Cultural Differences in Performance on Eriksen's Flanker Task

Cultural Differences in Performance on Eriksen's Flanker Task

Angela Gutchess
Brandeis University

John Ksander
Brandeis University

Peter R. Millar
Washington University
Brandeis University

Berna A. Uzundag
Kadir Has University
Bogazici University

Robert Sekuler
Brandeis University

Aysecan Boduroglu
Bogazici University

Contact information:
Dr. Angela Gutchess
Department of Psychology
Brandeis University
415 South Street, MS 062
Waltham, MA 02453
e-mail: gutchess@brandeis.edu
phone: (781)736-3247

Acknowledgements

Abstract

Eriksen's zoom model of attention implies a tradeoff between the breadth and resolution of representations of information. Following this perspective, we used Eriksen's flanker task to investigate culture's influence on attentional allocation and attentional resolution. In Experiment 1, the spatial distance of the flankers was varied, in order to test whether people from Eastern cultures (here, Turks) experienced more interference than people from Western cultures (here, Americans) when flankers were further from the target. In Experiment 2, the contrast of the flankers was varied. Although congruent and incongruent flankers (Experiment 1) as well as the degree of contrast of the flankers compared to the target (Experiment 2) interact with participants' cultural background to differentially influence accuracy or reaction times, In addition, we used evidence accumulation modeling to jointly consider measures of speed and accuracy. Results indicate that in order to make decisions in the Eriksen flanker task, Turks both accumulate evidence faster and require more evidence than Americans. These cultural differences in visual attention and decision making have implications for a wide variety of cognitive processes.

Keywords: culture, cognition, cross-cultural, visual interference, attention, flanker

Cultural Differences in Performance on Eriksen's Flanker Task

Culture is a multifaceted construct, encompassing the shared norms and mindsets of affiliated individuals ranging from small groups (e.g., one's family, the local Rotary Club) to nations or beyond (e.g., the former British colonies). A culture's shared perspective can impact myriad behaviors. For example, in the United States, one Northeastern city's sports teams are widely thought, perhaps unfairly, to have a culture of bending the rules or downright cheating, but that perspective is not shared by residents of that city. At the start of his long, illustrious career, Charles W. Eriksen used the term "culture" in a study that examined how personality variables (authoritarianism and neuroticism) reflected the sociocultural settings from which subjects were drawn, namely typical undergraduate institutions or U.S. naval training stations (Davids & Eriksen, 1957). Rather than following Davids & Eriksen's (1957) groups, the present study considers participants from different countries. Here, we use the term "culture" as Gutchess & Sekuler (2019) did, to signify groupings defined by a person's country of origin, and where and when s/he has spent the majority of his/her life. Defined this way, "culture" is a proxy for the multifaceted ways that an environment can sculpt a human brain and therefore shape the individual's cognitive world, how one views and interacts with the world. That perspective makes culture an interesting, useful portal into perception and attention. The trajectory of Charles Eriksen's career and contributions make us think he would have approved.

Throughout his career, Eriksen recognized the capacity limitations of visual attention, while also considering the precision, or resolution, of the mental representations of the information that results from attention's selectivity (B. A. Eriksen & Eriksen, 1974; C. W. Eriksen & St James, 1986; C. W. Eriksen & Yeh, 1985). His zoom model proposed that attention could be redistributed to have broader scope, but at the expense of reduced resolution or less efficient processing (C. W. Eriksen & St James, 1986; C. W. Eriksen & Yeh, 1985; see also Huang, 2010; Suchow, Fougnie, Brady, & Alvarez, 2014). Part of the work we report extends the ideas of Eriksen's zoom model.

Though much of Eriksen's career was devoted to understanding attentional processes, as we mentioned above, his early career investigated personality and sociocultural influences (Davids & Eriksen, 1957; C. W. Eriksen, 1954, 1957). The work presented in this manuscript occupies the intersection of these topics, investigating how attentional processes can be shaped by culture, which, like personality, represents relatively stable individual differences.

The question of whether cultural background can impact the resolution with which the visual environment is represented has gained substantial interest in recent years. Some studies have argued that culture can determine how broadly visual attention is allocated to the environment, thereby impacting the precision of representations (Boduroglu & Shah, 2017; Boduroglu, Shah, & Nisbett, 2009; Hakim, Simons, Zhao, & Wan, 2017; Lawrence, Edwards, Chan, Cox, & Goodhew, 2019; Lawrence, Edwards, Talipski, & Goodhew, 2020). For instance, Boduroglu & Shah (2017) demonstrated that East Asians performed not only more poorly than Americans on the functional field of view task, but, compared to Americans, their errors were less likely to be due to selecting a neighbor of the target. These findings suggest that East Asians distribute attention more broadly over space than Americans, but at the expense of representational precision. Others have argued that cultural differences in attentional processes may emerge because visual environments vary in their degree of clutter (e.g., Miyamoto, Nisbett, & Masuda, 2006), with lower levels of clutter fostering higher attentional selectivity and greater spatial focus, resulting in cultural differences in visual search and global-local bias (Cramer, Dusko, & Rensink, 2016; de Fockert, Caparos, Linnell, & Davidoff, 2011; Linnell & Caparos, 2011; Ueda et al., 2018). These studies demonstrate that research on cultural differences in visual attention has shifted away from an explanation based on cultural differences in holistic versus analytic cognitive style (e.g., McKone et al., 2010), toward more process-based accounts. However, most of these studies focused on demonstrating differences across cultures, and offering explanations that were post-hoc -- few attempted to identify when

and how these differences emerged (for an exception see Freeman, Ma, Han, & Ambady, 2013).

The goal of this study was to examine cultural differences in attentional allocation and attentional resolution, as well as the temporal dynamics of visual information accumulation that precede decisions. With this goal in mind, we tested Easterners and Westerners on a version of Eriksen's flanker task (B. A. Eriksen & Eriksen, 1974). In this task, participants respond based on the identity of a central target (e.g., pressing one key for an "E", another key for an "H"). The central target is flanked by stimuli to each side that are either compatible (i.e., same identity) or incompatible, which creates competition. This arrangement sets up a competition that increases reaction times compared to when the flankers are compatible with the target. We selected the flanker task because it been shown to support an integrated approach to diverse influences on visual processing, including influences identified in the voluminous literature on aspects of visual crowding (Levi, 2008). For example, visual crowding, which is produced by lateral interactions between adjacent stimuli, affects visual processing in real-world tasks (Rosenholtz, Yu, & Keshvari, 2019), limits reading speed (Pelli et al., 2007), is sensitive to time pressure in theoretically revealing ways (Dayan & Solomon, 2010), and reflects important consequences of abnormal visual experience (Farzin & Norcia, 2011). Importantly, although visual crowding is often described as a phenomenon of peripheral vision, there are good reasons to treat it, and results from Eriksen's flanker task, as a general characteristic of vision, including foveal vision (Coates, Levi, Touch, & Sabesan, 2018; Strasburger, 2020). Eriksen's flanker task offered several key advantages over other tasks that have been used to investigate visual interference across cultures, such as tests that used Navon figures (McKone et al., 2010; see Dale & Arnell, 2013). Specifically, Eriksen's flanker task has considerable sensitivity to well-understand attentional and visual phenomena, some of which were mentioned above. In particular, the task allowed us to assess cultural differences in interference in early visual attention (C. W. Eriksen & Yeh, 1985; Gaspelin, Ruthruff, & Jung, 2014), Additionally, the task's structure facilitated

parametric variation of multiple key variables, including the effects of distance separating the

flankers from the central target (e.g., B. A. Eriksen & Eriksen, 1974; C. W. Eriksen & St James,

1986; Miller, 1991).

Although some previous research has considered culture's influence on flanker

interference effects, that research lacked direct comparisons of Eastern and Western adults.

Rather, those studies investigated the cultural influences of a semi-nomadic lifestyle (de Fockert

et al., 2011), the contribution of independence/interdependence (Lin & Han, 2009) or analytic-

holistic processing (Hsieh, Yu, Chen, Yang, & Wang, 2020) adopting an individual differences

approach, or development in childhood with a focus on social stimuli (Senzaki, Wiebe, Masuda,

& Shimizu, 2018).

In Experiment 1, we manipulated the distance between targets and flankers to test

whether cultural groups differed in sensitivity to visual interference across different spatial

scales. In addition to carrying out inferential statistics-based comparisons across conditions, we

also investigated cultural differences in the speed-accuracy trade-off, a phenomenon with

demonstrated value for theories of visual processing (Heitz, 2014). Specifically, for each

condition and across culture groups, we determined the proportion of accurate responses, using

d', across different time bins of the RT distribution. Finally, we also used a modeling approach

that provides insight into information accumulation and decision processes. Evidence

accumulation models conceptualize decision-making as a noisy process in which people

accumulate evidence until some criterion is reached. Thus, reaction time is a function of both

the criterion (i.e., evidence threshold) and the rate at which individuals accumulate information

(i.e., the drift rate), as well as the non-decision time, which reflects the time to perceive and

respond to the stimulus. For instance, results from evidence accumulation modeling of the

flanker task favors single-process spotlight models of attention (B. A. Eriksen & Eriksen, 1974;

C. W. Eriksen & St James, 1986) over dual-process models (White, Ratcliff, & Starns, 2011), as

well as indicating that people narrow attention gradually rather than abruptly. In our study,

evidence accumulation modeling examined cultural differences in the task's component processes, including the quality of the information and the relative prioritization of speed versus accuracy (Ratcliff & McKoon, 2008).

Although we expected to replicate some previous results, namely that response to incongruent stimuli would be relatively slowed and that interference would be relatively greater with spatially proximate stimulus elements, we expected that cultural background would moderate both of these effects. If Easterners attend more broadly than Westerners, one would expect Easterners' target detection accuracy to be worse, particularly in the far condition. This might be driven by one of two factors. First, broader allocation of attention may make Easterners more prone to interference from far as opposed to near flankers; Westerners may be better at focusing on the centrally placed target. Second, attentional allocation could reduce representational precision. Should this be the case, Easterners would be poorer in identifying the target amongst the flankers and/or take longer to reach the criterion supporting the same level of accuracy. We would expect these types of trade-offs to impact performance, especially in these types of fast-pace detection tasks. Adopting an evidence accumulation modeling approach allows us to explore these possibilities. We compared Americans as our Western sample to Turks as our Eastern sample, because Turkey was shaped by a combination of Eastern and Western historical influences and is more Eastern in style of thought (Henrich, Heine, & Norenzayan, 2010; Schwartz, Boduroglu, & Gutchess, 2014; Uskul, Kitayama, & Nisbett, 2008).

