias scores and key study variables

	M (SD)
Standard computation (n = 276)	
Attention bias	−6.23 (118.35)
Orientation	2.28 (129.41)
Disengagement	−9.45 (137.56)
Response computation (n = 280)	
Vigilant bias	161.97 (92.47)
Avoidant bias	169.16 (101.67)
Vigilant orientation	168.25 (103.37)
Avoidant orientation	165.97 (99.59)
Slow disengagement	162.98 (98.84)
Fast disengagement	172.44 (112.69)
Anxiety/fear measures	
Anx-DOS (n = 260)	4.23 (2.60)
GAD (n = 259)	1.59 (1.37)
PAPA (n = 264)	0.88 (0.82)

also not used to calculate scores if there were less than 9 usable RTs for each emotion condition (n = 4), however these data were used for the response-based scoring. Results using the angry trials are reported in the manuscript and results using the happy trials are reported in the Supplemental Information section (see Tables S1–S4).

Anxiety/Fear Measures

Both observational and survey measures of anxiety were collected.

Observed Anxiety Observed anxiety was captured using the Anxiety Dimensional Observation System (ANX-DOS; Mian et al., 2015), an observational paradigm that that uses specific stimuli (e.g., remote-controlled spider, separation from parent) to elicit clinically salient behaviors. Both children and mothers were observed and global coding was used to capture normal to atypical indicators of fear behaviors and affect, physical avoidance and exaggerated startle (0 = no evidence, 1 = mild/normative, 2 = of concern, 3 = atypical). Interrater agreement for a randomly selected 20% of the sample was monitored in an ongoing fashion throughout the coding process. Interrater agreement, indexed by the ICCs ranged from 0.71 to 0.84 (Mian et al., 2015). The fear composite, which included the mean of the Fear Arousal, Physical Avoidance, and Exaggerated Startle codes of the ANX-DOS, was used for the current study.

Reported Anxiety Reported anxiety was obtained from mothers' reports in the Preschool-Age Psychiatric Assessment (PAPA; Egger & Angold, 2004), a semi structured diagnostic interview based on the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev.; DSM-IV-TR; American Psychiatric Association, 2000) that was administered by trained research assistants. A measure of total anxiety, calculated by summing specific phobia, social phobia, generalized anxiety, and separation anxiety symptoms (PAPA Fear), and a measure of total generalized anxiety symptoms (# GAD Symptoms) were used for the current study. Reliability of administration and scoring was monitored for 20% of interviews by an expert clinical psychologist (Percent agreement = 81–98%).

Data Scoring

Standard Computation Approach

Attentional bias to threat was computed by subtracting average reaction times for congruent trials (e.g., coin appears on the same side as the angy face) from average reaction times for incongruent (e.g., coin appears on the opposite side as the angry face) trials (i.e., $RT_{IncongruentMean} - RT_{CongruentMean}$). To calculate orientation and disengagement from threat, average reaction times on neutral trials were used. Positive and negative scores for orientation ($RT_{NeutralMean} - RT_{CongruentMean}$) indicate faster or slower orientation to threat, respectively. Positive and negative disengagement scores ($RT_{IncongruentMean} - RT_{NeutralMean}$), indicate slower disengagement or faster disengagement from threat, respectively.

Response-Based Computation Approach

Response-based scores were calculated using the approach outlined by Evans and Britton (2018), in which trial-level scores were referenced against each child's mean reference reaction time. For response-based attentional bias measures, the reaction time from congruent trials was individually indexed against the mean reaction time of incongruent trials as a reference (i.e., $RT_{IncongruentMean} - RT_{Congruent\,[Trial1\,\ldots\,Trial2\,\ldots\,Trialn]}$), which was used to assign each trial as a vigilance response (i.e., $RT_{difference} > 0$ ms) or avoidance response (i.e., $RT_{difference} < 0$ ms). Children had an average of 17.51 vigilant bias trials (*SD* = 4.82) and 12.38 avoidant bias trials (*SD* = 4.18). Trial-level scores were then averaged within response-based conditions to create a separate measure of average vigilance and avoidance for each child. The same trial-level approach, using the mean reaction time of neutral trials as a reference, was used to calculate response-based measures of orientation and disengagement. Children had on average 17.72 vigilant orientation trials (*SD* = 4.97) and 12.19 avoidant orientation trials (*SD* = 4.31). Children had on average

