RUNNING HEAD: Affective theory of mind inferences

Affective theory of mind inferences contextually influence the recognition of emotional facial

expressions

Suzanne L. K. Stewart (Corresponding author)

Department of Psychology, University of Chester, Parkgate Road, Chester CH1 4BJ

s.stewart@chester.ac.uk; 01244 511 680

Astrid Schepman

Department of Psychology, University of Chester, Parkgate Road, Chester CH1 4BJ

a.schepman@chester.ac.uk; 01244 511 658

Matthew Haigh

Department of Psychology, Northumbria University, Newcastle upon Tyne NE1 8ST

matthew.haigh@northumbria.ac.uk; 0191 227 3472

Rhian McHugh

Department of Psychology, University of Chester, Parkgate Road, Chester CH1 4BJ

r.mchugh@chester.ac.uk; 01244 513 144

Andrew J. Stewart

Division of Neuroscience and Experimental Psychology; Faculty of Biological, Medical, and

Human Sciences; University of Manchester; Oxford Road; Manchester M13 9PL

andrew.stewart@manchester.ac.uk; 0161 275 7331

Abstract

The recognition of emotional facial expressions is often subject to contextual influence, particularly when the face and the context convey similar emotions. We investigated whether spontaneous, incidental affective theory of mind inferences made while reading vignettes describing social situations would produce context effects on the identification of same-valenced emotions (Experiment 1) as well as differently-valenced emotions (Experiment 2) conveyed by subsequently presented faces. Crucially, we found an effect of context on reaction times in both experiments while, in line with previous work, we found evidence for a context effect on accuracy only in Experiment 1. This demonstrates that affective theory of mind inferences made at the pragmatic level of a text can automatically, contextually influence the perceptual processing of emotional facial expressions in a separate task even when those emotions are of a distinctive valence. Thus, our novel findings suggest that language acts as a contextual influence to the recognition of emotional facial expressions for both same and different valences.

Key words: theory of mind; inference; emotion; context; face processing

Word count: 7,994 (including the reference list)

Affective theory of mind inferences contextually influence the recognition of emotional facial expressions

Much of our everyday social interaction relies on our ability to understand the mental states of others, which is known as "theory of mind" (Premack & Woodruff, 1978). Theory of mind involves representation of knowledge, beliefs, and intentions (cognitive theory of mind) as well as emotions [affective theory of mind (aToM); e.g., Shamay-Tsoory, Tomer, Berger, Goldsher, & Aharon-Peretz, 2005]. Specifically, aToM inferences often rely on the contextual, communicative value of observed emotional facial expressions (e.g., the "Reading the Mind in the Eyes Test," Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). What is less clear, however, is the reverse – whether contextual aToM inferences can influence the subsequent identification of emotional facial expressions. We investigated this question by looking for evidence of this influence in measures of processing speed and accuracy when participants identified facial emotions after reading vignettes that implied an emotion that either was congruent or incongruent with the subsequent facial emotion.

Previous work has established that people generally do consider the surrounding context when identifying facial emotions (Aviezer et al., 2008; Aviezer, Trope, & Todorov, 2012; Barrett & Kensinger, 2010; Kayyal, Widen, & Russell, 2015; Righart & de Gelder, 2008; Schwarz, Wieser, Gerdes, Mühlberger, & Pauli, 2013; cf. Nakamura, Buck, & Kenny, 1990). "Context" pertains to anything separate from the facial emotion itself (Hassin, Aviezer, & Bentin, 2013; Wieser & Brosch, 2012) and includes (but is not limited to) pictorial scenes, body position and body language, individual emotional words, vignettes, and even the neurological processes occurring in parallel within the perceiver (Barrett, Lindquist, & Gendron, 2007).

However, there is good theoretical reason to believe that the degree of contextual influence on the identification of emotional facial expressions varies according to specific conditions. One such theoretical example is that of "limited situational dominance" (Carroll & Russell, 1996). In

this account, Carroll and Russell assert that observers rely on three dimensions of facial expressions in order to accurately classify the emotions they convey, and these are quasi-physical information, pleasantness, and arousal. Quasi-physical information pertains to the physical aspects of an expression that characterise that expression but which are not unique to it (Carroll and Russell give the example of a smile which can be recognised as such but does not determine whether the expression is of joy, embarrassment, nervousness, or a polite greeting). According to their view, pleasantness pertains to the positive or negative valence of the facial emotion, and arousal to its intensity. Carroll and Russell argue that when emotions portrayed by a face and by a situation are incongruent on all three dimensions (e.g., happiness and sadness), then the emotion in the face will take precedence, meaning that context has little to no influence. However, when the facial and situational emotional information are congruent on these aspects (e.g., fear and anger: aroused, unhappy, staring quasi-physical features; negative valence; high arousal), then the emotion portrayed by the situation will take precedence, meaning that the context has a strong influence and the facial emotion may, therefore, be mis-classified. Carroll and Russell found evidence for this theory in a series of experiments in which participants listened to the researcher read emotionally charged vignettes and then viewed photographs of differing emotional faces which were typically still congruent for quasi-physical information, pleasantness, and arousal (e.g., fear and anger). Participants chose what emotion (from several choices) the face was expressing, and their responses tended to be congruent with the vignette's emotion rather than the intended emotion of the face. Thus, Carroll and Russell's findings suggest that people extract affective information from narratives which seems to incidentally influence how emotional information in a subsequently presented face is interpreted, particularly when the emotions are relatively similar. However, it is impossible to directly attribute this effect to the written content of the vignettes because hearing the vignettes read aloud is a contextual influence in itself – the researcher could have unwittingly emphasised the emotion consistent with the vignette through prosodic factors and his/her own facial expressions and body language (Wieser & Brosch, 2012). Nonetheless, similar effects have been

uncovered with neuroimaging (Kim et al., 2004) and for ambiguous versus unambiguous facial emotions (Trope, 1986). Thus, the work presented here tested the theory of limited situational dominance in a new way through tightly controlled experiments that allowed participants to read the vignettes themselves rather than listening to them, that used a larger number of items, and that measured the effect on both processing and a subsequent classification task. Furthermore, as will be described, our vignettes were designed such that any context effect observed came from the implicit, spontaneous, incidental aToM inferences that participants made during narrative comprehension; and these were not confounded by the behaviour of the researcher or by explicit emotion words in the vignettes.

A contrasting theoretical view comes from Barrett and colleagues (Barrett & Kensinger, 2010; Barrett et al., 2007; Gendron, Lindquist, Barsalou, & Barrett, 2012; Lindquist, Barrett, Bliss-Moreau, & Russell, 2006), who developed a "language-as-context" hypothesis. Their evidence suggests that when people encounter individual emotional words, the comprehension of these words activates conceptual knowledge and sensory-related information in memory and that these simulations then act as top-down influences on the perception of simultaneously or subsequently presented facial stimuli (Gendron et al., 2012). Thus, these researchers suggest that the emotion word response options in many experiments contextually influence the perception of facial stimuli. In the work presented below, we were interested in linguistic contextual influences to the interpretation of facial emotions beyond the lexical level and, therefore, tested an extended version of the language-as-context hypothesis. We investigated whether spontaneous aToM inferences made at the pragmatic level about someone else's inner experiences could incidentally influence the perception of emotional facial expressions in an unrelated task. Thus, we tested predictions generated by the limited situational dominance account (broadly, that aToM inferences will only be influential in the conditions where these inferences are similar to the facial emotion to be identified, e.g., fear / anger but not happiness / sadness) versus an extension of the language-as-context account (broadly, that aToM inferences will be influential regardless of the similarity of the

inferences to the emotion in the face, e.g., fear / anger as well as happiness / sadness). We focused on the pragmatic level of narrative comprehension as psycholinguistic research suggests that people spontaneously make mental and emotional state inferences during reading, although the specificity of these inferences is debatable (Gernsbacher, Goldsmith, & Robertson, 1992; Gygax, Oakhill, & Garnham, 2003; Haigh & Bonnefon, 2015). Furthermore, this focus is similar to work involving the picture verification task used by Zwaan and colleagues (e.g., Zwaan, Stanfield, & Yaxley, 2002) which demonstrates that people mentally activate specific perceptual details of an object which are only implied by a preceding text.

Across two experiments, we investigated context effects of aToM inferences on the identification of facial emotions and determined whether the limited situational dominance account or the extended language-as-context account was better able to explain the findings. In Experiment 1, we explored what happens with emotions that are similar in terms of valence and arousal by examining congruent and incongruent combinations of situations and faces depicting fear and anger. Participants read vignettes that invited a "fear" or "anger" inference about the mental state of a character before being asked to identify the emotion of a subsequently presented face which portrayed either fear or anger. Combinations of fear and anger were also tested by Carroll and Russell (1996) because the affective signals for fear and anger are congruent for quasi-physical features (aroused, staring, unhappy expression), pleasantness, and arousal yet are discrepant for specific emotions. According to Carroll and Russell's theory of limited situational dominance, the situational emotion should dominate and so we expected that reaction times (RTs) recorded during the face classification task would be slower and that responses would be less accurate when the emotions of the situation and the face were incongruent compared to when they were congruent. The extended version of Barrett and colleagues' language-as-context hypothesis would suggest that aToM inferences made at the pragmatic level of the text unlock related sensory information and information from memory, producing a context effect on the identification of subsequent emotional faces. This account also makes a prediction that RTs will be slower and responses less accurate

when the emotions of the vignette and the face are incongruent. Thus, Experiment 1 tested whether the methodology is a valid, reliable, and sensitive way of detecting the expected context effects. Subsequently, Experiment 2 pitched the two theoretical models against each other by testing their differing predictions (described later) for congruent and incongruent combinations of differently-valenced emotions (happiness / sadness).

Experiment 1

Method

For both experiments, we report how our sample size was determined and all data exclusions, manipulations, and measures.

Participants

A power analysis indicated that 32 participants would be sufficient to find a medium effect size at approximately 80% power (Lenth, 2006-9). Thus, 32 participants aged between 18 and 65 were opportunity sampled from students and staff at the University of Chester (25 female; mean age = 25.50 years, SD = 9.65 years). A further two participants were tested but their data discarded due to one performing below chance for facial emotion recognition accuracy and one being inadvertently run on the wrong experimental list. All participants confirmed no serious visual impairments, no reading difficulties such as dyslexia, and a first language of English. Participants were eligible for a prize draw of one of ten £10 Amazon.co.uk vouchers and were awarded participation credits where suitable. The study was approved by the University of Chester Department of Psychology Ethics Committee.