EXPERIMENT 1

Methods

Participants

We tested 35 American undergraduates at Brandeis University, Waltham, USA, and 41 Turkish undergraduates at Bogazici University, Istanbul, Turkey. Sample sizes were selected to

exceed 30 per group, a value based on prior work and on G*Power calculations for a priori

power for a repeated measures within-between design (α = .05, and power = .95) with a

medium effect size (Cohen's d = .5). Participants at both locations were native to their

respective country of testing and had lived no more than 5 years abroad. Data from an

additional 19 participants (11 Turks, 8 Americans) were excluded; nine because their

demographics questionnaire indicated they were not eligible (e.g., lived out of the country;

exposed to both Eastern and Western cultures), seven because their math accuracy score did

not meet the cut-off for inclusion on the Operation Span task (Unsworth, Heitz, Schrock, &

Engle, 2005), and three because lighting conditions were not properly controlled. All participants

gave written consent and were reimbursed with either course credit or cash payment.

Materials

Stimuli were 300 Eriksen-type flanker images, each comprising either one or five letters,

capital *E*'s and/or *H*'s (B. A. Eriksen & Eriksen, 1974). All letters were shown in black block font

text (luminance = 0.8 cd/m^2) against a grey background (luminance = 43.6 cd/m^2). One hundred

of the stimuli were unflanked controls, consisting of a single letter subtending a 1.0* visual

angle, presented at the center of the screen. These served as a baseline condition for

assessing baseline cultural differences in reaction times. One hundred other stimuli were closely

flanked, consisting of five letters, each separated from its nearest neighbor by a visual angle of

0.5*, which made the set of span 7.0* from end to end. Of these closely flanked images, 50

were congruent, with all five letters matching, and 50 were incongruent, with the central letter

mismatched to the four flanker letters, for example four E's flanking a central H, or vice versa.

Finally, one hundred stimuli were widely flanked, consisting of five letters, each separated by a

visual angle of 1.5*, spanning 11.0* from end to end. Of these closely flanked images, 50 were

congruent and 50 were incongruent. Figure 1 provides some exemplars and an illustration of

task timing. The numbers of *E*'s and *H*'s in each position were equally matched in each

condition, and trials were randomly intermixed across conditions. Apparatus and experimental

set-ups (e.g., lighting levels) were carefully matched across sites, with stimuli presented using E-Prime 1.2 software (PSTNet, Pittsburgh, PA) on a Dell Optiplex 380 desktop PC with a Dell 2011H monitor positioned 1 m in front of the subject.

Procedure

Participants in both the US and in Turkey were tested in their respective native language and in their country of residence. Participants in each site followed identical procedures. After supplying informed consent and basic demographic information, participants completed 40 trials of a simple reaction time test, in which they pressed a key on the keyboard as quickly as possible in response to stimulus onset (Deary, Liewald, & Nissan, 2011).

Participants then completed the flanker task. Each trial of the task proceeded in the following manner: Participants focused on a black fixation point, positioned at the center of the screen, for 1400 ms. Then for 100 ms, the fixation point turned red, to serve as an attentional warning. Then one of the five stimuli types appeared at the center for 150 ms, immediately followed by a visual mask for 200 ms. In the final 2000 ms, an instruction screen prompted participants to press one key if they saw an *E* in the central target position or another key if they saw an *H*. Following each response, audio feedback was administered: a high-pitched beep signaled a correct response, while a low-pitched buzz signaled an incorrect response. Trials were presented in a random order of conditions.

Before doing the flanker task, participants first completed 30 trials, with the different conditions intermixed, as a practice. A threshold of 80% accuracy during this practice block was required before beginning the primary experimental task. Participants repeated the practice block up to three times, if needed, to reach that threshold.

After the flanker task, participants completed a choice reaction time task comprising 40 trials, in which they were instructed to press one of four keys whose spatial arrangement corresponded to the location in which a cross appeared on the display (Deary et al., 2011). Participants then completed an autobiographical memory task that lasted approximately 30

minutes. This memory task was designed for a separate study and will not be discussed in this paper.

Participants then completed a battery of neurocognitive measures and questionnaires. These measures included tests of processing speed (Digit Comparison; Hedden & Park, 2001) and working memory capacity (Automated Operation Span; Unsworth et al., 2005).

Experiment 1 Results

Participant Characteristics

Table 1a summarizes the participant characteristics[1]. Turkish participants were significantly older than American, $t(71) = 4.48$, $p < .01$, but the two groups did not significantly differ in years of formal education, $t(68) = p > .25$. These patterns reflect the tendency for our Turkish participants to have completed an additional year of preparatory school in English.

In terms of performance on the simple and choice reaction time tasks, American participants were faster than Turks in responding to both the simple, $t(74) = 4.79$, $p < .001$, and choice reaction time tests, $t(73) = 5.23$, $p < .001$. Supplementary Figure 1 shows the overlap across the two cultures in simple reaction times. The two groups did not significantly differ in processing speed, as measured by the digit comparison task, $t(73) = .65$, $p > .50$, or working memory, as measured by the automated operation span task, $t(67) = 1.71$, $p > .09$. The equivalence across cultures on these measures suggest that our samples are well-matched on these measures of cognitive ability.

Flanker Accuracy

A 2 x 2 x 2 ANOVA was conducted on accuracy with Culture (American, Turkish) as a between-subjects factor and Flanker Congruency (congruent, incongruent) and Flanker Distance (near, far) as within-subject factors. Results are shown in Figure 2, including performance on the unflanked control condition.

---

[1] The different degrees of freedom across these measures reflect measures that were not completed by some participants.

Participants were more accurate for congruent (M = .97, SD = .03) than incongruent trials (M = .92, SD = .04), $F(1, 74)$ = 92.63, p < .001, partial η^2 = .56. Accuracy was also higher for the far (M = .95, SD = .03) than the near (M = .93, SD = .05) trials, $F(1, 74)$ = 24.60, p < .001, partial η^2 = .25. The significant interaction of Congruency and Distance, $F(1, 74)$ = 18.01, p < .001, partial η^2 = .20, indicated higher levels of accuracy for the far (M = .94, SD = .04) than near (M = .90, SD = .08) incongruent trials, $t(75)$ = 4.84, p < .001, Cohen's d = .55, but no difference in accuracy between near and far for congruent trials (Ms = .97, SDs = .03), $t(75)$ = .97, p = .34, Cohen's d = .11.

Critically, the analyses considering culture revealed a significant main effect of Culture, $F(1, 74)$ = 11.65, p < .001, partial η^2 = .14, with Turks (M = .96, SD = .04) exhibiting greater accuracy than Americans (M = .93, SD = .04). This overall cultural difference in accuracy is consistent with performance on the unflanked control trials, for which the Turks performed significantly more accurately than the Americans, $t(74)$ = 4.49, p <.001, Cohen's d = 1.02. There was a significant interaction of Culture and Flanker Congruency, $F(1, 74)$ = 6.53, p < .02, partial η^2 = .08. Although the effect of incongruency was significant in both groups, the magnitude differed across groups. Americans' performance was disproportionately poorer on incongruent (M = .90, SD = .05) compared to congruent trials (M = .96, SD = .03), $F(1, 34)$ = 55.68, p < .001, partial η^2 = .62, relative to a smaller difference for Turks (Ms = .94 and .98; SDs = .06 and .03), $F(1, 40)$ = 33.89, p < .001, partial η^2 = .46. Neither the interaction between Culture and Distance, nor the three-way interaction of Culture, Distance, and Congruency reached significance, ($F(1, 74)$ = 3.00, p=.087, partial η^2 = .039; and $F(1, 74)$ = 3.57, p=.063, partial η^2 = .046, respectively).

Flanker Reaction Time

Outliers were trimmed such that trials above and below 2.5 standard deviations from each participant's mean, calculated across all conditions, were eliminated from analyses[2]. Any trials that were responded to faster than 100 ms were excluded. In order to adjust for skew in the data, analyses of reaction time data were conducted using the median of each participant's reaction times on correct trials,

A 2 x 2 x 2 ANOVA was conducted on median reaction times with Culture (American, Turkish) as a between-subjects factor and Flanker Congruency (congruent, incongruent) and Flanker Distance (near, far) as within-subject factors. Results are shown in Figure 3. As expected there was a significant main effect of Flanker Congruency, $F(1, 74) = 163.78$, $p < .001$, partial $\eta^2 = .69$, with faster reaction times on congruent trials ($M = 427.77$ ms, $SD = 49.79$) than on incongruent trials ($M = 459.36$ ms, $SD = 51.29$). The main effect of Flanker Distance was also significant, $F(1, 74) = 66.95$, $p < .001$, partial $\eta^2 = .48$, with faster reaction times on far-flanked trials ($M = 436.62$, $SD = 48.43$) than on near-flanked trials ($M = 450.51$, $SD = 51.39$). The interaction of Flanker Congruency and Flanker Distance was also significant, $F(1, 74) = 34.74$, $p < .001$, partial $\eta^2 = .32$. The incongruency effect was larger for near flankers (41 ms) than for far flankers (23 ms), although the effect was significant for both near and far trials: incongruent $M = 471.21$ ($SD = 54.12$) versus congruent $M = 430.49$ ($SD = 51.99$); $t(75) = 12.31$; Cohen's $d = 1.41$; and incongruent $M = 448.11$ ($SD = 50.32$) versus congruent $M = 425.84$ ($SD = 48.37$); Cohen's $d = 1.04$, respectively.

Although we had predicted that Turks would experience disproportionately more interference for the far flankers than Americans, culture did not significantly interact with any other variable, including flanker distance (all ps > .25). In addition, the main effect of culture was not significant, $F(1, 74) = .61$, $p > .40$, partial $\eta^2 = .01$, with no significant difference in reaction

---

[2] When used in combination with median reaction time for each participant, trimming reaction times represents a conservative approach to eliminating outlier data. However, we note that per participant, an average of only 5.26 (SD = 2.11) trials were trimmed in Experiment 1.

time between Americans (*M* = 439.16, *SD* = 49.23) and Turks (*M* = 447.98, *SD* = 49.23). The

lack of overall culture differences during flanker judgments is consistent with the lack of cultural

differences in reaction time to the unflanked controls, $t(74) = .84$, $p = .40$, Cohen's $d = .19$.

Speed-Accuracy Curves

To integrate our measures of speed and accuracy, we examined accuracy as a function

of reaction time. Particularly because effects of culture differ for measures of accuracy and

reaction time, with culture influencing accuracy more than reaction time, it is informative to

consider measures of speed and accuracy together. We pooled subjects' trial data within each

group and ordered trials by RT. Next, successive 100ms bins were created (e.g. 0-99 ms, 100-

199 ms, etc.) so that each bin contained at least 50 trials. Measures of discrimination accuracy,

d', were calculated for each time bin and each cultural group. Speed-accuracy curves were then

created for each group by plotting d' scores against time (Figure 4), with the score for each bin

plotted at the bin's midpoint (e.g. 50 ms, 150 ms, etc.). In general, faster RTs were associated

with lower accuracy rates, a speed-accuracy tradeoff. A surprising cultural difference appeared

in the near incompatible condition, in which Turkish participants were more accurate than

American participants across all levels of reaction times.