Table 2 Split half reliability estimates

	Spearman-brown corrected Split half reliability r (95% CI)	Permutation-based Split half reliability r (95% CI)
Standard scoring		
Angry bias	−0.14 (−0.32, 0.06)	−0.08 (−0.19, 0.03)
Response scoring		
Vigilant bias	0.88 (0.84, 0.91)	0.79 (0.73, 0.83)
Avoidant bias	0.68 (0.61, 0.74)	0.51 (0.44, 0.59)
Vigilant orientation	0.88 (0.84, 0.91)	0.78 (0.73, 0.83)
Avoidant orientation	0.75 (0.69, 0.80)	0.60 (0.53, 0.66)
Slow disengagement	0.67 (0.6, 0.73)	0.50 (0.42, 0.57)
Fast disengagement	0.90 (0.87, 0.92)	0.81 (0.77, 0.85)

12.19 slow disengagement trials ($SD = 4.31$) and 17.60 fast disengagement trials ($SD = 5.19$). Relative absolute magnitude scores for each type of response (e.g., Vigilance > Avoidance) were also calculated using a ratio index (e.g., |Vigilance|:|Avoidance|).

Data Analysis

The following analyses were conducted to examine internal consistency and criterion validity. All analyses were conducted in SPSS Version 28.

Internal Consistency

For standard scoring, split-half reliability was run by computing attention bias scores separately for odd and even trials and running split-half calculations with Spearman-Brown corrections. Since split-half estimates in dot probe tasks tend to be unstable (Parsons et al., 2019), we also obtained permutation-based split-half reliability estimates using the *splithalf* package in R (Version 0.8.2; Parsons, 2020), which averages the results of 5000 random splits. Both estimates, along with 95% confidence intervals are presented in Table 2.

Cronbach's alpha was computed by dividing RTs within conditions into 2-trial bins to form increasing numbers of bins up to a maximum of 24 bins (e.g., incongruent trial 1 and trial 2 formed a bin; Schmukle, 2005). Within each 2-trial bin, a single index score was calculated (e.g., [$RT_{IncongruentMean\ 1,2,3,4} - RT_{CongruentMean\ 1,2,3,4}$], see Fig. 2).

For response-based computation we computed response-based scores separately for odd and even trials and ran split-half calculations with Spearman-Brown corrections and also computed a permutation-based split-half reliability taking the average of the calculated score (e.g., vigilance bias scores) across 5000 random splits, as discussed above. Based on the score being analyzed, from 1 to 3 participants did not have adequate trial-level data for the analysis. To ensure maximum generalizability across estimates, we excluded the cases for the 6 participants who did not have adequate data on one of the scores from all estimates. For rigor, we also ran the split-half analyses with all eligible participants and results were similar to those with excluded participants, so the sample reported on for split-half reliability is the sample with all ineligible cases removed for consistency (n = 271).

For response-based computation, Cronbach's alpha was computed by computing the Cronbach's alphas for trials within a specific type (e.g., avoidant orientation), dividing RTs within conditions into 2-trial bins until the sample size fell below 25%, typically 38 trials (see Fig. 2). For example, we first calculated the Cronbach's alpha for every participant's first and second trial that was included under avoidant orientation, then we calculated reliability for every participant's first through fourth avoidant orientation trial. Since one trial might fall under avoidant orientation for one participant and vigilant orientation for another participant, the timing of "trial 1" and the number of trails between "trial 1" and "trial 2" for each participant would also vary. Internal consistency was defined as acceptable (0.70–0.79), good (0.80–0.89), and excellent (> 0.90).

Criterion Validity (Anxiety-Related Associations)

Criterion validity was compared by examining correlations with observed and parent-reported anxiety. We also controlled for overall reaction time ($RT_{Overall}$) and reaction time variability ($RT_{OverallStdDev}$), to adjust for the influence of overall reaction time variability on trial-level attention measures. As suggested by prior simulation studies (Kruijt et al., 2016), we included these variables as covariates to ensure that anxiety-related associations could not be attributed to RT variability more generally.