Materials

Vignettes (see *Anger example*) were composed which described social situations in which emotional reactions might be expected.[1] They comprised four sentences, and all vignettes involved social situations with a named main character interacting with or being affected by at least one other person (never named). Explicit descriptions of emotions or specific emotional words were avoided

---

[1] See online supplemental material for detailed descriptions of the vignette development for both experiments.

(Barrett et al., 2007; Wieser & Brosch, 2012); therefore, the emotion felt by the character had to be inferred. There was no instruction to make such an inference, and making an inference was not vital for comprehension of the vignette so any such inferences were spontaneous and elaborative.

Anger example:

Lucy worked part-time for a local newspaper and had been working on a big story about a campaign to save the historic town hall. She had even worked overtime and had spent her own money to interview lots of residents and do all the research. Her editor praised Lucy for all her hard work and told her it would be on the front page. When Lucy bought the paper the next day, she saw her editor had put his own name on the report.

Thirty-two angry and fearful vignettes were used in the experimental items (henceforth, "item" refers to the pairing of vignettes and faces). Thirty-two additional vignettes equally split across sadness, happiness, surprise, and disgust were used in the filler items. Vignettes within every emotion category were balanced for the main character's gender. Both experiments' vignettes are available online as supplemental material.

Colour photographs of faces were selected from the Karolinska Directed Emotional Faces database (Lundqvist, Flykt, & Ohman, 1998). Hit rates from Goeleven, De Raedt, Leyman, and Verschuere's (2008) study indicating good recognition were used to select 16 angry and 16 fearful faces. The hit rates for the angry (74.10%, SD = 11.03%) and fearful (73.13%, SD = 6.56%) faces did not differ, t (30) = 0.30, p = .765. However, the angry faces had a lower mean arousal rating than the fearful faces, 3.25 (SD = 0.26) versus 3.82 (SD = 0.39); t (30) = 4.90, p < .001. To examine the impact of this difference in arousal, we ran the main RT analysis with and without arousal as a co-variate. The models were not statistically different, $\chi2$ (2) = 2.113, p = .348, meaning that the

difference in arousal levels of the angry versus fearful faces did not impact the tested effect (and the key interaction between Vignette Emotion and Face Emotion remains significant when the covariate of arousal is added). Thus, the analysis in which arousal was free to vary is presented below. Thirty-two additional filler faces expressing happiness, sadness, disgust, and surprise were selected using high recognition hit rates. Faces representing each emotion were balanced for gender.

Design and procedure

All 32 experimental items were counterbalanced along a 2 (vignette: angry vs. fearful) x 2 (face: angry vs. fearful) design. Thus, two lists were created such that experimental vignettes paired with congruent faces in the first list were paired with incongruent faces in the second; equal numbers of participants viewed each list. Each list also contained 32 filler items that were a mix of congruent and incongruent combinations of vignettes and faces representing happiness, sadness, disgust, and surprise. The main character's gender was matched with the gender of the subsequently presented face.

The experiment was run in E-Prime 2 (Version 2.0.10.353; Psychology Software Tools, 2012). Participants sat comfortably at a desktop computer with a standard keyboard with their forefingers resting on the "A" and "L" keys. The first screen presented detailed instructions, while the second screen presented the key instructions in a numbered list, which emphasised that participants should identify the emotional expression of the face as quickly and accurately as possible. This was followed by a practice block of three trials and then the experimental block of 64 trials. For each trial, participants first saw a central fixation cross and pressed the spacebar to advance when ready. This was followed by the vignette presented in Arial size 12 font. After reading it at their own pace, participants pressed the spacebar to advance to the next screen, which immediately presented a centrally-located face at width = 50% and height = 60%. The face was flanked by two possible response options (e.g., Angry / Fearful) to the lower left and lower right in Arial size 18. Participants pressed either the "A" or "L" key as quickly as possible to make their response (correct answers were counterbalanced across left and right, so that response side was

balanced across emotion, gender, and congruency). The next screen immediately presented a comprehension question (Arial size 18) about a factual aspect of the vignette, which was flanked to the lower left by "yes" and to the lower right by "no," both displayed in Arial size 18. Again, participants pressed either the "A" or "L" key to respond. Half the questions should have been answered "yes" and half "no"; these were counterbalanced across facial emotion, gender, and congruency. A response caused the next trial to begin. Comprehension questions were used on every trial to encourage deeper processing of the text (e.g., Stewart, Holler, & Kidd, 2007). No feedback was given. A final block of five trials involving happy vignettes and faces was presented. These trials, which were not analysed, were presented so that participants would leave the lab in a positive frame of mind, which was a requirement of the ethics committee. Accuracy of responses to the faces and comprehension questions was recorded along with RTs in milliseconds from the onset of the face.

Analysis

To analyse the effect of Vignette Emotion and Facial Emotion on RTs we used linear mixed-effects models (LMMs; Baayen, Davidson & Bates, 2008) using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Development Core Team, 2017). For the accuracy data we used the *glmer* function under the binomial distribution. There are several advantages of the (G)LMM approach over factorial ANOVA, which is the statistical technique most frequently paired with 2x2 experimental designs. Two key advantages of (G)LMMs for 2x2 experimental designs are that (1) they are able to account for multiple random effects simultaneously (see Clark, 1973, for a discussion highlighting the importance of considering random effects related to items), allowing more of the error to be modelled, and (2) all the individual trials can be entered into the analysis rather than means for each participant, which gives more statistical power because (G)LMMs are, therefore, able to handle the interdependence of repeated observations (Baayen et al., 2008). The

code and data for our (G)LMM analyses can be found at https://osf.io/tne5b/. More details about the parameter estimates of the (G)LMMs are given below.

Results and discussion

The participants had a mean accuracy of 95.81% (SD = 5.12%) for the comprehension questions and a mean accuracy of 92.56% (SD = 7.21%) for facial emotion recognition. Data for 115 trials (11.23%) were excluded from the RT analysis for inaccuracy on either facial emotion recognition or comprehension question responses. RTs for a further 25 trials (2.75%) were slower than the mean plus three standard deviations for their pertinent conditions. These were replaced with the equivalent of the mean plus three standard deviations for the relevant conditions.

In our LMM analysis for the RT data, the fixed effects were Vignette Emotion (Fear, Anger), Facial Emotion (Fear, Anger), and the interaction between these factors. We used deviation coding for each of the experimental factors. Our model contained crossed random effects for participants, vignettes, and faces. The model with the most maximal effects structure that converged included random intercepts and additive slopes for both fixed factors by participants, and by vignettes, and random intercepts and slopes for the Facial Emotion factor by faces. Restricted maximum likelihood estimation was used when reporting the LMM parameters (see Table 1 for parameter estimates). The model revealed an interaction between Vignette Emotion and Facial Emotion that was significant at <.001 alpha level (estimated by approximating to the z-distribution).

(Table 1 about here)

The interaction was explored with pairwise comparisons performed using the emmeans package in R (Lenth, Love & Hervé, 2018) with degrees of freedom approximated using the Kenward-Roger method. The pairwise comparisons were interpreted using a Bonferroni-corrected alpha level of .025. Figure 1 contains the marginal means and standard errors calculated using the emmeans package. They show that RTs were slower for identifying a facial emotion when it mismatched the vignette emotion. Angry faces were recognised faster after angry vignettes than

after fearful vignettes, t (23.77) = 3.439, p = .002, while fearful faces were recognised faster after fearful vignettes than after angry vignettes, t (27.41) = 4.611 p < .001.[2]

(Figure 1 about here)

In our GLMM analysis for the accuracy data, the fixed effects were Vignette Emotion (Fear, Anger), Facial Emotion (Fear, Anger), and the interaction between these factors. Our model contained crossed random effects of participants and vignettes. The random effect of faces was dropped in order to arrive at a model that converged. The model with the most maximal effects structure that converged included only random intercepts by participants and by vignettes (see Table 2 for parameter estimates). The model revealed an interaction between Vignette Emotion and Facial Emotion that was significant at < .001 alpha level (based on the z-distribution).

(Table 2 about here)

The interaction was explored with contrasts performed using the emmeans package in R on the log odds ratio scale. The pairwise comparisons were interpreted using a Bonferroni-corrected alpha level of .025. Figure 2 contains the estimated marginal means and standard errors calculated using the emmeans package. They show that accuracy decreased when identifying facial emotions that mismatched the preceding vignette emotions. Angry faces were responded to with higher accuracy after angry vignettes than after fearful vignettes, z = 4.393, p < .001, while fearful faces were responded to with higher accuracy after fearful vignettes than after angry vignettes, z = 2.575, p = .01.

(Figure 2 about here)

The findings of Experiment 1 demonstrate that after spontaneously making aToM inferences about vignette characters, participants were slower to judge facial emotions that mismatched those inferences and were also more likely to make errors compared to when facial emotions and aToM

---

[2] The same pattern of results is found with a 2x2 ANOVA.

inferences matched. These results suggest that people represent more than the simple valence of a situation (e.g., positive / negative). Instead, they actively make richer, more specific aToM inferences about the characters, and this substantially impacts the processing and identification of a subsequent facial expression. These results support Carroll and Russell's (1996) theory of "limited situational dominance" as well as the extended language-as-context hypothesis of Barrett and colleagues. The findings establish that this methodology is sensitive to providing insight into the processing and judgements associated with aToM context effects.

Thus, Experiment 1 demonstrated that both theoretical models can explain the evidence that aToM inferences act as a strong contextual influence upon same-valenced yet discrepant emotions, which is also shown by much empirical work. This evidence supports the assertion that context assumes a more important role in discriminating between emotions when valence proves less useful. However, it is possible that context still remains an influential force even when valence is distinctive, as has been demonstrated by work on incongruent non-linguistic contexts (e.g., Aviezer et al., 2012; but see Hassin et al., 2013). Although this has not yet been conclusively established, it is possible that such an effect exists for linguistic contexts as well but has not been detected because of the tendency for research in the area to examine accuracy responses only rather than processing (e.g., Carroll & Russell, 1996, Kayyal et al., 2015; Schwarz et al., 2013). Thus, it is possible that this effect is evident during processing, but the information about distinctive valence may then override the cost to processing so that a correct accuracy response is made, which would explain results frequently found for linguistic contexts. Therefore, the crucial evidence for a context effect on processing would come from a difference in RTs between congruent and incongruent conditions. Additionally, a replication of findings from previous empirical work on linguistic contexts (e.g., Carroll & Russell, 1996, Schwarz et al., 2013) would demonstrate no difference in accuracy responses accompanying the critical context effect for RTs. Evidence of a context effect on RTs would strongly support the extended language-as-context hypothesis, which suggests that contextual emotional information has an influence regardless of the degree of similarity between the

contextual and target emotions. Subsequently determining whether there is an effect on accuracy would provide evidence of the strength of this effect in terms of whether valence information is able to override the contextual processing effect to produce correct accuracy responses (e.g., shown by no effect on accuracy) or not (e.g., shown by an effect on accuracy). In contrast, because the theory of limited situational dominance describes that emotions of positive versus negative valence may also be distinct on quasi-physical information and arousal (e.g., happiness / sadness), then according to this account, there should be no evidence of a context effect in the crucial analysis of the processing (RT) data. This would subsequently be accompanied by no evidence of an effect in the accuracy data. Thus, Experiment 2 set out to replicate Experiment 1 using happiness and sadness in order to investigate the extent of the influence of contextual aToM inferences and to determine whether the theory of limited situational dominance or the extended language-as-context hypothesis better accounted for the findings overall.