Drift Diffusion Models

In order to better understand cultural differences in flanker performance, we modeled

subjects' decisions as drift diffusion processes. Broadly, drift diffusion models (DDM) describe

decisions as processes where subjects sequentially sample evidence (i.e. information) until the

evidence accumulates past some threshold and triggers a decision. Researchers have often

modeled simple decision tasks with DDMs and have successfully explained many empirical

phenomena (Ratcliff & McKoon, 2008). Supplementary Figure 2 illustrates the DDM decision

process, and how key parameters influence the model's behavior. Each panel shows ten

simulated decisions as gray lines. In each decision, an individual accumulates evidence until

reaching a decision threshold; the thresholds are indicated by black horizontal lines. The

decision process terminates once evidence exceeds a threshold, and those instances are marked by black circles.

The top panel of Supplementary Figure 2 illustrates a situation where most responses are correct; the subject accumulated evidence until reaching the upper boundary. However, noisy evidence accumulation delays some responses, and even causes an error response (marked by a circle on the lower boundary). In this way, the top panel gives an intuition for how the model produces errors and heavy-tailed RT distributions.

The drift rate parameter specifies how quickly subjects accumulate evidence. The middle panel shows how doubling the drift rate (compared to that used in the top panel) influences the decision process. In this middle panel, the higher drift rate both produces faster responses and increases response accuracy. The subjects are accumulating information quickly, so subjects reach decisions faster and with less susceptibility to noise-driven errors.

The evidence threshold parameter specifies how much evidence subjects require for a decision. The bottom panel shows how doubling the evidence threshold from the top panel influences the decision process. In this situation subjects are not accumulating evidence any faster (i.e., drift rate remains the same), but they require more evidence. This can be thought of in terms of prioritizing accuracy over speed. As such, the bottom panel shows how higher evidence thresholds result in fewer errors compared to the top panel, but the decisions take much longer. Several decision processes actually fail to terminate within the plotted 8 seconds because not enough evidence has accumulated.

We are primarily interested in the parameters illustrated in Supplementary Figure 2: drift rate and threshold. In DDM these parameters specify how quickly subjects accumulate evidence, and how much evidence they require for decisions. Fitting DDM models to our data can show whether cultural groups differ in drift rate or evidence thresholds when performing the flanker task.

This modeling was implemented with the Bayesian Hierarchical Drift Diffusion Model (HDDM) Python software package (Wiecki, Sofer, & Frank, 2013). We selected this implementation of drift diffusion because the hierarchical modeling strategy respects our data's nested structure; that is, by nesting the individual subjects' data under cultural group, the group distributions constrain individual subjects' parameter estimates. Additionally, Ratcliff & Childers (2015) indicate that for group analyses where individual subjects have relatively low trial-counts per condition, HDDM outperforms non-hierarchical implementations.

HDDM generates posterior probability distributions for several drift diffusion parameters via Markov-chain Monte Carlo (MCMC) sampling. We estimated model parameters by running 20 independent sampling chains. Each chain sampled 65,000 times; we discarded the first 60,000 samples, and, to reduce autocorrelation, we discarded every fifth remaining sample. The individual chains were concatenated into one 20k sample chain and visually inspected for convergence issues. We also examined Gelman-Rubin $\hat{R}$ statistics for both subject and group-level parameters, and subsequently excluded one Turkish subject for nonconvergence ($\hat{R} >$ 1.02. We fit all models to RT distributions including both correct and incorrect responses, specifying that 5% of datapoints belong to an outlier distribution[3]. Furthermore, we only modeled the data from Experiment 1's near incompatible condition. This condition showed cultural differences in accuracy, and both groups committed enough errors to support modeling.

Our primary model of interest specified two group-level distributions for non-decision time (t), drift rate (v) and evidence threshold (a) parameters. That is, we estimated separate group t, v, and a distributions for Americans and for Turks. The model converged well, and supplementary figures 3-5 show MCMC sampling distributions. Comparing the posterior distributions for group-level parameters revealed strong evidence that compared to Americans, Turks have a higher mean decision threshold $p(a_{TR} > a_{US}) = .979$, and higher mean drift rate

---

[3] For the justification for this modeling decision see:
http://ski.clps.brown.edu/hddm_docs/tutorial_python.html#dealing-with-outliers

$p(v_{TR} > v_{US}) = .994$. However, there was little evidence for cultural differences in mean non-decision time $p(t_{TR} > t_{US}) = .383$. Figure 5 shows the posterior probability densities for these group parameters.

In order to verify the stability of these parameter estimates, we ran two additional models: one without culture-specific drift rate distributions, and another without culture-specific threshold distributions. Both models otherwise matched the primary model's specifications exactly. For instance, in the first control model v did not vary between cultural groups, but a and t varied between groups (just as in the primary model). Neither control model specification changed the pattern of results (e.g., omitting culture-specific a did not appreciably change group-differences in t or v). Furthermore, these models did not give appreciably better DIC values (-4458.089 and -4459.661) than the primary model (DIC = -4458.970) despite their lower complexity. We therefore concluded that estimates from original model describe the data well.

Experiment 1 Discussion

In Experiment 1 we replicated the typical findings regarding incongruent and far flankers, showing that incongruent and near flankers interfere with target detection. We did not, however, find the expected pattern cultural differences. Specifically, we had predicted that Turks, as a more Eastern culture, would be prone to greater interference from flankers, particularly for stimuli distributed more widely across space (i.e., the far condition) than Americans. Such a finding would be in line with literature indicating that Easterners attend more broadly than Westerners, with corresponding reductions in resolution that could increase interference effects (e.g., Boduroglu & Shah, 2017; Boduroglu et al., 2009; Cramer et al., 2016; Hakim et al., 2017; McKone et al., 2010; Ueda et al., 2018). The data from Experiment 1 do not support this hypothesis. The data suggest, however, that Turks are more accurate than Americans at the flanker task, and that Turks also exhibit a smaller difference in accuracy for incongruent versus congruent flanker trials than Americans (see Figure 2). The speed-accuracy curves for the near

incompatible condition indicate that Turks consistently perform more accurately than Americans across all reaction time bins (see Figure 4, top panel). These results do not indicate cultural differences in speed-accuracy tradeoffs. That is, although each group exhibits a speed-accuracy tradeoff, with higher levels of accuracy at slower reaction times, those tradeoffs do not differ between cultures (e.g., the lines are parallel for the two cultures).

The results from drift diffusion modeling converge with our other behavioral data, indicating that Turks set a higher threshold for making a decision and also accumulate evidence faster, making decisions that are more accurate than Americans in roughly the same amount of time (see Figures 2-5). In the drift diffusion process, higher drift rates typically yield better accuracy and faster responses, as evidence reaches threshold quickly. Conversely, higher decision thresholds give better accuracy, but at the expense of slower responses. People can increase their evidence thresholds to support more accurate responding. For example, older adults raise thresholds to reduce errors (e.g., Ratcliff & McKoon, 2008). In our data, Turks increase both parameters – having higher drift rates and higher thresholds – which is necessary to produce more accurate responses in the same amount of time (e.g., a higher drift rate alone should support high levels of accuracy for Turks at a *faster* response time than that exhibited by Americans). The results of the drift diffusion model reveal that these changes occur without a difference in the groups' non-decision times and without requiring elaborate parameter specifications (e.g. group-specific evidence starting points, or group-specific variance in different parameters).

In order to understand how Turks achieved more accurate performance than Americans at equivalent reaction times (as shown in Figure 4), it might be useful to conceptualize the flanker task in a different manner, one that allows for strategy or other types of attentional differences across cultures. It could be that Turks treat the array as more of a gestalt than Americans, in which case the central letter could stand out against the flankers. Such an idea is consistent with findings that uniform connectedness enhances performance; that is, perceptually

similar information can be treated as a single unit and thus processed faster (e.g., Han & Humphreys, 2005). To investigate this possibility, we conducted a follow-up study in which we varied the level of contrast between the flankers and the central target, similar to a manipulation by Eriksen & Schultz (1979). We predicted that a greater difference in luminance between the flankers and the central target would disrupt participants' ability to treat the entire array as a gestalt. As the luminance of the flankers is reduced we would expect the array to be processed as less of a unit or a "gestalt". In the condition in which flankers are at the same level of contrast as the targets, we would expect to replicate the pattern observed for the near condition in Experiment 1, in which Turks were more accurate than Americans on incongruent trials. For the low contrast trials we would expect the impact of the flankers to be reduced such that the cultural groups perform more similarly. Thus, the benefit Turks showed over Americans in Experiment 1 would be reduced for trials with low contrast flankers. In Experiment 2, we focused on the near flanker condition, which showed the largest effects of culture in Experiment 1.

EXPERIMENT 2

Methods

Participants

Our sample consisted of 33 American undergraduates at Brandeis University, Waltham, USA, and 34 Turkish undergraduates at Bogazici University, Istanbul, Turkey. None of the participants had taken part in Experiment 1. Participants at both locations were native to their respective country of testing and had lived no more than 5 years abroad. All participants gave written consent and were reimbursed with either course credit or cash payment. Data from an additional 15 participants (9 Turks, 6 Americans) were excluded, three because their demographics questionnaires revealed they were not eligible; nine because their math accuracy on the Operation Span task was below the cutoff; one because they were an outlier on the total span score; and two because of missing data.

Materials

Stimuli consisted of 400 Eriksen-type flanker images (B. A. Eriksen & Eriksen, 1974).

Each image consisted of either one or five capital *E*'s and/or *H*'s, with the same font and

luminance as used in Experiment 1. One-hundred of the stimuli were unflanked controls,

consisting of a single letter subtending a 1.0* visual angle presented at the center of the screen.

Three hundred other images were flanked by letters of varying visual contrast. All flankers were

separated by a visual angle of 0.5*, spanning 7.0* from end to end, as they were in Experiment

1's *near* condition. Of these flanked stimuli, one-hundred had *high-contrast* flankers, presented

as black text (luminance = 0.8 cd/m^2) viewed against the grey background (luminance = 43.6

cd/m^2). This produced a Weber contrast value of -0.98. Another 100 stimuli had *medium-contrast* flankers, presented in dark grey text (luminance = 16.8 cd/m^2) that produced a Weber

contrast value of -0.48. The final 100 stimuli had *low-contrast* flankers, presented in light grey

text (luminance = 24.7 cd/m^2), which produced a Weber contrast value of -0.23. In what follows,

we will use the terms "high," "medium," and "low" to denote the contrast of the flankers, with

"high" being the highest contrast flankers, that is, flankers whose luminance matched the

luminance of the target letters. Example stimuli are displayed in Figure 6.

In all flanked conditions, half were congruent, with all five letters the same (either all E's

or all H's), and half were incongruent, with the central letter mismatched with the four flanker

letters. The numbers of *E*'s and *H*'s in each position were equally matched in each condition.

The experiment was administered with the same apparatus and set-up as Experiment 1.

Selection of flanker contrast levels was based on the results of a preliminary experiment.

In that experiment, six American participants completed a modified flanker task in which five

different flanker-contrast levels were presented in random order, blocked by contrast. The *low-contrast* level was selected as the condition in which the flanker compatibility effect no longer

appeared significant, according to the pilot data. The *high-contrast* level had been preselected

as replicating the visual contrast of the stimuli in Experiment 1. The *medium-contrast* level was

then selected by calculating the geometric mean between the Weber contrast values of the *low*- and *high-contrast* conditions.