Results

Internal Consistency

Standard Attention Computation

Using the standard computation approach, all attention measures demonstrated unacceptable levels of internal consistency. Specifically, attention bias demonstrated unacceptable levels of internal consistency across estimates of both Spearman-Brown corrected ($r_{SB} = -0.14$) and permutation-based split-half reliability ($r_P = -0.08$) and Cronbach's alpha [max $\alpha = 0.26$ at 6 quartets; see Fig. 2b]. Similarly, both the orientation [r (277) = 0.03, $p = 0.603$; max $\alpha = 0.18$ at 9 doublets] and disengagement components [r (277) = 0.06, $p = 0.325$; max $\alpha = 0.41$

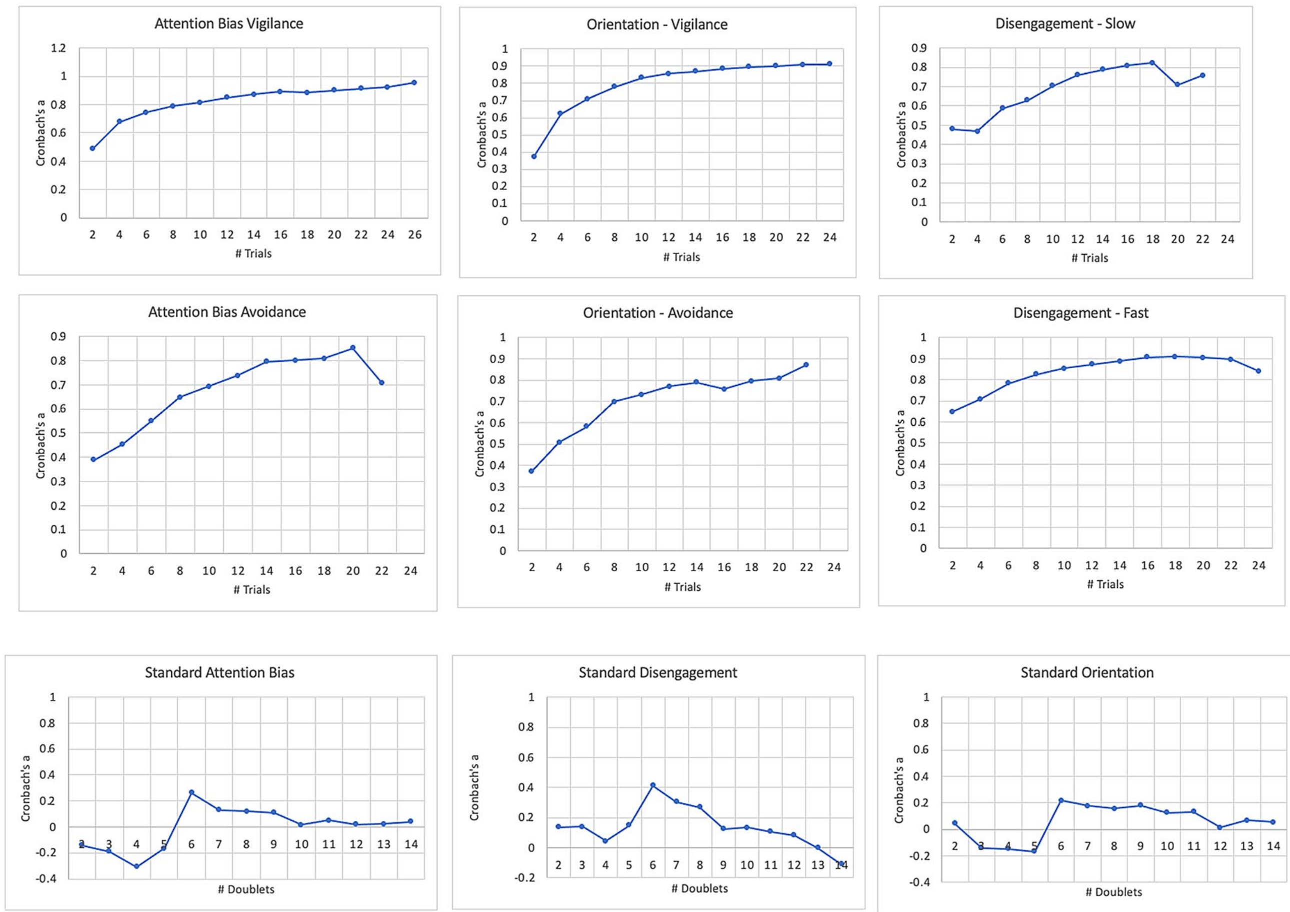


Fig. 2 Internal consistency across standard and response-based computation approaches

at 6 doublets] did not reach acceptable levels of internal consistency.