Experiment 2

Method

Participants

Thirty-two new individuals, selected in the same way as for Experiment 1, participated (20 female, 11 male, 1 undisclosed gender; mean age = 28.83 years, SD = 11.85 years). Data for two additional participants were excluded because they performed at or below chance on accuracy for the comprehension questions or facial recognition.

Materials

Thirty-two happy and sad vignettes were used in the experimental materials.[3] Thirty-two additional vignettes equally split across disgust, fear, anger, and surprise were used in the filler items.

---

[3] See the online supplemental material for a description of the vignettes' development.

Photographs of 16 happy and 16 sad faces, balanced for gender, were selected in the same way as for Experiment 1. The happy faces had a higher mean recognition rate than the sad faces, 99.61% (SD = 0.70%) versus 94.92% (SD = 2.71%); t (16.99) = 6.71, p < .001, but this is unsurprising given the recognition advantage ascribed to happy faces (e.g., Leppänen & Hietanen, 2004). Importantly, this bias was controlled for by the nature of the experimental effect we investigated (i.e., an interaction between the vignette's intended emotion and the intended facial emotion). As expected, the happy and sad faces differed on arousal, with the happy faces having a higher mean arousal rating than the sad faces, 4.03 (SD = 0.37) versus 3.38 (SD = 0.36); $t (30)$ = 5.01, p < .001. An additional 32 faces spread equally across anger, fear, disgust, and surprise and balanced for gender were selected for the filler items.

Design, procedure, and analysis

The design, procedure, and analysis were the same as in Experiment 1. In addition, we calculated the Bayes Factor for both the RT and accuracy measures to determine whether the data were more supportive of Model 1 (which predicts a congruency effect for the RT and accuracy data) or Model 2 (which predicts no congruency effect for the RT and accuracy data). This was done following the procedure based on the BIC values for the two possible models (Raftery, 1995; Wagenmakers, 2007). With respect to the theoretical models, the extended language-as-context account would be supported by evidence for Model 1 for the RTs (regardless of whether Model 1 or Model 2 is supported for accuracy), while the limited situational dominance account would be supported by evidence for Model 2 for the RTs (and subsequently by evidence for Model 2 for accuracy).

Results and discussion

For the comprehension questions, participants had a mean accuracy of 93.44% (SD = 5.47%); and for the facial emotion recognition, participants had a mean accuracy of 97.25% (SD =

4.36%). Data for 91 trials (8.89%) were excluded from the RT analysis because of inaccuracy on

the comprehension questions or for facial recognition. Overall, RTs for 16 trials (1.71%) were

slower than the mean plus three standard deviations for their pertinent conditions; these were

replaced with the equivalent of the mean plus three standard deviations for the relevant conditions.

A LMM was fitted to the RT data, following the same procedure used in Experiment 1. The

fixed effects were Vignette Emotion (Happy, Sad), Facial Emotion (Happy, Sad), and the

interaction between these factors. We used deviation coding for each of the experimental factors.

Our model contained crossed random effects for participants, vignettes, and faces. The model with

the most maximal effects structure that converged included random intercepts and slopes for both

fixed effects and the interaction between them by participants, and by vignettes, and random

intercepts and slopes for the fixed factors additively (i.e., dropping the interaction term) by faces.

Restricted maximum likelihood estimation was used when reporting the LMM parameters (see

Table 3 for parameter estimates). The model revealed an interaction between Vignette Emotion and

Facial Emotion that was significant at <.001 alpha level (estimated by approximating to the *z*-

distribution).

(Table 3 about here)

As with Experiment 1, the interaction was explored with pairwise comparisons performed

using the emmeans package in R with degrees of freedom approximated using the Kenward-Roger

method. The pairwise comparisons were interpreted using a Bonferroni corrected alpha level

of .025. Figure 3 contains the marginal means and standard errors calculated using the emmeans

package. Happy faces were recognised faster after happy vignettes than after sad vignettes, t

$(25.71) = 3.996$, $p < .001$, while sad faces were recognised faster after sad vignettes than after

happy vignettes, t $(22.80) = 3.918$, $p < .001$.[4] In order to calculate the Bayes Factor to compare our

Model 1 (i.e., with additive effects of Vignette Emotion and Face Emotion and with the interaction

between them) to Model 2 (i.e., with additive effects of Vignette Emotion and Face Emotion, but

---

[4] The same pattern of results is found with a 2x2 ANOVA.

with no interaction between them) we had to simplify the random effects structure; for both models we had random intercepts and slopes involving additive effects of Vignette Emotion and Face Emotion for participants and vignettes, and (uncorrelated) random intercepts and slopes involving Face Emotion for faces. This produced a Bayes Factor (BF12) of 8,886,111 in support of Model 1 over Model 2.

(Figure 3 about here)

In our GLMM analysis for the accuracy data, the fixed effects were Vignette Emotion (Happy, Sad), Facial Emotion (Happy, Sad), and the interaction between these factors. Our model contained crossed random effects of participants and vignettes. The random effect of faces was dropped in order to arrive at a model that converged. The model with the most maximal effects structure that converged included only random intercepts by participants and by vignettes. The model revealed no main effects nor an interaction between Vignette Emotion and Face Emotion (see Table 4 for parameter estimates). Figure 4 contains the marginal means and standard errors calculated using the *emmeans* package. The Bayes Factor comparing Model 1 (i.e., with additive effects of Vignette Emotion and Face Emotion and with the interaction between them) to Model 2 (i.e., with additive effects of Vignette Emotion and Face Emotion, but with no interaction between them) was calculated to be 6 (BF21) in support of Model 2 over Model 1.

(Table 4 about here)

(Figure 4 about here)

The RT results of Experiment 2 replicated those of Experiment 1. However, participants did not make more errors in recognising facial emotions when they were preceded by incongruent inferences compared to congruent ones. The Bayes Factor calculations suggest that we have positive evidence for a congruency effect in RTs (Model 1) and positive evidence for a lack of a congruency effect in accuracy (Model 2; Raftery, 1995, Wagenmakers, 2007). Thus, the critical RT analysis in Experiment 2 provided evidence for Barrett and colleagues' theoretical extended

"language-as-context" account over Carroll and Russell's (1996) theory of limited situational

dominance, which predicted no difference in RTs. Consistent with previous empirical work on

linguistic contexts, we also found no evidence for an accompanying difference in accuracy rates.

This suggests that context is still influential on processing even when contexts and faces are of

different valences (which strongly supports the extended language-as-context account) but that

valence information is then likely used to override this processing cost to produce correct accuracy

responses. This conclusion will be discussed further in the General Discussion with regards to both

experiments.[5]

Comparison of the RT Congruency Effect between Experiments 1 and 2

Upon observing the apparent difference in the size of effects on RTs between the two

experiments, we decided to examine whether the magnitude of the congruency effect associated

with the RT data differed between Experiments 1 and 2 by conducting an additional, post-hoc LMM

analysis. As the congruency effect was symmetrical for each of our pairs of factors in Experiments

1 and 2, we re-coded our conditions as Congruent vs. Incongruent. We added Experiment as a new

fixed effect, and re-coded our participant, vignette, and face factors uniquely. The other fixed effect

was Congruency. The model included crossed random effects for Congruency by participants,

vignettes, and faces, each with random intercepts and slopes. We used deviation coding for each of

the fixed factors and restricted maximum likelihood estimation was used when reporting the

parameters (see Table 5). This revealed effects of Congruency, Experiment and an interaction

between them that was significant at the < .05 alpha level (estimated by approximating to the z-

distribution). The congruency effect was larger in Experiment 1 (719 ms) than it was in Experiment

2 (343 ms). Figure 5 contains the marginal means and standard errors calculated using the emmeans

package.

[5] One possible alternative explanation is that certain words in the vignettes primed the emotion word response options at a lexical level. We re-ran the analyses excluding vignettes that contained possible cue words, and the pattern of effects was the same. See the online supplemental material for full details of these analyses.

(Table 5 about here)

(Figure 5 about here)

General Discussion

In both experiments, RTs were slower when participants identified facial emotions that conflicted with spontaneous aToM inferences made while reading previously presented vignettes that described social situations, regardless of whether the stimuli were of the same valence (anger / fear, Experiment 1) or of different valences (happiness / sadness, Experiment 2). However, participants only made more errors when the stimuli were of the same valence. Similarly, both Gendron et al. (2012) and Aviezer et al. (2008, Experiment 3) found evidence of context effects occurring during the perceptual process. Using eye-tracking, Aviezer et al. found that scanning patterns of faces reflected the emotion conveyed by the context (e.g., focusing on the lower face for a disgust context even when the face was intended to convey anger). The pattern of our accuracy and RT findings as well as our comparison of the congruency effects across the two experiments support Aviezer et al.'s contention that context effects are likely to be stronger the more "confusable" the emotions are (Hassin et al., 2013). However, our findings show that exposure to two successive, differently-valenced emotions can also significantly affect the perceptual process (Aveiezer et al.'s eye-tracking experiment used same-valenced emotions).

It appears that even when contexts are dissimilar from emotional facial expression targets in terms of valence, arousal, and quasi-physical information, processing takes longer when the context and the target facial expression convey different emotions compared to when they convey the same emotion. This suggests that context cannot be easily discounted even when valence information is useable for discriminating emotions. Indeed, a comparison of the congruency effects across the two experiments shows that the effect is stronger when the contextual and facial emotions match on quasi-physical features, pleasantness, and arousal (anger / fear). This comparison means that

although context does not have the same degree of influence on processing when the context and face are incongruent on quasi-physical information, pleasantness, and arousal (happiness / sadness), it still has a significant impact on processing and the time taken to make a correct accuracy response. However, even this variation in the degree of contextual influence on processing does not support the limited situational dominance account (Carroll & Russell, 1996), which predicts no influence of context when the emotions portrayed by the context and the face differ on quasi-physical information, pleasantness, and arousal (e.g., happiness / sadness). It must be noted that this finding was a result of a post-hoc observation rather than an a priori prediction and should be interpreted with caution.