An additional preliminary test confirmed that flanker letters could be easily seen at each of the selected contrast levels, even the lowest. Five American participants completed another modified flanker task in which the three selected flanker contrast levels and unflanked controls were presented, intermixed, in a random order. Instead of identifying the central target letter, participants were instructed to respond whether or not flanker letters were present on the screen. In each contrast level, participants were able to detect the flankers with over 90% accuracy.

Procedure

Procedures were the same as in Experiment 1, including the same sequence, time course, instructions, and data cleaning to remove outliers for the flanker task. Experiment 2 introduced just two differences to the flanker task. First, this version included 400 trials (instead of 300). Second, this version manipulated the visual contrast of the flanker letters (instead of manipulating the distance between flankers and targets).

Experiment 2 Results

Participant Characteristics

Table 1b summarizes the participants' characteristics[4]. Samples were similar to those in Experiment 1, although Turkish participants were significantly older, $t(64) = 4.11$, $p < .001$, and had completed more years of education, $t(65) = 3.47$, p=.001, than Americans. In terms of reaction time and cognitive tasks, Turks outperformed Americans on the digit comparison speed of processing task, $t(64) = 5.80$, p<.001. For reaction time tasks, American participants responded to the simple, $t(62) = 3.10$, $p = .003$, but not the choice, $t(61) = 1.35$, $p = .18$, reaction time tests faster than the Turks. The groups did not differ on the automated operation span task,

---

[4] The different degrees of freedom reflect measures that were not completed by some participants.

$t(59) = .86$, $p = .39$, suggesting that our samples are well-matched on this task of cognitive ability.

Accuracy

A 2 x 2 x 3 ANOVA was conducted on accuracy with Culture (American, Turkish) as a between-subjects factor and Flanker Congruency (congruent, incongruent) and Flanker Contrast (low, medium, high) as within-subject factors. Results are displayed in Figure 7 including performance on the unflanked control condition. There was a significant main effect of Congruency, $F(1, 65) = 53.80$, $p<.001$, partial $\eta^2 = .45$, with higher levels of accuracy on congruent trials ($M = .97$, $SD = .02$) than on incongruent trials ($M = .93$, $SD = .06$). In addition, there was a significant main effect of Contrast, $F(2, 130) = 5.91$, $p=.004$, partial $\eta^2 = .08$, such that participants were more accurate on low contrast flankers ($M = .96$, $SD = .03$) than medium or high (Ms = .94, SDs = .05) flankers. Finally, there was a significant interaction of Congruency x Contrast, $F(2,130) = 9.14$, $p<.001$, partial $\eta^2 =.12$. For the incongruent trials, accuracy differed across levels of contrast, $F(2, 132) = 11.10$, $p < .001$, partial $\eta^2 = .14$, with accuracy highest for low contrast flankers ($M = .95$, $SD = .05$) than for the medium or high (Ms = .92, SDs = .07) ones. For congruent trials, accuracy did not significantly differ across levels of contrast, $F(2, 132) = .89$, $p = .41$, partial $\eta^2 = .01$, was relatively consistent for congruent trials across high ($M = .97$, $SD = .03$), medium, and low (Ms = .96, SDs = .03) contrast flankers. None of the main effects or interactions involving Culture approached significance (ps > .11). Performance on the unflanked control condition also did not differ across cultures, $t(65) = 1.59$, $p = .12$, Cohen's $d = .39$.

Reaction Time

A 2 x 2 x 3 ANOVA was conducted on median reaction times with Culture (American, Turkish) as a between-subjects factor and Flanker Congruency (congruent, incongruent) and Flanker Contrast (low, medium, high) as within-subject factors. Results are depicted in Figure 8,

including the unflanked control condition. As expected, there was a main effect of Congruency,

F(1, 65) = 208.19, p<.001, partial η^2 = .76, with faster reaction times on congruent (*M* = 443.87

ms; *SD* = 48.26) than on incongruent (*M* = 470.12 ms; *SD* = 49.47) trials. There was also a

main effect of Contrast, F(2, 130) = 10.25, p<.001, partial η^2 = .14, with the fastest RTs for low

contrast flankers (*M* = 452.44, *SD* = 47.16) than medium (*M* = 458.51, *SD* = 49.96) or high (*M* =

460.04, *SD* = 49.91) ones. The interaction of Congruency x Contrast reached significance, F(2,

130) = 21.70, p<.001, partial η^2=.25. Although across the different level of contrast reaction

times were significantly longer for incongruent than congruent trials, the RT cost for incongruent

trials was the smallest for the low contrast condition (*M* congruent = 445.94, *SD* = 48.61; *M*

incongruent = 458.93, *SD* = 48.05, *t*(66) = 5.02, p < .001, Cohen's *d* = .62) compared to the

medium (*M* congruent = 441.66, *SD* = 49.09; *M* incongruent = 475.35, *SD* = 53.66, *t*(66) =

11.39, p < .001, Cohen's *d* = 1.39) and high (*M* congruent = 443.98, *SD* = 50.54; *M* incongruent

= 476.07, *SD* = 51.45, *t*(66) = 12.33, p < .001, Cohen's *d* = 1.50) contrast conditions.

Counter to our predictions, only two culture-related effects reached significance: the

main effect, F(1, 65) = 4.95, p=.03, partial η^2 = .07, with Americans (*M* = 443.86 ms, *SD* =

48.29) responding faster than Turks (*M* = 470.13 ms, *SD* = 48.30), and the interaction of Culture

x Contrast, F(2, 130) = 5.63, p=.005, partial η^2 = .08. For Americans, contrasts indicated a

significant quadratic effect, F(1, 32) = 10.55, p = .003, partial η^2 = .25, consistent with a steeper

drop-off in RT for low contrast flankers (*M* = 436.89, *SD* = 47.15) compared to medium (*M* =

448.71, *SD* = 49.95) and high (*M* = 445.99, *SD* = 49.90) flankers. The pattern differed for the

Turks: the quadratic effect was not significant, F(1, 33) = 1.36, p=.25, partial η^2 = .04; RTs were

more similar across the three conditions, with the high contrast flankers (*M* = 474.08, *SD* =

49.91) slightly slower than the medium (*M* = 468.31, *SD* = 49.95) or low (*M* = 467.99, *SD* =

47.15) contrast flankers (See Figure 8). None of the other effects approached significance, *p*s>

.26. On the unflanked control condition, Americans were significantly faster than Turks, *t*(65) =

2.87, p =.006, Cohen's d = .70, a pattern consistent with the main effect that emerged from analyses including flanker congruency and levels of contrast.

Speed-Accuracy Curves

Speed-accuracy curves are plotted in Figure 9, in the same manner as in Experiment 1. Results from the high-contrast incompatible condition for the slower bins replicate what was found in the near incompatible condition in Experiment 1. In this condition, Turks achieve higher accuracy at the same level of RT as Americans, at least for the slower reaction times. There is also a suggestion of higher levels of accuracy for the Turks for the middle bins of reaction times in the low contrast compatible condition.

Experiment 2 Discussion

Reducing the contrast level of the flankers reduced the magnitude of the flanker effect. In the high contrast condition - identical to the near incompatible condition in Experiment 1 - we replicated the pattern of findings from Experiment 1, with incongruent trials slower than congruent trials. The cultural differences in speed-accuracy graphs for high contrast incompatible flankers were also consistent with the findings in Experiment 1, in that Turks achieved higher levels of accuracy than Americans, at least in the slower bins of reaction times. In contrast, interference effects were reduced for low contrast flankers, producing higher levels of accuracy and faster reaction times. Participants from the two cultures were affected differently by the manipulation of contrast with Americans deriving a somewhat larger benefit for reaction time from the low contrast flankers than Turks, but cultural differences across contrast levels did not interact with compatibility. Comparisons of speed-accuracy curves across cultures also did not consistently identify larger cultural differences for the incompatible condition; although there was a trend for this in the high contrast condition, the only other condition with a notable difference across the cultures was the low contrast compatible condition. Thus, any effects of culture in the speed-accuracy curves did not appear to mitigate the flanker interference effects on incompatible trials.

Thus, the results of Experiment 2 do not support the explanation we proposed for why Turks achieved greater accuracy at faster reaction times than Americans in Experiment 1. Turks' performance on incompatible flankers was not disproportionately disrupted by changing the contrast level of the flankers. We had predicted that reducing the level of contrast would hinder the Turks' ability to treat the array as a gestalt, but this possibility was not supported; we will consider alternate explanations for cultural differences.

Some aspects of the data may reflect changes in the task from Experiment 1. Although we partially replicated the finding from the near incompatible condition (Exp 1; Figure 4) in the equivalent high contrast incompatible condition (Exp 2; Figure 9), the cultural groups only appear to differ in the slower reaction time bins, converging in the faster bins. Moreover, accuracy generally increased for slower reaction time bins in Experiment 1, but that did not occur consistently in Experiment 2. The reduced number of trial types in each condition could account for some variance in the findings in Experiment 2. But it may be the case that changing the contrast level of the flankers altered the task beyond what we had predicted. Having all of the flankers change contrast might allow the target to pop out more and thus might also encourage the use of broad attention to detect the singleton. It is possible that Turks were differentially impacted by these changes to the visual display, perhaps adjusting their breadth of attention or expectations for scanning space. The finding that Turks had overall slower reaction times and were not significantly more accurate than Americans in Experiment 2, patterns that differ from Experiment 1, could provide some evidence for cultural shifts in strategy or other changes in attentional processes. Study designs that allow for direct comparisons of cultural groups across different groupings of trial types would support comparison of strategies. In addition, using different display types, spatial positions, contrasts for unflanked stimuli, as well as flankers of different contrasts within each array (e.g. one low, one medium, one high) could be used to test for cultural differences in multiple types of processes, such as expectations, pop out effects, and the use of gestalt grouping factors.

General Discussion

In two experiments, we demonstrate cross-cultural differences in performance on the flanker task, which suggests that Turks extract more information and do so at a faster rate than Americans. We investigated the underlying processes contributing to cross-cultural differences in multiple ways, including comparison of the distance of flankers, the contrast level of flankers, the interplay between speed and accuracy, and the application of drift diffusion modeling. These studies provide the first direct evidence for cultural differences between Easterners and Westerners on Eriksen's flanker task. Our finding that culture can impact the magnitude of flanker effects is consistent with prior findings, stemming from three different approaches to the study of culture. One of these approaches assessed individual differences in cultural values, finding that those individuals with a more holistic processing style experienced more interference than those individuals with a more analytic style (Hsieh et al., 2020). Another study in this vein used an approach that primes cultural knowledge within an individual (Hong, Morris, Chiu, & Benet-Martinez, 2000). Results indicate that interdependent, Eastern self-construals are associated with greater incongruency effects than independent, Western self-construals (Lin & Han, 2009). A second approach investigates cultural differences by comparing the Himba, a semi-nomadic group, with other cultures. The Himba show a reduced flanker effect in that they are less prone to interference from incompatible flankers than Westerners (de Fockert et al., 2011). They exhibit a local processing bias (de Fockert et al., 2011) and superior selective attention capabilities than other cultural groups (Caparos, Linnell, Bremner, de Fockert, & Davidoff, 2013). A third approach employed the flanker talk to investigate the development of executive control in children, finding that Japanese children exhibited more interference than Canadian children on a social flanker task (Senzaki et al., 2018). Thus, we extend cross-cultural work on the Eriksen flanker task with a direct comparison of adults from Eastern and Western cultures.