Response-Based Computation

Using incongruent trials as a reference to compute vigilance and avoidance bias, vigilance demonstrated acceptable to good levels of split-half reliability ($r_{SB}=0.88$, $r_P=0.79$) and avoidance did not demonstrate acceptable levels of split-half reliability ($r_{SB}=0.68$, $r_P=0.51$; see Table 2 for 95% CIs). Vigilant bias reached excellent levels of internal consistency [acceptable ($\alpha=0.74$, $p<0.001$) at 6 trials, good ($\alpha=0.81$, $p<0.001$) at 8 trials, and excellent ($\alpha=0.91$, $p<0.001$) at 22 trials] and avoidant bias reached good levels of internal consistency [acceptable ($\alpha=0.74$, $p<0.001$) at 12 trials and good ($\alpha=0.80$, $p<0.001$) at 16 trials].

Using neutral trials as a reference to compute vigilance and avoidance orientation, vigilant orientation reached good and acceptable levels of split-half reliability ($r_{SB}=0.88$, $r_P=0.78$) and avoidant orientation demonstrated acceptable levels of split-half reliability with only the Spearman-Brown corrected split-half reliability ($r_{SB}=0.75$, $r_P=0.60$). Vigilant orientation reached excellent levels of internal consistency [acceptable ($a=0.71$, $p<0.001$) at 6 trials, good ($\alpha=0.83$, $p<0.001$) at 10 trials, and excellent ($\alpha=0.90$, $p<0.001$) at 18 trials] and avoidant orientation reached good levels of internal consistency [acceptable ($a=0.70$, $p<0.001$) at 8 trials and good ($\alpha=0.80$, $p<0.001$) at 18 trials].

Fast disengagement demonstrated excellent and good levels of split-half reliability ($r_{SB}=0.90$, $r_P=0.81$) and slow disengagement did not reach acceptable levels of split-half reliability ($r_{SB}=0.67$, $r_P=0.50$). Slow disengagement reached good levels of internal consistency [acceptable ($\alpha=0.70$, $p<0.001$) at 10 trials and good ($\alpha=0.81$, $p<0.001$) at 16 trials] and fast disengagement reached excellent levels of internal consistency [acceptable ($\alpha=0.71$, $p<0.001$) at 4 trials, good ($\alpha=0.83$, $p<0.001$) at 8 trials, and excellent ($\alpha=0.91$, $p<0.001$) at 16 trials].

Criterion Validity (Anxiety-Related Associations)

Using standard approaches, no significant associations emerged between anxiety measures and standard scores (see Table 3, all p's > 0.25).

Using the response-based approach, no significant associations emerged between observed and parent-reported anxiety measures and response-based scores (all p's > 0.20; see Table 4). Among relative magnitude measures, greater vigilant bias *relative* to avoidant bias (i.e., Vigilant Bias > Avoidant Bias) was marginally associated with higher Anx-DOS scores [r (258) = 0.12, $p = 0.051$] after controlling for RT_{Mean} and RT_{StdDev} [r (254) = 0.15, $p = 0.016$] and when controlling for standard bias scores [r (253) = 0.14, $p = 0.030$] (see Table 5). No other significant associations with response-based measures and anxiety were noted (see Table 4). Similarly, no significant relationships were found between relative magnitude of response-based measures and either GAD symptoms or PAPA-Fear scores (all $ps > 0.09$). Anxiety measures were not associated with happy response-based measures (all $ps > 0.17$, see Table S3).

Discussion

This study demonstrated support for the use of response-based computation approaches for the preschool dot probe over standard approaches. The response-based approach, which compares trial-level reaction times to mean reference reaction times, produced attention measures that exhibited comparatively superior internal consistency and detected an anxiety-related association not observed using standard computation measures. These findings confirm work done with adult populations (Evans & Britton, 2018; Meissel et al., 2021) and support the use of response-based computation approaches across developmental periods.

Internal Consistency

As identified in prior work with adults (Evans & Britton, 2018), standard computation approaches were characterized by unacceptable levels of internal consistency, whereas response-based approaches exhibited strong internal consistency. Specifically, standard measures failed to reach acceptable internal consistency regardless of the number of trials considered, whereas response-based measures required relatively few trials to achieve

Table 3 Bivariate correlations among anxiety measures and standard attention measures

	n	1	2	3	4	5
1. Anx-DOS	260	–				
2. # GAD Sx	259	−0.01	–			
3. PAPA fear	264	−0.01	0.54***	–		
4. Attention bias	276	0.05	0.06	−0.04	–	
5. Orientation	276	0.07	0.06	−0.02	0.43***	–
6. Disengagement	276	−0.02	−0.01	−0.02	0.53***	−0.60***