Thus, our findings fit best with an extended version of Barrett and colleagues' "language-as-context" hypothesis (Barrett & Kensinger, 2010; Barrett et al., 2007; Gendron et al., 2012; Lindquist et al., 2006). Their work suggests that activation of emotional concepts and related sensory information in memory through actions such as reading individual emotion labels shapes a person's interpretation of an emotional facial expression. While it is possible that the option words flanking the faces in our experiments influenced the perception of the facial expressions in the way that Barrett et al. suggest (though these were held constant across the congruent and incongruent conditions so would not have influenced the congruency effect we observed), our findings clearly demonstrated that the broader simulations of social vignettes influenced the perception of the faces. Thus, it appears that it is not only simulations and concept activations generated at a lexical level that produce context effects but also operations at a pragmatic level. Indeed, while Barrett et al. suggest that context effects result from the re-activation of a person's own previous emotional experiences through actions like reading emotion labels, our findings demonstrate that context effects can also result from inferences about the emotional experiences of other people in social situations. In other words, a person's engagement of his/her aToM during the construction of the situation model of someone else's circumstances can actively influence the decoding and interpretation of emotional facial expressions, regardless of valence.

At a fine-grained level, how does text comprehension lead to a specific aToM inference? Dynamic models of emotion processing, such as the Component Process Model (CPM; see Scherer, 2009), may be instructive. In the CPM account, events (such as those described in our vignettes) are appraised along several dimensions proceeding through increasingly complex levels of processing. These dimensions determine relevance for the individual in terms of needs and goals, implications for the individual and his/her needs and goals in terms of the perceived outcomes of the events and their probability of occurrence (for example), how the individual might cope with the event given his/her relative control and power over it, and the normative significance of the event as compared to self and social standards (Scherer, 2001). According to Scherer, this appraisal process leads to changes in motivation and contributes to generating an active or potentially physical response; these also influence the ongoing appraisal process. The integration of these continually updating appraisals, motivational changes, and potentials for action constitute the emotional experience, which includes subconscious and conscious levels (Scherer, 2009). Although the CPM describes an individual's own experience of emotion, we cautiously speculate that it may also support aToM inferences by allowing simulations and appraisals of others' experiences. Thus, in our vignettes, we suspect the simulated emotional experience continually updates as events unfold over the course of the narrative and as the reader appraises those events in terms of relevance, implications, coping potential, and normative significance for the character and his/her needs and goals based on the available information. Given the relative brevity of our vignettes and the lack of nuanced detail about the characters, the outcome emotions derived from the vignettes may approximate "modal emotions" that represent more commonly experienced patterns of appraisals, motivational changes, and potentials for action (e.g., anger, fear).

These emotional patterns, particularly the conscious aspects of them (which are labelled "feelings" by Scherer), may be found to be congruent or incongruent with the subsequent face. In the *Anger example,* readers would simulate and appraise Lucy's situation. They would assess it as being highly relevant for her, involving an intentional thwarting of her goals, and being significant

for her self-esteem. These appraisals would be more typical of anger than of other modal emotions (Scherer, 2001).

This possible accounting of emotion derivation via the CPM fits, broadly, with the creation and updating of a reader's situation model of the text. General and social knowledge are thought to form part of situation models, but this knowledge is also factored into the appraisal process described above (Scherer, 2009). Thus, the situation model may contain the comprehension of the text, general and social knowledge, and simulated emotional experiences which integrate to allow the production of elaborative aToM inferences about the character which then also enter the integrated situation model (see Zwaan & Radvansky, 1998). Indeed, emotional aspects of a text seem to activate brain regions that process affective information above and beyond language comprehension (Ferstl, Rinck, & von Cramon, 2005). The situation model is also thought to enter long-term memory (Zwaan, 1999). Thus, the enduring, integrated situation model which is formed and updated, in part, at the pragmatic level of the text and which includes incidental and elaborative aToM inferences is likely a significant contextual factor that then affects the processing and response to a subsequently viewed face. This has important implications for our understanding of both empathy and mentalising, as well as how context affects the perception of facial emotions.

Barrett and Kensinger (2010) concluded that context is particularly salient when participants are asked to decide upon a specific emotion (e.g., giving a specific label such as fear or disgust), rather than general affect (e.g., a more general positive or negative feeling such as approaching or avoiding a face). What is less clear, however, is whether, and under what circumstances, context can influence affective judgements. Gygax et al. (2003) demonstrated that readers make general affective inferences in an automatic and subconscious way, and it also appears that observers make general affective inferences about faces automatically and subconsciously (Barrett & Kensinger, 2010). Kim et al. (2004) found that context effects produced different activation patterns in the amygdala even during passive exposure to stimuli. Because this is an under-explored area, future research should investigate what occurs when the affective information of context and face conflict

using a paradigm similar to ours but where the judgement task is unrelated to affect. This would demonstrate whether context influences processing at an implicit level when attention is specifically directed to non-emotional aspects of the stimuli (Kim et al.'s participants may have directed their own attention to the stimuli's emotional aspects). We predict that there would still be evidence of a context effect observed in processing measures (e.g., RTs, eye movements) when affective information in the context and the face is incongruent.

Despite our clear findings, we must acknowledge that it is possible that participants thought that they were to categorise the emotion presented in the vignette rather than the face, which could account for our findings. However, misunderstanding the task would likely have led to much lower accuracy rates of around 50% rather than the very high rates observed. Also, the experimental task was clearly detailed in the participant information sheet as well as two instruction screens immediately preceding the experiment, with the emphasis on judging the emotion in the face. Furthermore, our experiments did not include a baseline condition, meaning that we cannot conclude whether the context effects observed are evidence of facilitation in the congruent conditions or suppression in the incongruent conditions. However, this is a question that can be investigated in future work. Instead, our experiments successfully met our aims of (1) exploring whether, and the extent to which, spontaneous, incidental aToM inferences act as a contextual influence on the identification of subsequent emotional facial expressions and (2) discovering which theoretical model could better account for our pattern of findings.

The experiments presented here demonstrated that people make inferences about someone else's mental and emotional state that are richer and more detailed than superficial impressions about affective valence. These inferences are extracted from the pragmatic level of a text, and, therefore, our novel findings suggest that language acts as a contextual influence to the recognition of emotional facial expressions for both same and different valences. To borrow the phrasing of the "Reading the Mind in the Eyes Test" (Baron-Cohen et al., 2001), although a general belief exists that people "read" others' facial expressions in order to make inferences about their

emotional states, our findings show that people also "write" such inferences onto the emotional

faces of others.

Acknowledgement

Disclosure statement

The authors report no conflicts of interest.

References

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A.,…Bentin, S. (2008).

   Angry, disgusted, or afraid? Studies on the malleability of emotion perception.

   Psychological Science, 19, 724-732. doi: 10.1111/j.1467-9280.2008.02148.x

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate

   between intense positive and negative emotions. Science, 338, 1225-1229. doi:

   10.1126/science.1224313

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed

   random effects for subjects and items. Journal of Memory and Language, 59, 390-

412.

   doi: 10.1016/j.jml.2007.12.005

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). "Reading the Mind in the

   Eyes" Test Revised Version: A study with normal adults, and adults with Asperger

   syndrome or high-functioning autism. Journal of Child Psychology and Psychiatry,

   42, 241-251. doi: 10.1111/1469-7610.00715

Barrett, L. F., & Kensinger, E. A. (2010). Context is routinely encoded during emotion perception.

   Psychological Science, 21, 595-599. doi: 10.1177/0956797610363547

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of

   emotion. Trends in Cognitive Sciences, 11, 327-332. doi: 10.1016/j.tics.2007.06.003

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using

   lme4. *Journal of Statistical Software, 67*, 1-48. doi:10.18637/jss.v067.i01

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging

   emotion from the face in context. Journal of Personality and Social Psychology,

   70, 205-218. doi: 10.1037/0022-3514.70.2.205

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in

psychological research. Journal of Verbal Learning and Verbal Behavior, 12,

335-359. doi: 10.1016/S0022-5371(73)80014-3

Ferstl, E. C., Rinck, M., & von Cramon, D. Y. (2005). Emotional and temporal aspects of situation

model processing during text comprehension: An event-related fMRI study. *Journal of

Cognitive Neuroscience, 17*, 724-739. doi: 10.1162/0898929053747658

Gendron, M., Lindquist, K. A., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape

perception. Emotion, 12, 314-325. doi: 10.1037/a0026007

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. W. (1992). Do readers mental represent

characters' emotional states? Cognition and Emotion, 6, 89-111. doi:

10.1080/02699939208411061

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed

Emotional Faces: A validation study. Cognition and Emotion, 22, 1094-1118. doi:

10.1080/02699930701626582

Gygax, P., Oakhill, J., & Garnham, A. (2003). Emotion in text comprehension: Do readers infer

specific emotions? Cognition and Emotion, 17, 413-428. doi:

10.1080/02699930244000048

Haigh, M., & Bonnefon, J.-F. (2015). Eye movements reveal how readers infer intentions from the

beliefs and desires of others. Experimental Psychology, 62, 2016-213. doi:

10.1027/1618-

3169/a000290

Hassin, R. R., Aviezer, H., & Bentin, S. (2013). Inherently ambiguous: Facial expressions of

emotions, in context. Emotion Review, 5, 60-65. doi: 10.1177/1754073912451331

Kayyal, M., Widen, S., & Russell, J. A. (2015). Context is more powerful than we think: Contextual

cues override facial cues even for valence. Emotion, 15, 287-291. doi:

10.1037/emo0000032

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., & Whalen, P. J. (2004). Contextual modulation of amygdala responsivity to surprised faces. Journal of Cognitive Neuroscience, 16, 1730-1745. doi: 10.1162/0898929042947865

Lenth, R. V. (2006-9). Java Applets for Power and Sample Size [Computer software]. Retrieved from http://www.stat.uiowa.edu/~rlenth/Power.

Lenth, R., Love, J. & Hervé, M. (2018). Package emmeans. https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Leppänen, J. M., & Hietanen, J. K. (2004). Positive facial expressions are recognized faster than negative facial expressions, but why? Psychological Research, 69, 22-29. doi: 10.1007/s00426-003-0157-2

Lindquist, K. A., Barrett, L. F., Bliss-Moreau, E., & Russell, J. A. (2006). Language and the perception of emotion. Emotion, 6, 125-138. doi: 10.1037/1528-3542.6.1.125

Lundqvist, D., Flykt, A., & Ohman, A. (1998). The Karolinska Directed Emotional Faces – KDEF, CD ROM from Department of Clinical Neurscience, Psychology section, Karolinska Institutet, ISBN 91-630-7164-9.