The potential mechanisms supporting cultural differences in selective attention have not been conclusively identified. Although congruent and incongruent flankers (Experiment 1) as well as the degree of contrast of the flankers compared to the target (Experiment 2) interact with cultural background of the participants to differentially influence accuracy or reaction times, we cannot fully account for the pattern of results. We designed the experiment to test whether cultures differed in the resolution of information over space, based on the idea that Eastern cultures represent attended information at a lower resolution than Westerners (Boduroglu & Shah, 2017); although lower representational resolution sacrifices detail it makes it possible for attention to be distributed more broadly over space (Lawrence et al., 2020). Such ideas are consistent with Eriksen's zoom model (C. W. Eriksen & St James, 1986; C. W. Eriksen & Yeh, 1985). However, the evidence does not suggest that more spatially distant information interferes with Turks' performance more than Americans, similar to previous work that did not find cultural differences in covert spatial attention using an inhibition of return task (Lawrence et al., 2019).

It is possible that our task design lacked the sensitivity to detect such effects. Accuracy was quite high - above 90% in most conditions - and was particularly high for Turks. Such "ceiling" effects could have limited our ability to detect differences across cultural groups, and contributed to our finding in Experiment 1 of a larger effect of congruency in the Americans than the Turks, due to the high level of performance for the Turks. In addition, it could have been the case that more elements in our stimuli were foveal than was the case for previous work that used stimuli with more widely separated elements (such as in functional field of view or visual change detection tasks; Boduroglu & Shah, 2017), limiting the ability of the far flankers to differentially impact Turks compared to Americans. In addition, the Eastern participants differ from previous studies that recruited East Asians from China (e.g., Boduroglu & Shah, 2017).

There are other candidate explanations that cannot be ruled out based on the design of the flanker task. For example, it may be that broader allocation of attention alters processing of multiple aspects of the visual display, rather than disproportionately impacting the flanker

incompatibility effect. That pattern emerged in a study by Hsieh et al. (2020) for participants who scored high in holistic processing, compared to those high on analytic processing. For example, the central location of the target and use of a central alerting cue (i.e., fixation cross) in all our trials also may have encouraged a narrow scope of attention. Additionally, the schedules of trials representing different conditions could have impacted attentional bias. Both experiments reported here contained trials consisting of a single, unflanked target. In fact, in Experiment 1, these unflanked control trials together with the near flanker trials constituted the majority of trials. This distribution of trials may have promoted a relatively narrow distribution of attention; future work investigating potential cultural differences in the flexibility of attentional allocation would be necessary to address responsiveness to different contingencies. It is also possible that cultures differ in the tendency to occasionally attend to the wrong part of the stimulus (Gaspelin et al., 2014), such that Americans mistakenly allocate attention to the flanker more than Turks in some circumstances; varying the spatial position of the target or other modifications to the visual display would be necessary to assess this possibility. Cultures could also differ in their sensitivity to visual crowding. Although this possibility still needs to be tested directly, any differences in the visual complexity of the environments typical for Turks and Americans could have influenced the ways in which the two cultures allocate visual attention. After all, differences in environmental complexity between two other cultures has been shown to influence the allocation of visual attention (Miyamoto, Nisbett & Masuda, 2006; Gutchess & Sekuler, 2019).

The finding in both experiments that Turks extract more information faster than Americans may be the most interesting takeaway from our study. Applying a drift diffusion modeling approach allows for finer-grained conclusions to be drawn about the nature of cross-cultural differences, based on our finding that Turks both accumulate evidence faster (higher drift rate) and require more evidence (higher threshold) than Americans in order to make a decision. Although differences in general cognitive ability can impact flanker performance through differences in conflict control (e.g., Liu, Xiao, Shi, Zhao, & Liu, 2011), we can eliminate

this explanation in accounting for overall cultural differences in task performance, as our cultural groups are matched on Operation Span scores. Drift diffusion modeling add a more mechanistic explanation to previous studies reporting cross-cultural differences in overall cognitive effort, such as East Asians' tendency to exhibit higher levels of cognitive persistence than Americans in continuing to improve their performance across task blocks (Telzer, Qu, & Lin, 2017). Future work manipulating decision thresholds and evidence quality would be helpful to further understand cultural differences in these processes, including potentially identifying conditions in which cultures converge.

Our findings of cultural differences are consistent with the continuous flow model (C. W. Eriksen & Schultz, 1979), emphasizing the dynamic nature of information processing, including in attention. This model emphasizes the gradual accumulation of information about stimuli in the visual system, which allows for responses to be primed or partially activated. Building on Eriksen & Schultz's (1979) demonstration that target discriminability is affected by external manipulations (e.g., location in space, figure-ground contrast, size of target relative to noise), our results indicate that the internal factor of culture also contributes to target discriminability. Compared to Americans, Turks' faster accumulation of information allows them to respond faster with a high level of accuracy.

Further research is needed to investigate the potential links between the mechanisms that lead to cultural differences across tasks, and what common mechanisms are impacted across disparate tasks. For example, are the findings reported in the present study related to cultural differences in prioritization of low versus high spatial frequency information (Tardif et al., 2017)? Additionally, the extent to which task demands and constraints induce cultural effects represents a rich domain for future work. More systematic delineation of different cultural groups is also important to consider; in the present study, our hypotheses are based on prior work with East Asian participants, but further work is necessary to assess the convergence of Turkish and East Asian participants' performance.

Despite these limitations, our results offer some conceptual and methodological points that are important to consider for future work across cultures. Flanker compatibility and the contrast level of interfering information affect performance differently across cultures. Moreover, the data from both of our experiments illustrate the importance of considering accuracy as well as speed measures, and the utility of applying drift diffusion modeling in cross-cultural research. Some research argues that cultural differences do not simply reflect differences in speed-accuracy tradeoff, as cultural differences trended in the same direction for both speed and accuracy or groups did not differ on measures of accuracy or corrected measures of speed divided by accuracy (Caparos et al., 2013; McKone et al., 2010). Our results are in line with this argument. However, in order to understand cultural differences on this task more completely, it was important to *integrate* measures of accuracy and speed. The fact that Turks achieved higher levels of performance than Americans was revealed most clearly when performance was viewed through the lens of an evidence accumulation framework.

Open Practices Statement

      The data, but not the materials, are available at: https://osf.io/t46w8/. The experiments

were not preregistered.

References

Boduroglu, A., & Shah, P. (2017). Cultural differences in attentional breadth and resolution. *Culture and Brain, 5*(2), 169-181. doi:10.1007/s40167-017-0056-9

Boduroglu, A., Shah, P., & Nisbett, R. E. (2009). Cultural Differences in Allocation of Attention in Visual Information Processing. *Journal of Cross-Cultural Psychology, 40*(3), 349-360. doi:10.1177/0022022108331005

Caparos, S., Linnell, K. J., Bremner, A. J., de Fockert, J. W., & Davidoff, J. (2013). Do local and global perceptual biases tell us anything about local and global selective attention? *Psychol Sci, 24*(2), 206-212. doi:10.1177/0956797612452569

Coates, D. R., Levi, D. M., Touch, P., & Sabesan, R. (2018). Foveal Crowding Resolved. *Scientific Reports, 8*(1), 9177. doi:10.1038/s41598-018-27480-4

Cramer, E. S., Dusko, M. J., & Rensink, R. A. (2016). Group-level differences in visual search asymmetry. *Atten Percept Psychophys, 78*(6), 1585-1602. doi:10.3758/s13414-016-1137-0

Dale, G., & Arnell, K. M. (2013). Investigating the stability of and relationships among global/local processing measures. *Attention, Perception, & Psychophysics, 75*(3), 394-406. doi:10.3758/s13414-012-0416-7

Davids, A., & Eriksen, C. W. (1957). Some social and cultural factors determining relations between authoritarianism and measures of neuroticism. *Journal of Consulting Psychology, 21*(2), 155-159. doi:10.1037/h0047132

Dayan, P., & Solomon, J. A. (2010). Selective Bayes: attentional load and crowding. *Vision Res, 50*(22), 2248-2260. doi:10.1016/j.visres.2010.04.014

de Fockert, J. W., Caparos, S., Linnell, K. J., & Davidoff, J. (2011). Reduced distractibility in a remote culture. *PLoS One, 6*(10), e26337. doi:10.1371/journal.pone.0026337

Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. *Behav Res Methods, 43*(1), 258-268. doi:10.3758/s13428-010-0024-1

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of Noise Letters upon Identification of a Target Letter in a Nonsearch Task. *Percept Psychophys, 16*(1), 143-149. doi:10.3758/bf03203267

Eriksen, C. W. (1954). Some personality correlates of stimulus generalization under stress. *The Journal of Abnormal and Social Psychology, 49*(4, Pt.1), 561-565. doi:10.1037/h0054094

Eriksen, C. W. (1957). Personality. *Annual Review of Psychology, 8*, 185-210. doi:10.1146/annurev.ps.08.020157.001153

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25*(4), 249-263. doi:10.3758/BF03198804

Eriksen, C. W., & St James, J. D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. *Percept Psychophys, 40*(4), 225-240. doi:10.3758/bf03211502

Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of Attention in the Visual Field. *Journal of Experimental Psychology: Human Perception and Performance, 11*(5), 583-597. doi:10.1037/0096-1523.11.5.583

Farzin, F., & Norcia, A. M. (2011). Impaired visual decision-making in individuals with amblyopia. *J Vis, 11*(14), 6-6. doi:10.1167/11.14.6

Freeman, J. B., Ma, Y., Han, S., & Ambady, N. (2013). Influences of culture and visual context on real-time social categorization. *Journal of Experimental Social Psychology, 49*(2), 206-210. doi:10.1016/j.jesp.2012.10.015

Gaspelin, N., Ruthruff, E., & Jung, K. (2014). Slippage theory and the flanker paradigm: An early-selection account of selective attention failures. *Journal of Experimental Psychology: Human Perception and Performance, 40*(3), 1257-1273. doi:10.1037/a0036179

Gutchess, A. H., & Sekuler, R. (2019). Perceptual and Mnemonic Differences across Cultures. *Psychology of Learning and Motivation, 71*, 131-174.

Hakim, N., Simons, D. J., Zhao, H., & Wan, X. (2017). Do easterners and westerners differ in visual cognition? A preregistered examination of three visual cognition tasks. *Social Psychological and Personality Science, 8*(2), 142-152. doi:10.1177/1948550616667613

Han, S., & Humphreys, G. W. (2005). Perceptual organization at attended and unattended locations. *Sci China C Life Sci, 48*(2), 106-116.