***$p < 0.001$

Table 4 Partial correlations among anxiety measures and response-based attention measures (n = 280)

	1	2	3	4	5	6	7	8
1. Anx-DOS	–							
2. # GAD Sx	–.01	–						
3. PAPA Fear	−0.01	0.54***	–					
4. Anger—Vigilant Bias	0.04	0.08	−0.03	–				
5. Anger—Avoidant Bias	−0.06	−0.03	0.06	−0.69***	—			
6. Anger—Vigilant Orientation	0.04	0.08	−0.02	0.20*	−0.26***	–		
7. Anger—Avoidant Orientation	−0.08	−0.04	0.03	−0.20**	0.65***	−0.66***	–	
8. Anger—Slow Disengagement	0.01	−0.08	−0.09	0.67***	−0.40***	−0.48***	0.32***	–
9. Anger—Fast Disengagement	0.03	−0.05	−0.05	−0.43***	0.21***	0.69***	−0.50***	−0.67***

Controlling for RT_{Mean} and RT_{StdDev}

**$p < 0.01$, ***$p < 0.001$

Table 5 Partial correlations among anxiety measures and response-based magnitudes (n = 280)

	1	2	3	4	5
1. Anx-DOS	–				
2. # GAD Sx	−0.01	–			
3. PAPA Fear	−0.01	0.54***	–		
4. Vigilant:Avoidant Bias	0.12+	0.04	−0.09	–	
5. Vigilant:Avoidant Orientation	0.08	0.07	−0.02	0.45***	–
6. Fast:Slow Disengagement	0.03	−0.13	−0.11	0.24**	−0.23*

Controlling for RT_{Mean} and RT_{StdDev}

****p* < 0.05, ***p* < 0.01, ****p* < 0.001

acceptable internal consistency. For example, avoidant bias reached acceptable levels of internal consistency at only 12 trials and vigilant bias reached acceptable levels of internal consistency at only 6 trials. This is a particularly important finding in considering the use of the preschool dot probe task with young children, as prior work has found some challenges in implementing 180-trial versions of the dot probe task with young preschool children due to children's attention span (e.g., aged 3; Loomis, 2020), and may provide support for an abbreviated version of the dot probe task.

The improved psychometric properties of response-based computation measures are consistent with variable attention states within an individual. Specifically, response-based measures may achieve superior psychometric properties by separately assessing distinct components of attention (e.g., vigilant orientation and avoidant orientation), rather than combining these components of attention using standard computation approaches. Thus, it is possible that the poor psychometric properties of standard computation measures may mask anxiety-related associations, whereas the superior psychometric properties of response-based measures approach may facilitate detection of anxiety-related associations.

Criterion Validity

In the current study, we found modest associations between observed anxiety and response-based measures, which were not observed using standard computation approaches. Among relative magnitude ratio measures, greater vigilant bias *relative* to avoidant bias (i.e., Vigilant Bias > Avoidant Bias) was associated with higher Anx-DOS scores ($p = 0.051$), where children with greater ratio of vigilant bias relative to avoidant bias (i.e., vigilant bias > avoidant bias) exhibited more *observed* anxious behaviors. Although this association was small in magnitude, it highlights relative ratio/magnitude variables as bring more robust against potential confounds like general RT variability in detecting threat-related attention bias (Evans & Britton, 2018). Of note, the relationship between response-based measures of attentional bias and anxiety measures were unique to angry faces and were not observed to happy faces (see Tables S2–S4). Thus, anxiety-related associations did not reflect attention in response to emotional stimuli more generally, but were instead somewhat specific to threat. These exploratory findings suggest that response-based computation approaches may help to uncover trends related to threat-related stimuli that are missed from standard approaches to computation. Research in the future should explore whether and how response-based approaches can provide more precision for those wishing to distinguish specific relationships to threat-related stimuli. Future research should continue to incorporate multi-method observations of anxiety to parse the potentially distinct roles of threat-related attention in parent-reported and observed anxiety symptomatology.