Nakamura, M., Buck, R., & Kenny, D. A. (1990). Relative contributions of expressive behavior and contextual information to the judgment of the emotional state of another. Journal of Personality and Social Psychology, 59, 1032-1039. doi: 10.1037/0022-3514.59.5.1032

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1, 515-526. doi: 10.1017/S0140525X00076512

Psychology Software Tools, Inc. [E-Prime 2.0]. (2012). Retrieved from http://www.pstnet.com.

R Development Core Team. (2017l). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org

Raftery, A. E. (1995). Bayesian model selection in social research. Sociological Methodology, 25,

111-163. doi: 10.2307/271063

Righart, R., & de Gelder, B. (2008). Recognition of facial expressions is influenced by emotional

scene gist. Cognitive, Affective, & Behavioral Neuroscience, 8, 264-272. doi:

10.3758/CABN.8.3.264

Scherer, K. R. (2001). Appraisal considered as a process of multilevel sequential checking. In K. R.

Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods*

research (pp.92-120). Oxford: Oxford University Press.

Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the Component Process

Model. *Cognition and Emotion, 23,* 1307-1351. doi: 10.1080/02699930902928969

Schwarz, K. A., Wieser, M. J., Gerdes, A. B. M., Mühlberger, A., & Pauli, P. (2013). Why are you

looking like that? How the context influences evaluation and processing of human faces.

Social Cognitive & Affective Neuroscience, 8, 438-445. doi: 10.1093/scan/nss013

Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., Goldsher, D., & Aharaon-Peretz, J. (2005).

Impaired "affective theory of mind" is associated with right ventromedial prefrontal

damage. Cognitive and Behavioral Neurology, 18, 55-67. doi:

10.1097/01.wnn.0000152228.90129.99

Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence

for delay. *Quarterly Journal of Experimental Psychology, 60,* 1680-1696. doi:

10.1080/17470210601160807

Trope, Y. (1986). Identification and inferential processes in dispositional attribution.

Psychological   Review, 93, 239-257. doi: 10.1037/0033-295X.93.3.239

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values.

Psychonomic Bulletin & Review, 14, 779-804. doi: 10.3758/BF03194105

Wieser, M. J., & Brosch, T. (2012). Faces in context: A review and systematization of contextual

influences on affective face processing. Frontiers in Psychology, 3. doi:

10.3389/fpsyg.   2012.00471

Zwaan, R. A. (1999). Five dimensions of narrative comprehension: The Event-Indexing Model. In

S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension,*

causality, and coherence: Essays in honor of Tom Trabasso (pp. 93-110). Mahwah, NJ,

USA: Lawrence Erlbaum Associates, Inc.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and

memory. *Psychological Bulletin, 123,* 162-185. doi: 10.1037/0033-2909.123.2.162

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehenders mentally

represent the shapes of objects. Psychological Science, 13, 168-171. doi:

10.1111/1467-9280.00430

Table 1: Parameter estimates following the linear mixed effects model reaction time analysis for Experiment 1.

	b	SE	t
Intercept	2302.02	211.45	10.887
Vignette Emotion	-45.35	136.91	-0.331
Facial Emotion	-116.17	153.21	-0.758
Interaction	-1399.70	218.17	-6.416*

Table 2: Parameter estimates following the generalized linear mixed effects model accuracy analysis for Experiment 1.

	b	SE	z
Intercept	3.21	0.27	11.677
Vignette Emotion	0.46	0.37	1.263
Facial Emotion	-0.37	0.33	-1.136
Interaction	3.45	0.66	5.236*

Table 3: Parameter estimates following the linear mixed effects model reaction time analysis for Experiment 2.

	b	SE	t
Intercept	1488.07	85.02	17.503
Vignette Emotion	-31.15	50.13	-0.621
Facial Emotion	-71.95	48.87	-1.472
Interaction	-689.66	142.83	-4.828*

Table 4: Parameter estimates following the generalized linear mixed effects model accuracy analysis for Experiment 2.

	b	SE	z
Intercept	4.72	0.54	8.691
Vignette Emotion	-0.001	0.57	-0.002
Facial Emotion	-0.92	0.52	-1.756
Interaction	1.79	1.04	1.718

Table 5: Parameter estimates following the linear mixed effects model reaction time analysis comparing Experiments 1 and 2.

	b	SE	t
Intercept	1897.48	118.32	16.037
Congruency	-531.30	90.57	-5.866
Experiment	819.22	236.63	3.462
Interaction	-376.36	181.14	-2.078*
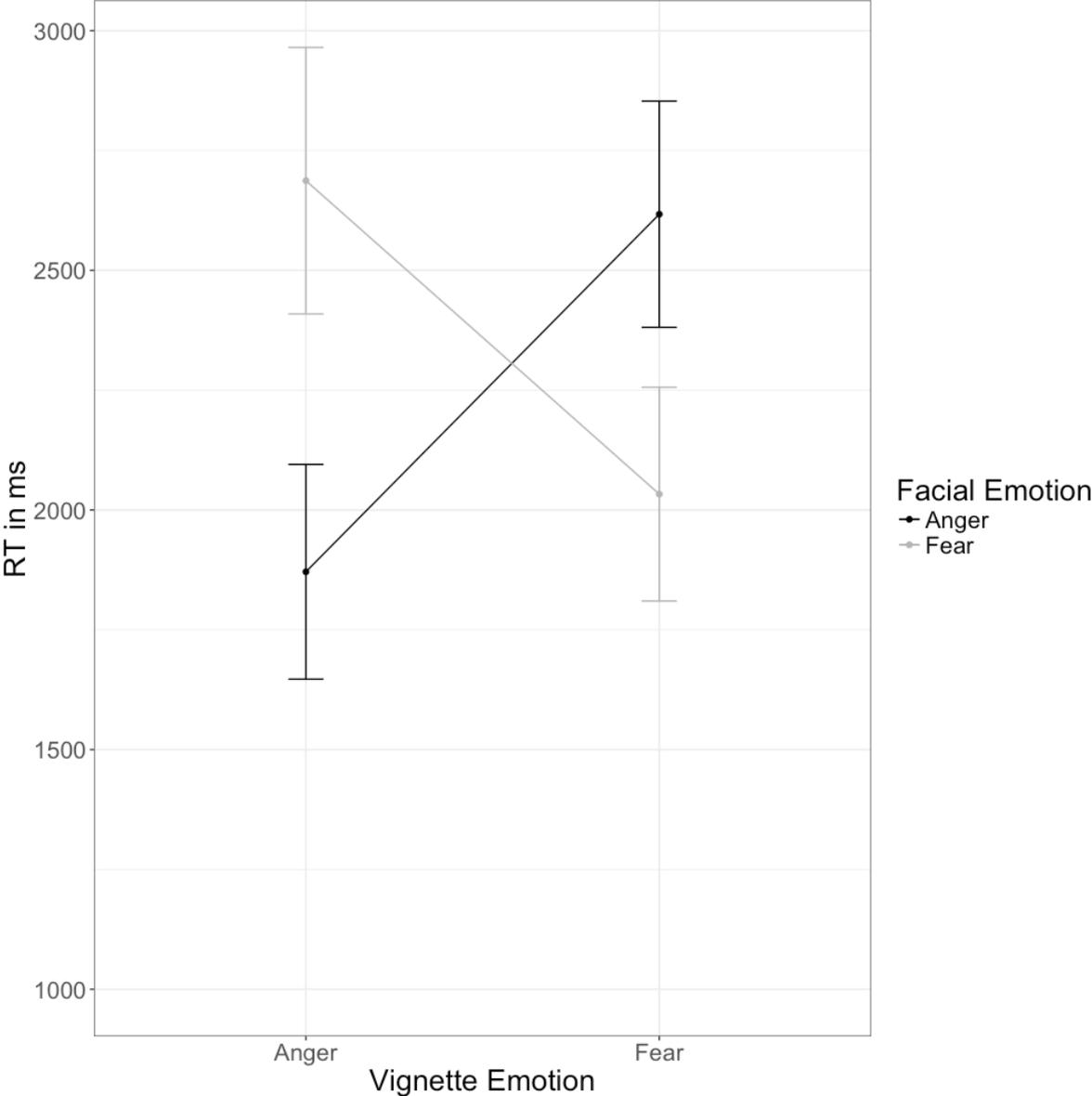
Figure 1

Means and SEs for RTs for Experiment 1
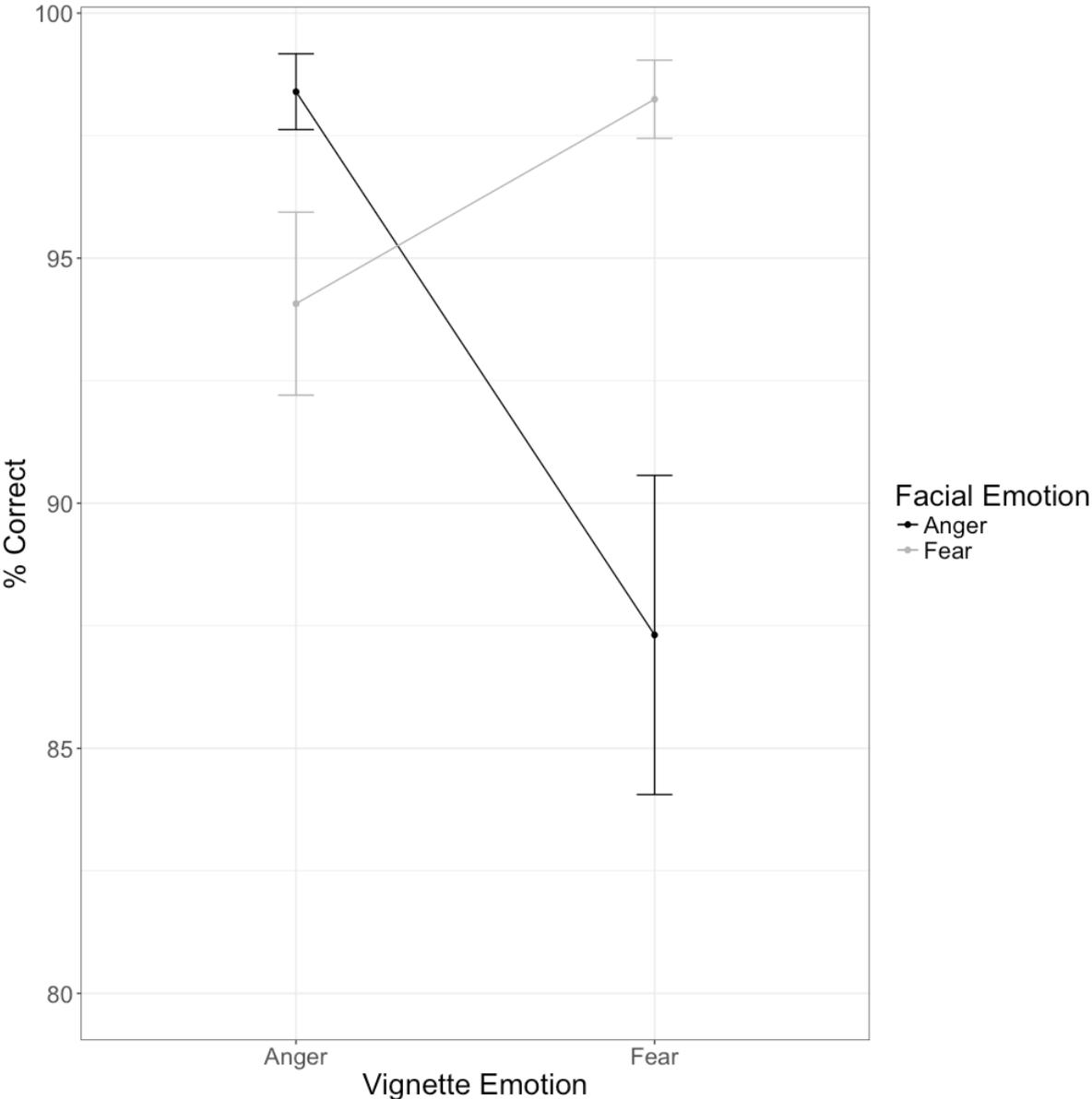
Figure 2