Hedden, T., & Park, D. (2001). Aging and interference in verbal working memory. *Psychol Aging, 16*(4), 666-681.

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frontiers in Neuroscience, 8*, 150.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2-3), 61-83. doi:10.1017/s0140525x0999152x

Hong, Y. Y., Morris, M. W., Chiu, C. Y., & Benet-Martinez, V. (2000). Multicultural minds. A dynamic constructivist approach to culture and cognition. *Am Psychol, 55*(7), 709-720.

Hsieh, S., Yu, Y.-T., Chen, E.-H., Yang, C.-T., & Wang, C.-H. (2020). ERP correlates of a flanker task with varying levels of analytic-holistic cognitive style. *Personality and Individual Differences, 153*, 109673. doi:https://doi.org/10.1016/j.paid.2019.109673

Huang, L. (2010). Visual working memory is better characterized as a distributed resource rather than discrete slots. *J Vis, 10*(14), 8. doi:10.1167/10.14.8

Lawrence, R. K., Edwards, M., Chan, G. W. C., Cox, J. A., & Goodhew, S. C. (2019). Does cultural background predict the spatial distribution of attention? *Culture and Brain*. doi:10.1007/s40167-019-00086-x

Lawrence, R. K., Edwards, M., Talipski, L. A., & Goodhew, S. C. (2020). A critical review of the cognitive and perceptual factors influencing attentional scaling and visual processing. *Psychonomic Bulletin and Review*. doi:10.3758/s13423-019-01692-9

Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635-654. doi:https://doi.org/10.1016/j.visres.2007.12.009

Lin, Z., & Han, S. (2009). Self-construal priming modulates the scope of visual attention. *Quarterly Journal of Experimental Psychology, 62*(4), 802-813. doi:10.1080/17470210802271650

Linnell, K. J., & Caparos, S. (2011). Perceptual and cognitive load interact to control the spatial focus of attention. *J Exp Psychol Hum Percept Perform, 37*(5), 1643-1648. doi:10.1037/a0024669

Liu, T., Xiao, T., Shi, J., Zhao, D., & Liu, J. (2011). Conflict control of children with different intellectual levels: an ERP study. *Neurosci Lett, 490*(2), 101-106. doi:10.1016/j.neulet.2010.12.035

McKone, E., Aimola Davies, A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., & Platow, M. J. (2010). Asia has the global advantage: Race and visual attention. *Vision Res, 50*(16), 1540-1549. doi:10.1016/j.visres.2010.05.010

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. *Percept Psychophys, 49*(3), 270-288. doi:10.3758/bf03214311

Miyamoto, Y., Nisbett, R. E., & Masuda, T. (2006). Culture and the physical environment - Holistic versus analytic perceptual affordances. *Psychological Science, 17*(2), 113-119.

Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *J Vis, 7*(2), 20-20. doi:10.1167/7.2.20

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision, 2*(4), 237-279. doi:10.1037/dec0000030

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873-922. doi:10.1162/neco.2008.12-06-420

Rosenholtz, R., Yu, D., & Keshvari, S. (2019). Challenges to pooling models of crowding: Implications for visual mechanisms. *J Vis, 19*(7), 15-15. doi:10.1167/19.7.15

Schwartz, A. J., Boduroglu, A., & Gutchess, A. H. (2014). Cross-Cultural Differences in Categorical Memory Errors. *Cognitive Science, 38*(5), 997-1007. doi:10.1111/cogs.12109

Senzaki, S., Wiebe, S. A., Masuda, T., & Shimizu, Y. (2018). A cross-cultural examination of selective attention in Canada and Japan: The role of social context. *Cognitive Development, 48*, 32-41. doi:https://doi.org/10.1016/j.cogdev.2018.06.005

Strasburger, H. (2020). Seven Myths on Crowding and Peripheral Vision. *i-Perception, 11*(3), 2041669520913052. doi:10.1177/2041669520913052

Suchow, J. W., Fougnie, D., Brady, T. F., & Alvarez, G. A. (2014). Terms of the debate on the format and structure of visual memory. *Attention, Perception, & Psychophysics, 76*(7), 2071-2079. doi:10.3758/s13414-014-0690-7

Tardif, J., Fiset, D., Zhang, Y., Estephan, A., Cai, Q., Luo, C., . . . Blais, C. (2017). Culture shapes spatial frequency tuning for face identification. *J Exp Psychol Hum Percept Perform, 43*(2), 294-306. doi:10.1037/xhp0000288

Telzer, E. H., Qu, Y., & Lin, L. C. (2017). Neural processes underlying cultural differences in cognitive persistence. *Neuroimage, 156*, 224-231. doi:10.1016/j.neuroimage.2017.05.034

Ueda, Y., Chen, L., Kopecky, J., Cramer, E. S., Rensink, R. A., Meyer, D. E., . . . Saiki, J. (2018). Cultural Differences in Visual Search for Geometric Figures. *Cognitive Science, 42*(1), 286-310. doi:10.1111/cogs.12490

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behav Res Methods, 37*, 498-505.

Uskul, A. K., Kitayama, S., & Nisbett, R. E. (2008). Ecocultural basis of cognition: Farmers and fishermen are more holistic than herders. *Proceedings of the National Academy of Sciences of the United States of America, 105*(25), 8552-8556. doi:10.1073/pnas.0803874105

White, C. N., Ratcliff, R., & Starns, J. J. (2011). Diffusion models of the flanker task: Discrete versus gradual attentional selection. *Cognitive Psychology, 63*(4), 210-238. doi:10.1016/j.cogpsych.2011.08.001

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Front Neuroinform, 7*, 14. doi:10.3389/fninf.2013.00014

Table 1a.

*Demographic information and mean (with standard deviation) test scores for participants in Experiment 1. * denotes effects significant at p<.05.*

	American	Turkish	Significance
Age	18.5 (.8)	20.7 (2.7)	$p < .01$ *
Years of Education	13.2 (1.2)	12.9 (1.2)	$p > .20$
Median Simple RT(ms)	267 (30)	297 (25)	$p < .01$ *
Median Choice RT(ms)	417 (52)	492 (67)	$p < .01$ *
Digit Comparison	79.9 (16.9)	77.1 (15.8)	$p > .50$
Operation Span Score	46.7 (15.6)	52.7 (13.6)	$p > .09$

Table 1b.

*Demographic information and mean (with standard deviation) test scores for participants in Experiment 2. * denotes effects significant at p<.05.*

	American	Turkish	Significance
Age	19.3 (1.4)	20.9 (1.7)	$p < .01$ *
Years of Education	12.7 (1.0)	13.5 (.8)	$p < .01$ *
Median Simple RT	286 ms (21)	307 ms (31)	$p < .01$ *
Median Choice RT	458 ms (52)	476 ms (54)	$p > .18$
Digit Comparison	78.9 (11.2)	105.5 (23.9)	$p < .01$ *
Operation Span Score	50.2 (11.7)	47.2 (15.1)	$p > .39$