Even with the better psychometric properties resulting from the response-based computing, anxiety related associations were still weak in magnitude. It may be that a single time point of measuring attention bias (even with trial-based measures) is insufficient, as intra-individual variability in attention bias is likely to also shift over periods of days. In Evans and Britton (2018), response-based computations demonstrated strong internal consistency, but the test–retest reliability was still essentially zero over a 1-week period. This suggests that attention bias may fluctuate on both a moment-to-moment basis as on a longer time-scale across days/weeks. Thus, it is possible that anxiety symptoms simply do not strongly correlate with a "static" (i.e., cross-sectional measure from a single session) measure of attention bias. This is particularly meaningful given findings that young children's behavior fluctuates substantially over time (Wakschlag et al., 2015). Multiple time points of assessment of both anxiety/behavior and attention bias patters are likely optimal for understanding these relationships within the developmental context of early childhood.

In addition to examining different ways to compute attention bias scores from raw data, researchers have also attempted to address the issues in psychometric reliability of the dot probe task through adaptations to the approach itself. For example, Heitmann et al. (2021) compared psychometric properties of attentional bias measures in relation to alcohol use when collected in a bar (alcohol relevant) context compared a traditional lab-based implementation. Another adaptation introduced a task-related prompt along with the more traditional picture prompt, allowing researchers to experimentally control for task-related bias, which resulted in improved reliability and associations with anxiety (Gladwin et al., 2021). These novel approaches should be examined in future efforts to improve the psychometric properties in the preschool dot probe; for example, by comparing implementation in a home-based setting where past violence may have taken place to implementation in a traditional lab-based setting.

Threat-related attention may also only capture implicit threat processes in isolation, which work additively/interactively with explicit threat processes in anxiety disorders. Robust anxiety-related differences may only be seen when considering the interaction between implicit and explicit threat processes. It may be that a multivariate approach in which orientation and disengagement are considered together would produce stronger anxiety-related associations. For example, higher anxiety may be associated with a *multivariate* combination of high vigilance, moderate avoidance, high slow disengagement, and low fast disengagement. Thus, it is possible that univariate analytic approaches (even if psychometrically sound) may not comprehensively capture an individual's "profile" of threat-related attention.

Limitations and Future Directions

This study is the first to use the novel response-based computation approach in a preschool sample, demonstrating improved psychometric properties consistent with work in adult samples. Despite notable contributions, there are some limitations that should be considered. Prior work found no evidence of test–retest reliability using standard computation approaches, but varying evidence using response-based measures (albeit none reaching acceptable levels; Evans & Britton, 2018). We were not able to examine test–retest reliability with the current study, but future work may examine whether findings related to test–retest reliability differs among younger samples due to developmental changes. Additionally, we did not employ mixed models for analyzing data which essentially treats the stimuli (e.g., the faces in the task) as random and does not consider how attentional bias may be evoked differently based on the stimuli (e.g., angry pictures of males versus females), increasing the risk of Type 1 errors and likely contributing to inconsistent results (Judd et al., 2012). Future work with the preschool dot probe that integrate mixed models, regardless of the computation approach used, would elucidate factors related to variance in attentional biases among young children that traditional analytic approaches ignore.

It should also be noted that our finding differs somewhat from other work with preschoolers, which found cross-sectional links between anger bias using the standard scoring approach to the dot probe and observed fear/anxiety using the Anx-DOS (Briggs-Gowan et al., 2015; Mian et al., 2015). However, these studies differed slightly in the age of preschoolers included (e.g., age 3 and up versus age 4 and up), the use of the extended (360 trial) vs. standard (180 trial) version of the dot probe, and covariates (e.g., prior work controlled for age, sex, race/ethnicity, and non-verbal reasoning), which may explain differences in the results related to the standard scoring approach. Of note, the current paper contributes to the current literature by demonstrating that response-based scoring can capture links between attention bias and parent-reported anxiety, which was not identified in earlier work using standard approaches. These differences also reflect a need for more work identifying anxiety-related associations using both scoring approaches and with different preschool populations.

Additionally, despite the improved psychometric properties using the response-based scoring approach, relations between attention bias and observed/parent-report measures of anxiety were still not robust. Further, we did not apply a correction method thus these results should be interpreted as exploratory and would benefit replication with larger samples of young children. Response-based approaches are not a panacea for addressing inconsistent anxiety-related associations across studies. The response-based approach offers some information about intra-individual variability during the timing of the lab session, however we did not examine intra-individual variability across longer periods of time, such as multiple days or weeks, which would provide more nuanced understanding of fluctuations in attention bias to threat. Since there may be developmental differences in how attention bias to threat unfolds across time, particularly in early childhood, it is important to compare response-based scoring approaches to anxiety and threat-related symptoms across time in future longitudinal studies.