Means and SEs for Accuracy for Experiment 1
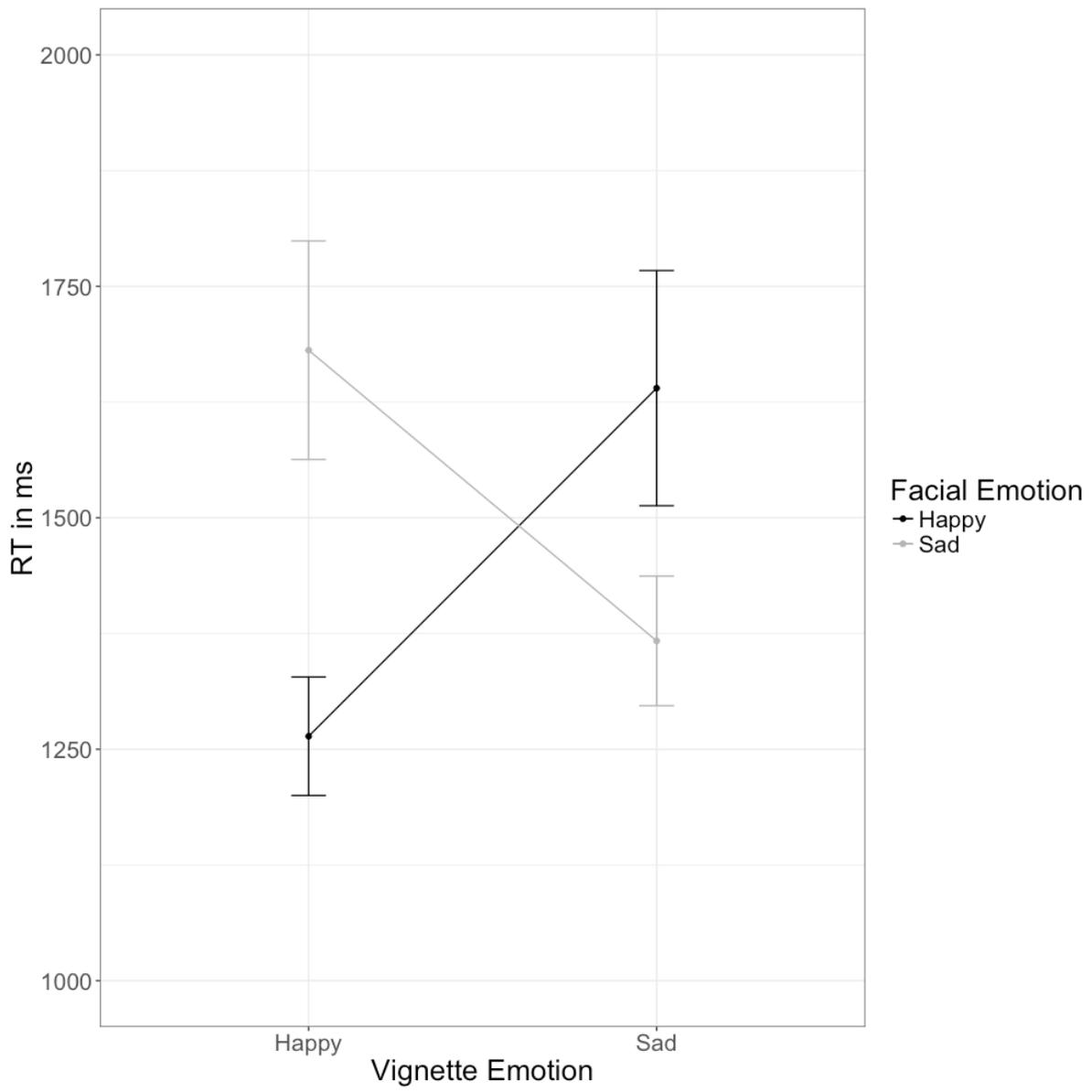
Figure 3

Means and SEs for RTs for Experiment 2
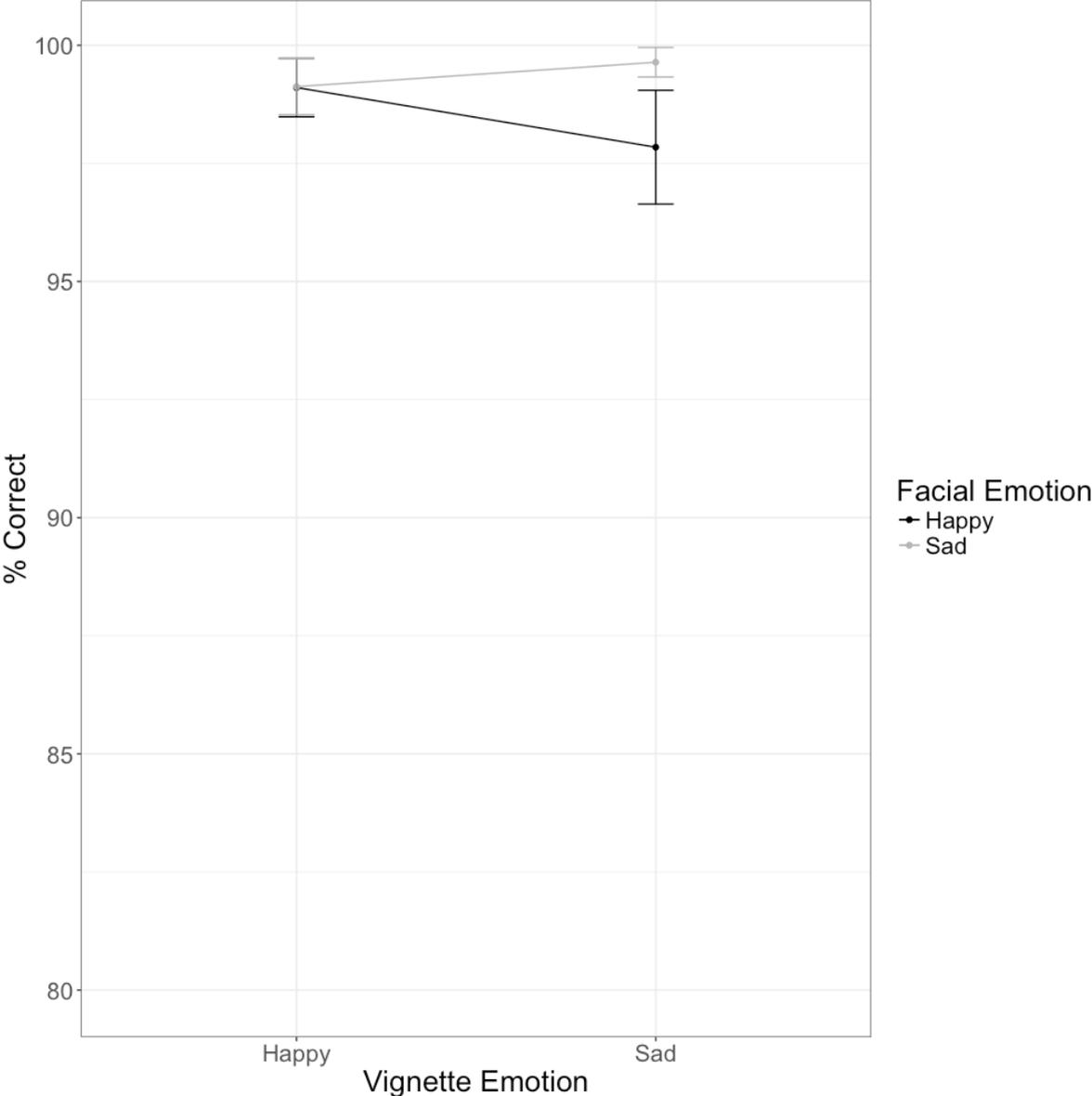
Figure 4