Figure 1. Trial time course and stimuli for Experiment 1. The top panel illustrates the timing of the different trial components, displaying a stimulus for an unflanked control trial. The middle and bottom panels display example stimuli for the different conditions of flanker trials.
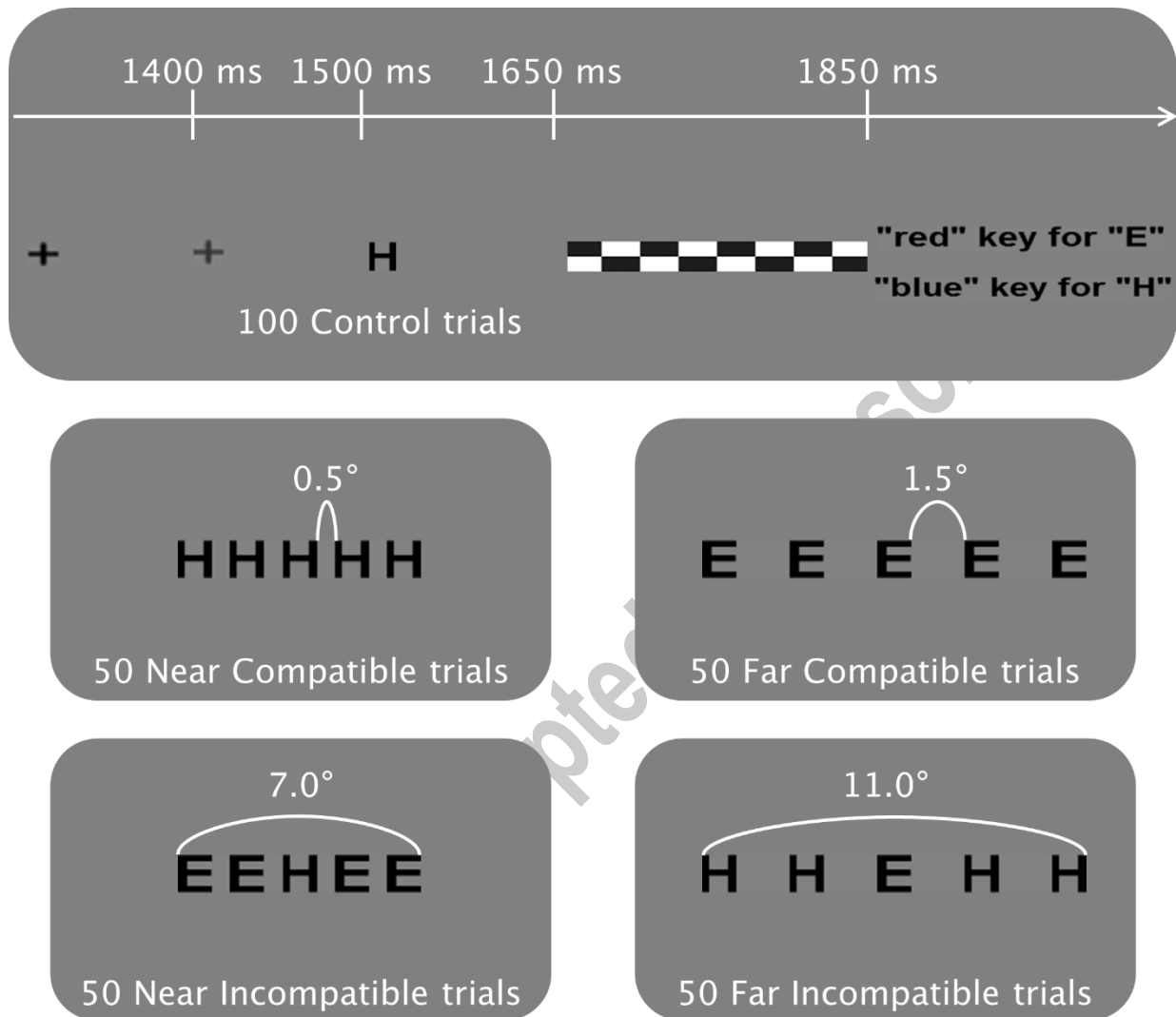
Figure 2. Accuracy for Americans and Turks in Experiment 1. The US mean for unflanked controls is marked by a dashed line (--) and Turkish mean for unflanked controls is marked by a dotted line (•••). Error bars represent the standard error of the mean.
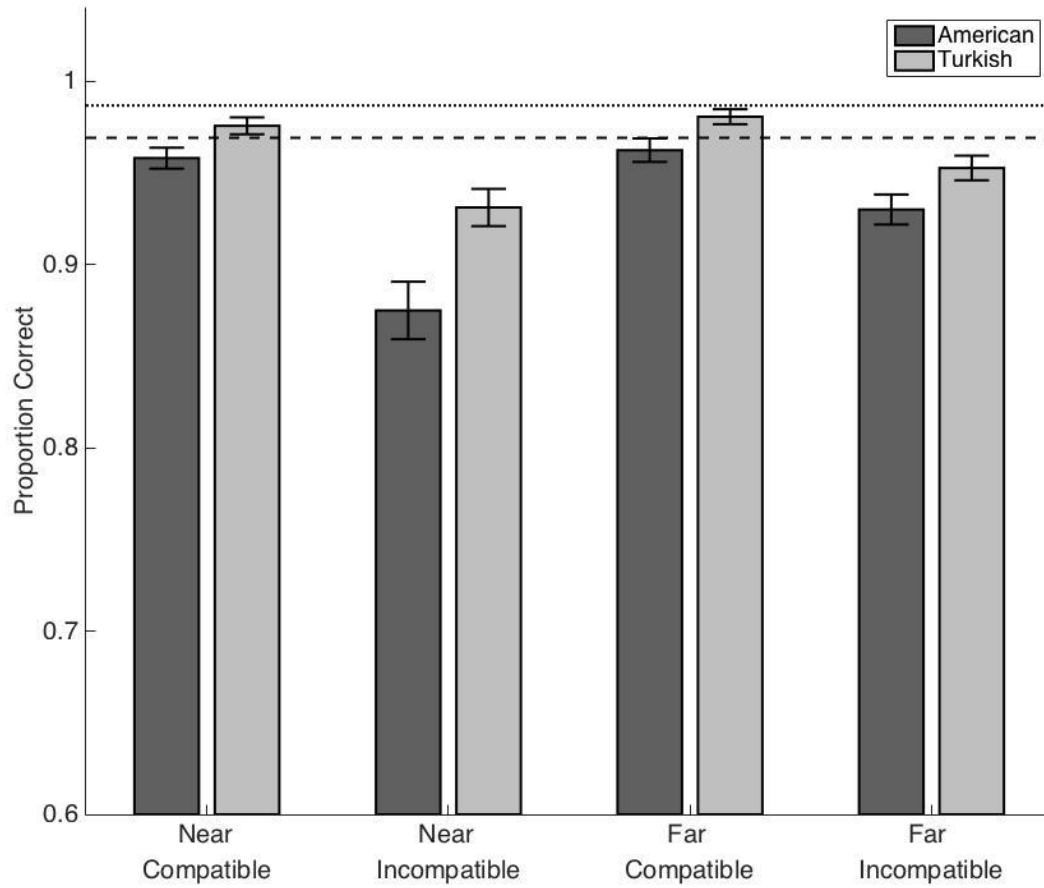
Figure 3. Reaction times for Americans and Turks in Experiment 1. The figure displays Tukey box plots, for which the whiskers represent 1.5x the interquartile range. US median reaction time for unflanked controls is marked by a dashed line (--) and Turkish median reaction time for unflanked controls is marked by a dotted line (•••).
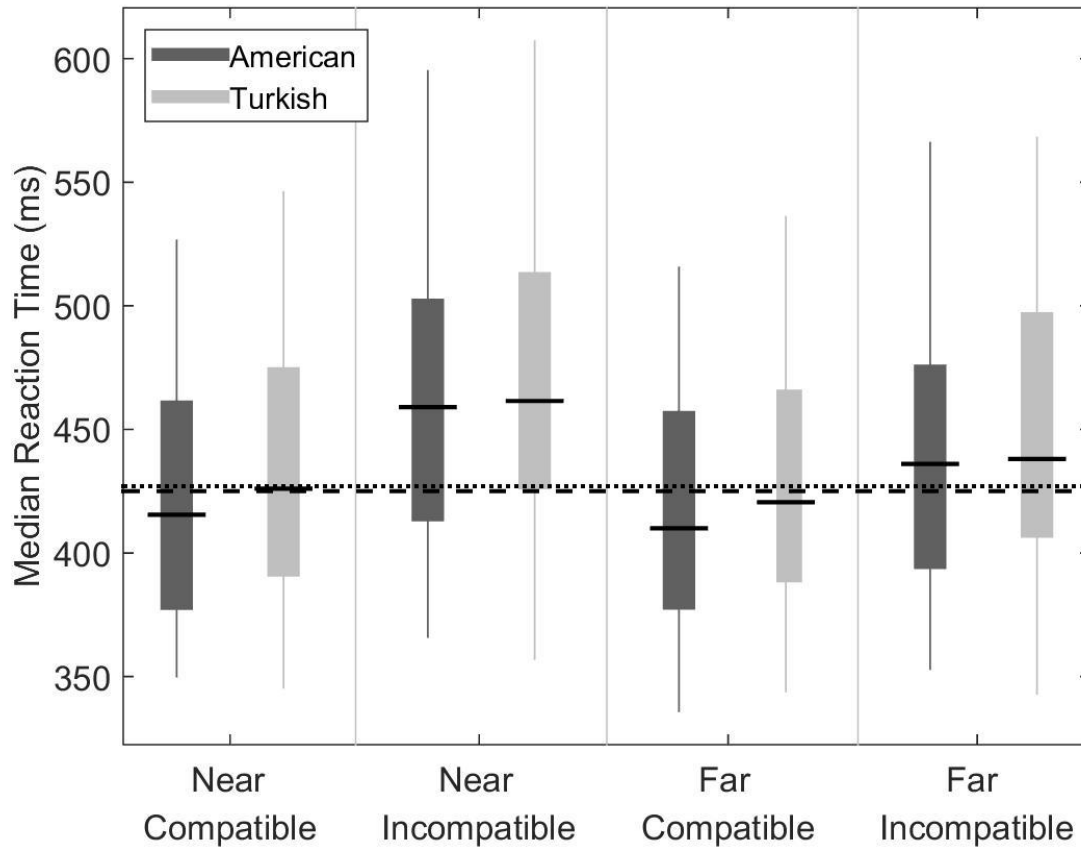
Figure 4. Discrimination (d') plotted as a function of reaction time for Experiment 1. d' scores are plotted for each 100 msec bin with at least 50 trials. Notable in the near incompatible condition is that Turks (light) achieve higher d' scores than Americans (dark).
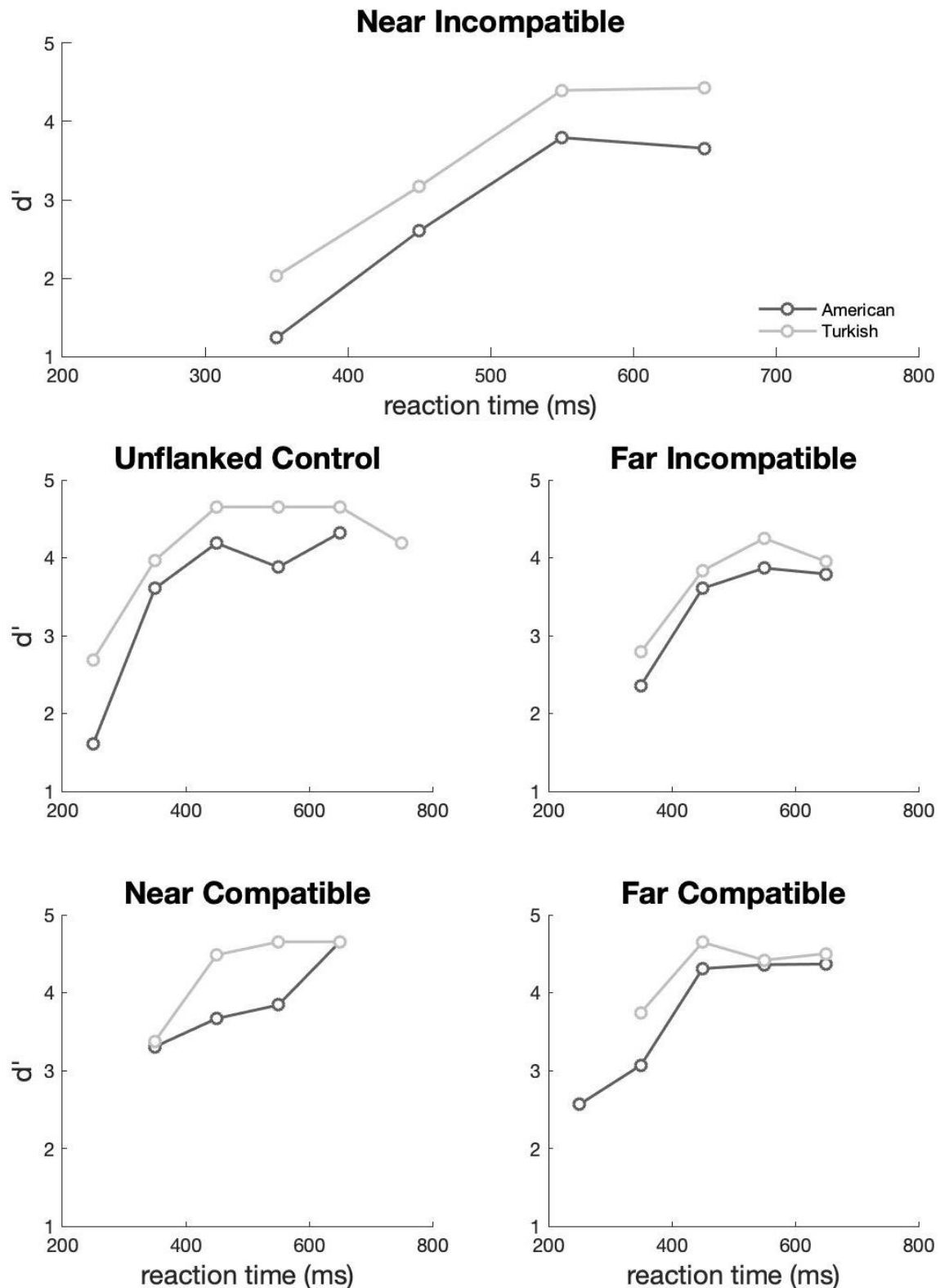
Figure 5. The differences in group mean parameter posterior distributions are displayed in three panels, corresponding to the drift rate (top), threshold (middle), and non-decision time (bottom).
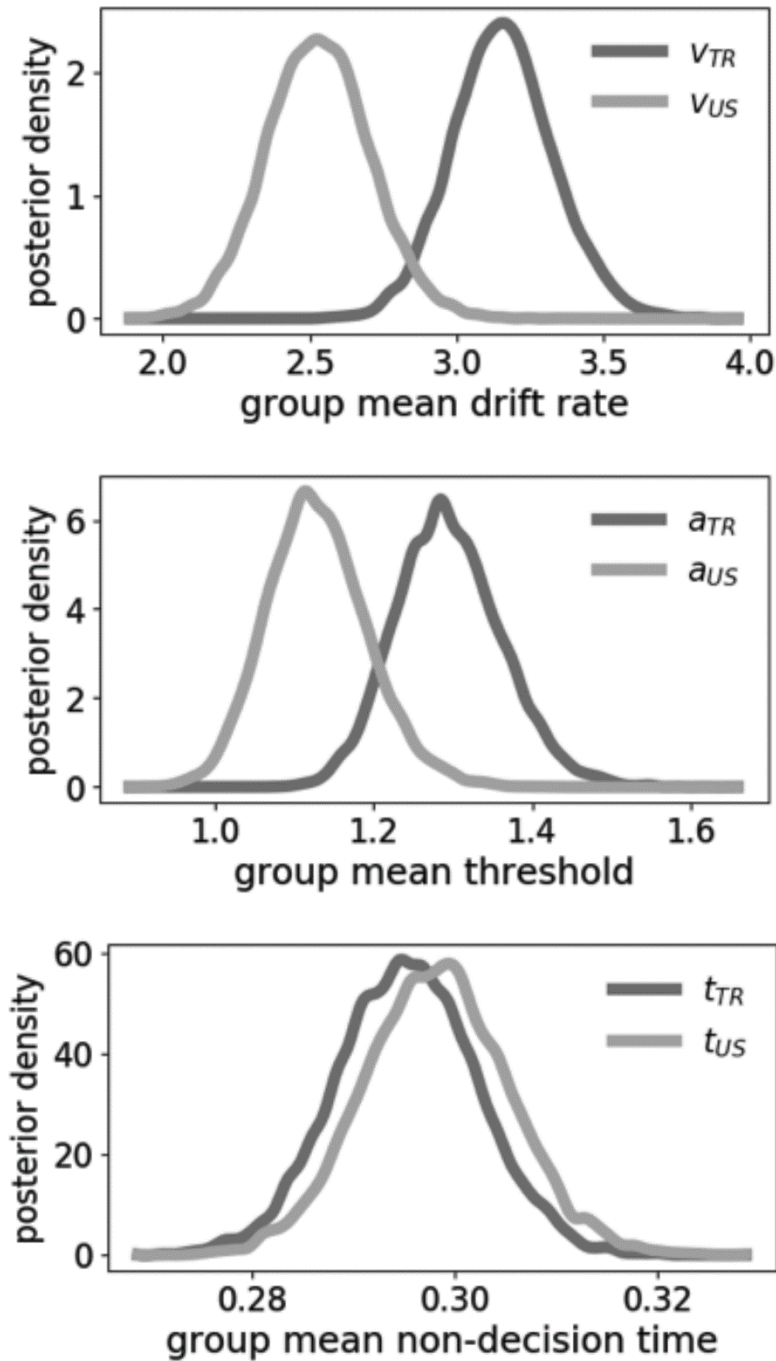
Figure 6. Example of stimuli presented in Experiment 2. Disclaimer: in this reproduction, figures are not to scale and the luminances do not reproduce the actual luminances presented in the experiment.