Conclusion

Dot probe tasks have long been critiqued for inconsistent findings putatively related to poor psychometric properties, which may give rise to inconsistent results across studies. Prior work with adult populations has found trial-level computation approaches to dot probe measures produces superior psychometric properties compared to standard scoring approaches. The current study advances the field by confirming the superiority of response-based scoring approaches with younger samples by comparing internal consistency and criterion validity using both response-based and standard scoring approaches. The response-based computation approach demonstrated superior psychometric properties and lower trial numbers compared to standard approaches and may improve the detection of anxiety-related associated in early childhood samples.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10608-023-10409-1.

Funding This study includes secondary analysis of previously funded research. The initial study was funded by the National Institute of Mental Health (MH090301, Principal Investigator [PI]: Margaret J. Briggs-Gowan; MH082830, PI: Lauren S. Wakschlag).

Data Availability Data for this study is not publicly available.

Declarations

Conflict of interest Alysse M. Loomis, Travis C. Evans, Damion J. Grasso, Margaret Briggs-Gowan declare that they have no conflict of interest. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Ethical Approval All study protocols were approved by the relevant institutional review boards.

Informed Consent Informed consent was obtained from all individual participants included in the initial study, from which secondary data analysis was conducted.

Animal Rights No animal studies were carried out by the authors for this article.

References

Aktar, E., Van Bockstaele, B., Pérez-Edgar, K., Wiers, R. W., & Bögels, S. M. (2019). Intergenerational transmission of attentional bias and anxiety. *Developmental Science, 22*(3), 1–19. https://doi.org/10.1111/desc.12772

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision). Washington, DC.

Bar-Haim, Y. (2010). Research review: Attention bias modification (ABM): A novel treatment for anxiety disorders. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 51*(8), 859–870. https://doi.org/10.1111/j.1469-7610.2010.02251.x

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*(1), 1–24. https://doi.org/10.1037/0033-2909.133.1.1

Briggs-Gowan, M. J., Pollak, S. D., Grasso, D., Voss, J., Mian, N. D., Zobel, E., McCarthy, K. J., Wakschlag, L. S., & Pine, D. S. (2015). Attention bias and anxiety in young children exposed to family violence. *Journal of Child Psychology and Psychiatry, 56*(11), 1194–1201. https://doi.org/10.1111/jcpp.12397

Briggs-Gowan, M. J., Grasso, D., Bar-Haim, Y., Voss, J., McCarthy, K. J., Pine, D. S., & Wakschlag, L. S. (2016). Attention bias in the developmental unfolding of post-traumatic stress symptoms in young children at risk. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 57*(9), 1083–1091. https://doi.org/10.1111/jcpp.12577

Cremone, A., Kurdziel, L. B. F., Fraticelli-Torres, A., McDermott, J. M., & Spencer, R. M. C. (2017). Napping reduces emotional attention bias during early childhood. *Developmental Science, 20*(4), 1–12. https://doi.org/10.1111/desc.12411

Dudeney, J., Sharpe, L., & Hunt, C. (2015). Attentional bias towards threatening stimuli in children with anxiety: A meta-analysis. *Clinical Psychology Review, 40*, 66–75. https://doi.org/10.1016/j.cpr.2015.05.007

Egger, H. L., & Angold, A. (2004). The Preschool Age Psychiatric Assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children. In: *Handbook of infant, toddler, and preschool mental health assessment* (Vol. 42, no. 5, pp. 223–243).

Evans, T. C., & Britton, J. C. (2018). Improving the psychometric properties of dot-probe attention measures using response-based computation. *Journal of Behavior Therapy and Experimental Psychiatry, 60*, 95–103. https://doi.org/10.1016/j.jbtep.2018.01.009

Gade, P. R., Molloy, A., & Anderson, P. L. (2021). The use of trial-level bias scores to examine attention bias and attention bias variability among people with and without social anxiety disorder. *Journal of Clinical Psychology*. https://doi.org/10.1002/jclp.23259

Gladwin, T. E., Halls, M., & Vink, M. (2021). Experimental control of conflict in a predictive visual probe task: Highly reliable bias scores related to anxiety. *Acta Psychologica, 218*, 103357.

Grasso, D. J., Petitclerc, A., Henry, D. B., McCarthy, K. J., Wakschlag, L. S., & Briggs-Gowan, M. J. (2016). Examining patterns of exposure to family violence in preschool children: A latent class approach. *Journal of Traumatic Stress, 29*(6), 491–499.