Means and SEs for Accuracy for Experiment 2
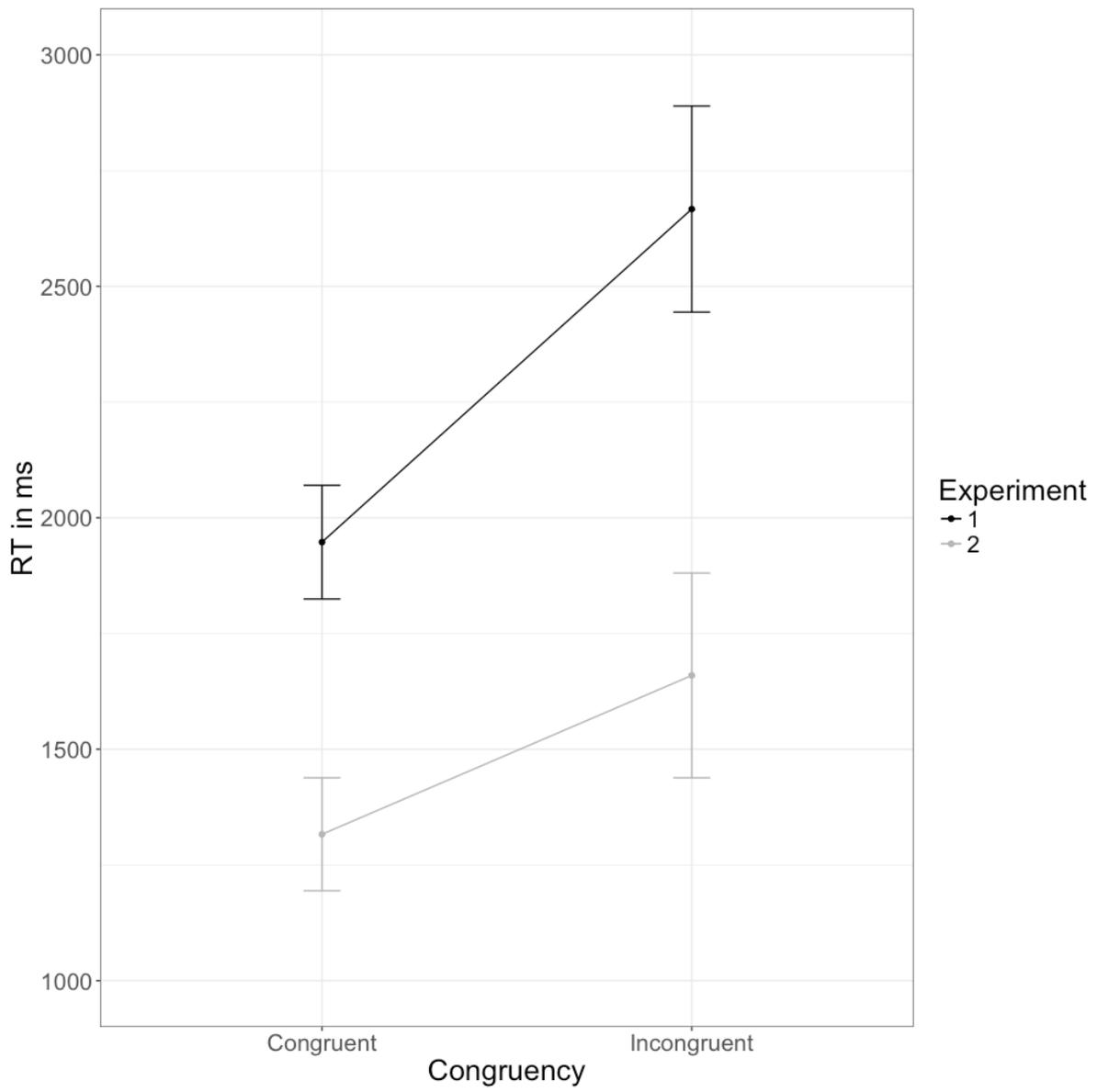
Figure 5

Means and SEs for RTs comparing Experiments 1 and 2

Affective theory of mind inferences contextually influence the recognition of emotional facial expressions

SUPPLEMENTAL ONLINE MATERIAL

Description of vignette development

Experiment 1

During the vignette development phase, 60 vignettes implying anger, fear, or sadness were subjected to an online pre-test with 30 different participants (23 female; mean age = 28.10 years, SD = 9.08 years). The exclusion criteria were the same as for the main experiment. They were eligible to enter the voucher prize draw and earned participation credits where suitable. Participants were asked to read each vignette and select the single best emotion (from sad, angry, afraid, or surprised) that described what the main character was feeling at the end of the vignette. Accuracy rates for the vignettes were calculated; and, for each category, the 16 vignettes with the highest accuracy ratings were selected. In order to determine which two categories were best matched in terms of their implied emotions, the accuracy rates for each vignette category were subjected to a one-way within-subjects ANOVA with emotion as the independent variable and accuracy as the dependent, F (2, 58) = 16.42, p < .001. Pairwise comparisons showed that the sad vignettes had a worse mean accuracy rate (82.24%, SD = 12.63%) than both the angry (95.15%, SD = 7.07%; p < .001) and the fearful (91.85%, SD = 7.74%; p < .001) vignettes. The accuracy rates of the angry and fearful vignettes did not differ, p = .35. Thus, the 32 angry and fearful vignettes were used in Experiment 1.

Experiment 2

In the development phase, 40 vignettes conveying happiness or sadness were subject to an online pre-test in which 30 Psychology students (26 female; mean age = 23.21 years, SD = 9.28 years) took part for participation credits; they did not participate in the main experiment or in Experiment 1 or its pre-test. The exclusion criteria were the same as for Experiment 1. One

additional pre-test participant's data were excluded because his/her participation duration was less than four minutes (other participants' mean participation duration = 20.45 minutes, SD = 9.49), and the researchers agreed that it was unlikely that 40 vignettes could have been thoroughly read in such a short time period. After reading each vignette, participants were asked to select the single best emotion (from surprised, happy, sad, and afraid) that described how the main character felt at the end of the vignette. The 16 happy and 16 sad vignettes with the highest rates of recognition were selected, and these were not significantly different, t (29) = 0.143, p = .89; happy mean = 85.21%, SD = 11.31%; sad mean = 84.79%, SD = 12.47%.

Vignettes for Experiment 1 and Experiment 2

<u>Practice items (both experiments):</u>

(ANGER) Max's girlfriend was passionate but had a big temper. One Saturday, she got into an argument with him and accused him of going after her best friend. Max tried to calm her down, but she only started shouting more loudly. Then, she picked up Max's prize digital camera and smashed it on the floor.

(SADNESS) Matthew was getting married to his long-time girlfriend, and the wedding day had arrived. He and his best man got dressed and got the rings together, and Matthew's parents came to congratulate him and wish him well. Then Matthew's phone rang. It was his wife-to-be, saying that she just couldn't go through with the wedding.

(DISGUST) Claire was a modest person who always tried to be polite. She had recently started University and wanted to meet new people and make friends. She attended a party at the student

union pub and fell into conversation with a few other new students. One of the others told a very crude and racist joke, and everyone laughed except Claire.

Experiment 1

Experimental items:

FEAR MALE:

Tom was walking home from his Saturday evening shift as a chef in a busy city centre restaurant. The quickest route home took him through some alleyways, but they were usually quiet. On this evening, he came upon two men in an alley who were arguing and holding broken bottles. Then the men saw Tom and began walking towards him.

Simon took his new girlfriend out to a National Park for a long walk. After a while, they lost their way and ended up missing lunch. Simon's new girlfriend began to feel faint and started to look pale. She explained that she was diabetic and that she had left her insulin back in the car, which was nearly a mile away.

Allan was looking after his 3 year old nephew for the day, so he took the little boy to visit the beach. The beach was very crowded when they arrived so they found a spot near the seafront to make sand castles. Allan bent down to fill a bucket with sand. When he stood up, he did not see his nephew anywhere.

Richard was out late one night at a pub when he realised he had run out of money. He told his

friends that he was just going out to a cashpoint. There were very few people around. As Richard

was typing in his PIN, he heard footsteps behind him and very loud breathing.

Henry's best friend had bought a new car and asked Henry to come along for the first drive. As they

got onto the motorway, the friend moved into the fast lane and said he wanted to see how fast the

car could go. Henry said there were too many other cars on the motorway and that it would be really

dangerous. Henry's friend laughed and pushed the accelerator to the floor.

Mike went to the supermarket one evening for his weekly shop. As he went around the aisles, he

noticed a man in a black hoodie who always seemed to be close by. On the last aisle, Mike looked

straight at the man who immediately turned away. As Mike left and walked through the dark car

park, he noticed the man walking some distance behind him.

David was at a large political rally, protesting student fees. The man standing next to him had

gotten increasingly loud and agitated over the past hour. As David looked on, the man started

pushing a police officer. Others started to join in, yelling and shoving the other officers.

Dan had joined the university rugby team at the beginning of his first semester. When he joined, the

older boys told him he and the other freshers would have to go through initiation tests before he

could play in a match. Dan had heard that the initiations involved a lot of alcohol and being made to

strip, and that most years, someone ended up in hospital. When he arrived for his initiation, he was

told that he would have to walk along a railway track for a mile without stopping or getting off.

FEAR FEMALE:

Lindsey had been on her way to a party, when she realised she was lost in an unfamiliar part of town. It was dark, and there were very few people around. She saw a man standing at a bus stop and went to ask directions. When she approached him, she saw that he was cleaning a large knife.

Amanda had arranged to meet her friend at a local pub. She arrived early and decided to wait outside the pub. Just then, a man came and stood very close to her and began asking her all kinds of questions. Amanda told him that she had to go and she started walking off home. As she walked away, she noticed that the man was following her.

Vicky and her boyfriend had stayed out too late to catch a bus, so they decided to walk home instead. Vicky's boyfriend decided to climb over a high wall as shortcut, even though Vicky had said they should walk the long way around. When Vicky's boyfriend was at the top, he slipped and fell to the ground on the other side. She heard a loud cracking sound as he landed.

Liliana had to work late one bank holiday. When she was finished, she realised that there were no buses running and she would have to walk home. The shortest route was through a park, but it was dark and there were hardly any other people around. As she crossed the park, she heard heavy footsteps coming up behind her.

Michelle and her brother had been out late to a party, and they were very drunk as they walked home. Michelle stopped to tie her shoe while her brother kept walking. As she looked up, she saw her brother stumble towards the busy road. She shouted as her brother stepped out into the road.

Lisa was working late at the library one night. There was hardly anyone around, but then a guy came and sat down right at her table. A few minutes later, Lisa got up to use the toilet. When she came back, the guy started to stare at her.

Gemma and her boyfriend were apart for a few days as he had travelled to London with a group of friends. They were all interested in archaeology and had tickets to a special exhibition that day at a large museum. Gemma sat down in her living room to watch some telly. When she turned on the news, she heard the newsreader report that there was a bomb threat to the museum and that special police units were going in.

Janine and her friends were out at a pub late one night when a very loud group of drunk men came in. They shouted at the barman and made some lewd comments to some women sitting near the bar. Janine and her friends decided to go. As they stood up and grabbed their coats, one of the men turned around and headed towards Janine's group.

ANGER MALE:

Tim was in a band that was doing a small tour around Europe. They took a flight to Paris, where their first gig was going to be the following evening. At baggage claim, Tim's guitar never came out. The airline representative told Tim the guitar would arrive three days later and they could not compensate him for the delay.

Steven purchased a new bike after saving for months to get the best bike he could afford, as he was a keen cyclist and frequently cycled into work. One morning, Steven locked his new bicycle in the bike shed outside his workplace. When he returned after work, he found that all the bicycle seats had been stolen. Steven had to push his bike for two miles to get home.

After several months of unemployment, Noah decided to use a recruitment agency to improve his chances of finding a job. He forwarded his CV to them and paid a large upfront fee. After a few weeks of no contact, Noah telephoned the agency. The woman told him that it appeared that his CV had been lost and, thus, he had not been put forward for any jobs.