Figure 7. Accuracy for Americans and Turks in Experiment 2. The US mean for unflanked controls is marked by a dashed line (--) and Turkish mean for unflanked controls is marked by a dotted line (•••). Error bars represent the standard error of the mean.
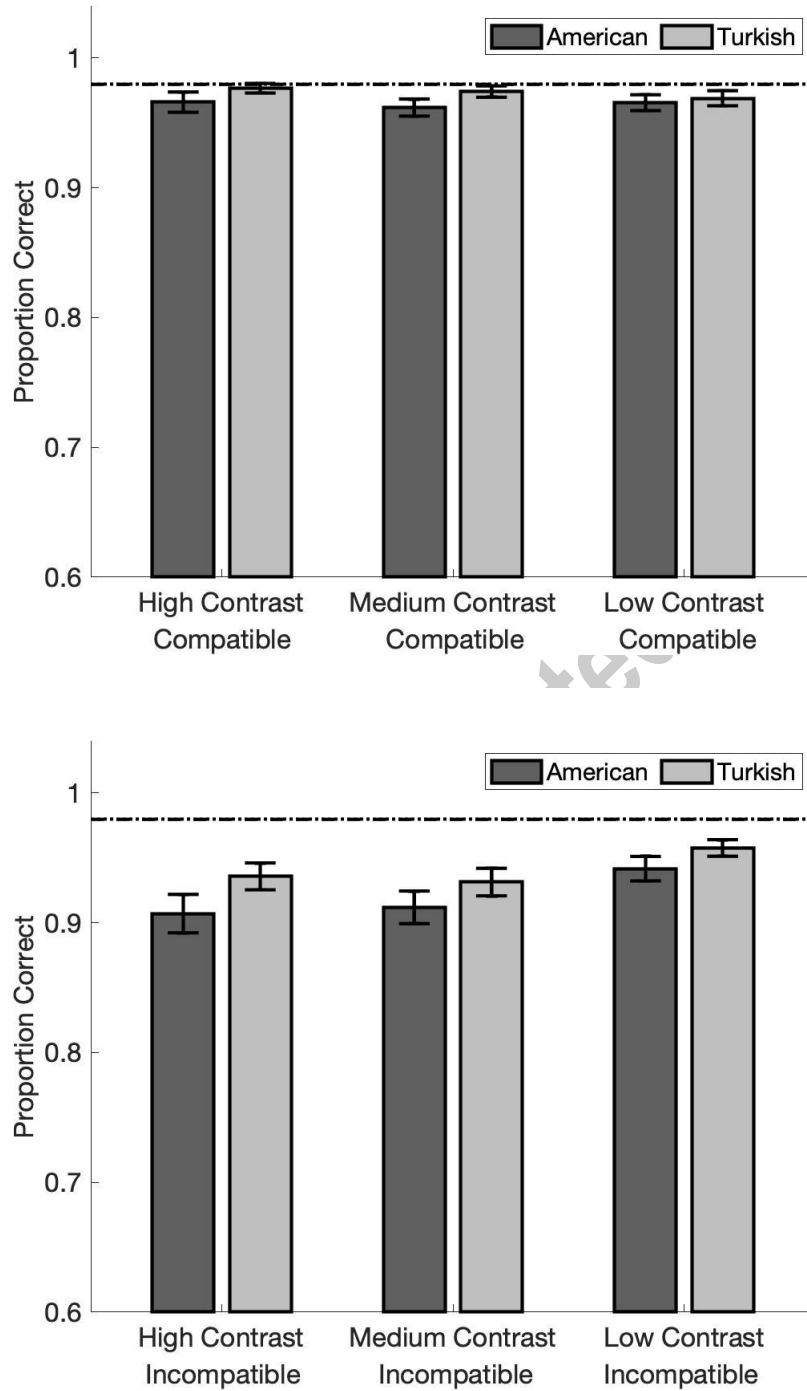
Figure 8. Reaction times for Americans and Turks in Experiment 2. The figure displays Tukey box plots, for which the whiskers represent 1.5x the interquartile range. US median reaction time for unflanked controls is marked by a dashed line (--) and Turkish median reaction time for unflanked controls is marked by a dotted line (•••).
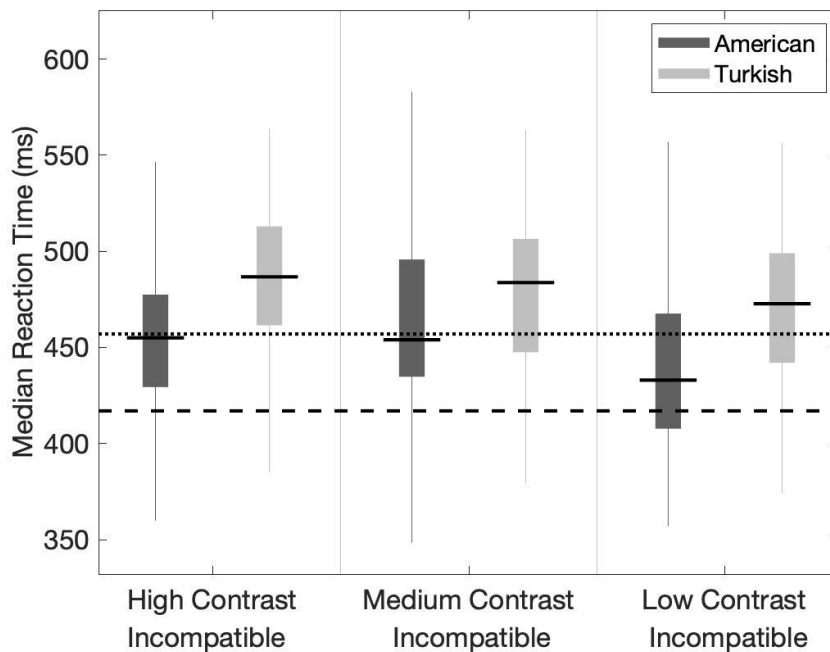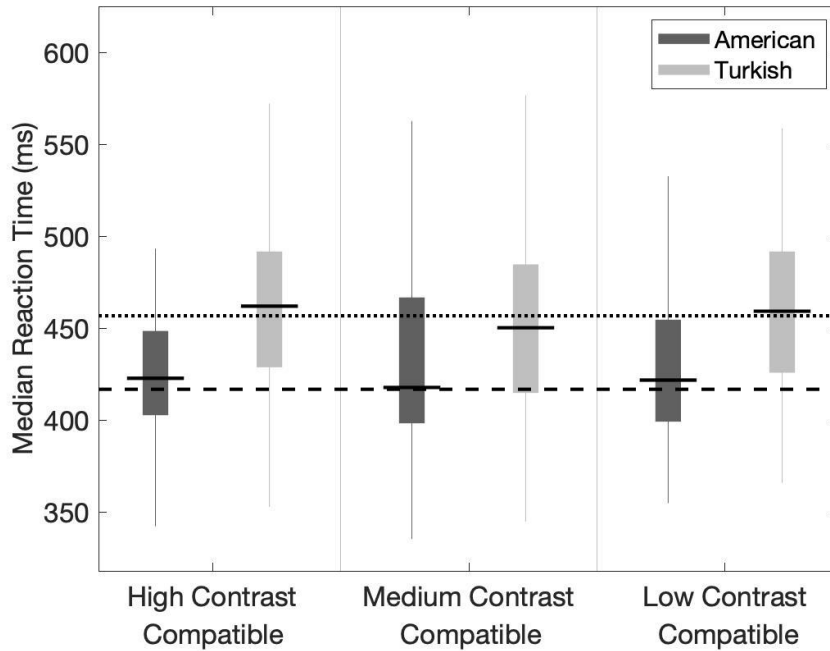
Figure 9. Accuracy (d') plotted as a function of reaction time for Experiment 2. d' scores are plotted for each 100 msec bin in which there were at least 50 trials.