Heitmann, J., Jonker, N. C., & De Jong, P. J. (2021). A promising candidate to reliably index attentional bias toward alcohol cues–an adapted odd-one-out visual search task. *Frontiers in Psychology, 12*, 630461.

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*(1), 54.

Kruijt, A. W., Field, A. P., & Fox, E. (2016). Capturing dynamics of biased attention: Are new attention variability measures the way forward. *PLoS ONE, 11*(11), 1–22. https://doi.org/10.1371/journal.pone.0166600

Kruijt, A. W., Parsons, S., & Fox, E. (2019). A meta-analysis of bias at baseline in RCTs of attention bias modification: No evidence for dot-probe bias towards threat in clinical anxiety and PTSD. *Journal of Abnormal Psychology, 128*(6), 563–573. https://doi.org/10.1037/abn0000406

Loomis, A. M. (2020). Effects of household and environmental adversity on indices of self-regulation for Latino and African American preschool children: Closing the school readiness gap. *Early Education and Development*. https://doi.org/10.1080/10409289.2020.1745513

Meissel, E. E. E., Liu, H., Stevens, E. S., Evans, T. C., Britton, J. C., Letkiewicz, A. M., & Shankman, S. A. (2021). The reliability and validity of response-based measures of attention bias. *Cognitive Therapy and Research*. https://doi.org/10.1007/s10608-021-10212-w

Mian, N. D., Carter, A. S., Pine, D. S., Wakschlag, L. S., & Briggs-Gowan, M. J. (2015). Development of a novel observational measure for anxiety in young children: The Anxiety Dimensional Observation Scale. *Journal of Child Psychology and Psychiatry* *and Allied Disciplines, 56*(9), 1017–1025. https://doi.org/10.1111/jcpp.12407

Miller, J., & Ulrich, R. (2013). Mental chronometry and individual differences: Modeling reliabilities and correlations of reaction time means and effect sizes. *Psychonomic Bulletin & Review, 20*(5), 819–858.

Parsons, S. (2020). Splithalf: Robust estimates of split half reliability. figshare. Software. https://doi.org/10.6084/m9.figshare.11956746.v4

Parsons, S., Kruijt, A. W., & Fox, E. (2019). Psychological science needs a standard practice of reporting the reliability of cognitive-behavioral measurements. *Advances in Methods and Practices in Psychological Science, 2*(4), 378–395.

Pine, D. S., Mogg, K., Bradley, B. P., Montgomery, L. A., Monk, C. S., McClure, E., Guyer, A. E., Ernst, M., Charney, D. S., & Kaufman, J. (2005). Attention bias to threat in maltreated children: Implications for vulnerability to stress-related psychopathology. *American Journal of Psychiatry, 162*(2), 291–296. https://doi.org/10.1176/appi.ajp.162.2.291

Schmukle, S. C. (2005). Unreliability of the dot probe task. *European Journal of Personality, 19*(7), 595–605. https://doi.org/10.1002/per.554

Susa, G., Pitică, I., Benga, O., & Miclea, M. (2012). Anxiety-related attention biases in preschoolers: An investigation using the pictorial dot-probe task. *Procedia—Social and Behavioral Sciences, 33*, 637–641. https://doi.org/10.1016/j.sbspro.2012.01.199

Ursache, A., & Blair, C. (2015). Children's cortisol and salivary alpha-amylase interact to predict attention bias to threatening stimuli. *Physiology and Behavior, 138*, 266–272. https://doi.org/10.1016/j.physbeh.2014.10.002

Wakschlag, L. S., Briggs-Gowan, M. J., Choi, S. W., Nichols, S. R., Kestler, J., Burns, J. L., Carter, A. S., & Henry, D. (2014). Advancing a multidimensional, developmental spectrum approach to preschool disruptive behavior. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*, 82–96.

Wakschlag, L. S., Estabrook, R., Petitclerc, A., Henry, D., Burns, J. L., Perlman, S. B., Voss, J. L., Pine, D. S., Leibenluft, E., & Briggs-Gowan, M. L. (2015). Clinical implications of a dimensional approach: The normal:abnormal spectrum of early irritability. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(8), 626–634.

Zvielli, A., Bernstein, A., & Koster, E. H. W. (2015). Temporal dynamics of attentional bias. *Clinical Psychological Science, 3*(5), 772–788. https://doi.org/10.1177/2167702614551572