John lived in a shared house while at university with four other lads. Every six months the landlord came round for an inspection of the property, so they all pitched in and cleaned the house from top to bottom. When the next inspection day arrived, John's housemates filled his room with all the rubbish they had collected. When the landlord saw the mess, he told John that he was terminating his contract.

George was interested in his family history, and his prized possessions were his great-grandfather's World War I medals. One day, George went to his room to find the medals bent and broken off their

chains. George's flatmate told him he had no idea what had happened. A few weeks later, the flatmate told George that he had broken them when he was showing some friends.

Sean took his elderly mother to the local supermarket every Saturday for her weekly shop. The car park was very full and Sean had to drive around a few times until he noticed someone leaving. As Sean was waiting for the car to move out of the way, another person drove around Sean's car and into the parking space. When Sean honked his horn, the other driver didn't move.

Kevin's flatmate was unemployed and was really struggling with money. Kevin had offered to cover his rent for a while, and had even starting giving some money to the flatmate to buy some food instead of putting it into his savings. One day, the flatmate came home from the shops with two big bags. When Kevin looked in, there was no food but several DVDs, some alcohol, and a computer game instead.

Craig was waiting for his sister to give him a lift to an important doctor's appointment, but she was late. The GP's surgery was oversubscribed and appointments were always difficult to come by. When Craig phoned his sister to see where she was, she said that she had only just woken up and would be there in thirty minutes. Craig missed his appointment.

ANGER FEMALE:

Bernadette worked in a busy city centre shoe shop during the summer break from university. Her manager asked her to reorganise the stock room into alphabetical order to help other staff members find stock quickly. Bernadette spent two full days organising the stock. When she returned to work after a day off, she found that the stock room was a huge mess and that other staff members were putting stock in all the wrong places.

Stacey had waited for months for the release of a new science fiction film. While waiting in the queue with her younger brother to buy tickets for the 3D showing, Stacey saw three men push into the line in front of her. When Stacey got to the counter, the cashier told her that he had just sold the last tickets to the three men. Stacey and her brother had to wait until the next day.

Louise had arranged for a repairman to service her boiler as it was faulty. Because the repairman was unable to specify an appointment time, Louise had to take the whole day off from work as annual leave. Late in the day, the repairman still had not come so Louise phoned his company to ask when he might turn up. The man on the end of the phone laughed and told her that the repairman had taken the day off.

Jacquie had an important interview, and her flatmate was going to drive her there. As she got ready, Jacquie noticed that the flatmate was only just getting out of bed. Jacquie reminded her flatmate that she had to get there in a hurry and did not want to be late. The flatmate replied that he'd forgotten to tell Jacquie that his car had broken down a few days ago and he wouldn't be able to drive her to the interview.

Steph really wanted to buy the new iPhone on its release day and had camped out overnight be first in the queue. As the opening time approached, Steph realised she couldn't wait any longer to use the loo. The guy behind her promised to hold her spot, but when she came back, the guy was sitting on the pavement playing with his new phone while hundreds of others flooded into the shop. He laughed and wished her better luck next time.

Emma had started a romantic relationship with a colleague at her new job. As Christmas was fast approaching, Emma and her colleagues decided to go into town for a few festive drinks. Throughout the evening, Emma noticed that a female co-worker kept sitting close and touching her boyfriend's thigh. The boyfriend kept stroking the woman's hair.

Violet spent a lot of money on a beautiful new dress for the uni ball. When she got home and tried it on with her new shoes, she discovered a big yellow stain on the back of the dress. She brought it back to the shop and showed the shop manager. The manager said the dress had not been like that when Violet had bought it and refused to give her a refund.

Lucy worked part-time for a local newspaper and had been working on a big story about a campaign to save the historic town hall. She had even worked overtime and had spent her own money to interview lots of residents and do all the research. Her editor praised Lucy for all her hard work and told her it would be on the front page. When Lucy bought the paper the next day, she saw her editor had put his own name on the report.

<u>Experiment 2</u>

Experimental items:

HAPPY FEMALE:

Danielle entered a national poetry contest. One day a letter arrived in the post which said that she was a finalist and was invited to London to read her poem to the judges of the final round. At the event, all the final poems were very good. After the voting was over, the chief judge approached her and told her that she had written a wonderful poem that had really moved all the panel members.

Leigh had been helping her son study for his GCSE maths exam, as it was a subject that he really struggled with. They studied and worked together for nearly two hours every day in the weeks leading up to the exam. In August, Leigh's son went to his school to find out his results while she waited at home. Leigh's phone rang, and it was her son telling her that he had received a B.

Nicole and her young daughter decided to go to the zoo. Once they were there, they walked around and saw lots of interesting animals. They talked about why the animals all looked so different, and then they had a picnic lunch near the giraffes. Finally, they shared some ice cream before heading home, hand in hand.

Vivian's family was having a big family reunion weekend. It had been longer than anyone could remember that all the extended family had been together. Vivian hosted the first event, a barbecue, at her house; and everyone stayed late into the night, sharing stories and talking. Vivian herself spent a long time chatting with everyone, catching up and getting to know her relatives better.

Winnie was the captain of her University's quiz show team. The team had practiced for months and knew each other's strengths and weaknesses well. The final round was tough, but they managed to work out many difficult answers together. When the final bell rang, the host announced that Winnie and her team had won by five points.

Hannah and her mum shopped together and had lunch together once a month. This past Saturday, her mum took Hannah to a special restaurant for a particularly fancy meal. Hannah offered to pay for half, but her mum wouldn't accept the offer. Her mum said she was so proud of Hannah's achievements at work recently that she wanted to give her a special treat.

Laura and her partner were looking for a house to buy. They had seen a lot, but none had felt right. At the end of a long day, they had arrived at the last house they were to view. As soon as Laura and her partner stepped in, Laura knew it was the right house. The owner described all the wonderful memories she'd had in the house and said she hoped whoever bought it would make lots of memories there too.

Becky lived in the centre of London, where it rarely snowed. One week after Christmas, it was very cold and snow blanketed the area. Becky heard some shouting outside and saw some of her neighbours having a snowball fight in the street. She headed outside where she joined in with everyone, and later they made a snowman and some snow angels too.

HAPPY MALE:

Johnny was on the University's football team. They had recently won an important match because Johnny has scored the winning goal. The team decided to go out to the pub to celebrate. Johnny's team members congratulated him and bought his drinks all evening.

Andrew was in danger of failing his introductory history module. He had decided to become more studious and had seen his tutor on a few occasions to discuss how he could improve his marks. When it was time for the next hand-back, the tutor asked Andrew to come by her office. She told him that he had received a 65.

Grant's University had its own weekly student newspaper, and he had worked as a reporter and copy editor on it for two years. He decided to apply to become the newspaper's editor when the position became open in his final year. During the interview, he spoke passionately about his journalism work. A few hours later, the panel telephoned him to say that he was going to be the new editor.

Joe and three of his friends were competing in a relay marathon run for charity. Joe had the last leg and as he approached the finish line, he looked at the race clock. He saw that they were going to set a new best team time. His teammates ran onto the route screaming wildly and proceeded to run the last hundred metres with him.

Thomas attended the induction events at his new University. The students' union put on a treasure hunt around campus and randomly assigned the students to teams. Thomas and his team got on well, helping each other out with the clues as they raced around the campus to the various locations.

They didn't win, but they all planned to meet that night for drinks, and Thomas exchanged phone numbers with all of them.

Harry had a close-knit group of friends from school, with whom he still kept in touch even after many years. They agreed to meet for a weekend in their hometown after several years of not seeing each other. As Harry got off the train, he saw his three friends on the platform, talking and laughing together. As they noticed him, they all shouted to him and started waving.

Ethan had planned a special day for his girlfriend's birthday. He collected her in the afternoon with a bouquet of roses. Then, he took her rowing on the river. Finally, he took her to a cocktail bar and a fancy restaurant. As they walked home, she told him that she loved him.

Mitch and his son often played football together. They went to the park every Saturday afternoon to play for a few hours. One Saturday, Mitch's son did a particularly impressive header which Mitch praised. When they were leaving, Mitch's son told him that he'd had a really great time.

SAD FEMALE:

Jessica was interested in drama and decided to audition for the drama society's new play. She practiced a monologue and even had her friends critique her acting. At the audition, Jessica gave it her all. When the audition was over, the drama society thanked her for coming but said that she would not receive a role in the play.

Helen had a favourite independent coffee shop that she visited about every other day. She had gotten to know the shopkeepers well because usually there were only a few other customers in the shop. One day, Helen saw a sign which said that the coffee shop would be closing in a week. When she enquired about it with the shopkeeper, he told her that they just were not making enough money and had to close.

Jane and her brother had always been very close. While Jane had stayed close to home, her brother had lived in Spain for several years. Her brother was going to arrive the next day for a long visit, and it was the first time in a year that they would see each other. That evening, Jane's brother rang to say that he had to postpone his trip for a couple of weeks because his son had taken ill.

Melissa and her best friend were very close and always did everything together. They had even decided to go to the same University, although they studied different courses. While Melissa had found her work and exams to be okay, the friend had always struggled. One day, the friend told Melissa that she had failed too many times and that she had decided to withdraw.

Sarah and her best friend wanted to rent a flat together for their final year at University. Although Sarah's parents supported her financially, her friend was almost always broke. They began searching for cheap flats. But, the next time they met for coffee, the friend told Sarah that she didn't have enough money to help rent the flat and that she was going to live at home.

Heather had volunteered at a summer holiday camp for children. She had gotten to know lots of the children well and had made good friends with the other volunteers. Now that it was the end of the

summer, everyone was packing up to go home. She watched as the children that she had looked after found their parents and climbed into their cars.

Alice's grandmother was very old and had lived in a nursing home for the past five years. Alice visited her grandmother weekly, and they often played card games together. Recently, the grandmother had developed pains in her legs. The next time that Alice visited, she saw that her grandmother could not get out of bed.

Joanna played doubles tennis with her best friend. They made a good team, but Joanna had injured her shoulder badly in a match last year. Although she continued to play, her shoulder was hurting more and more. After having a chat with her friend and seeing her GP, Joanna and her friend decided that it was better for her health that Joanna stop playing.

SAD MALE:

Oliver was a caring person and got to know a few of the Big Issue sellers around town. He always stopped to have a chat and buy a new issue. One day he realised that he had not seen one of the guys for quite a while. He inquired about him with one of the other sellers, who said that the man had caught pneumonia and was very poorly in hospital.

Sam had been unemployed for some time despite sending out hundreds of job applications. He lived frugally although sometimes he did not have money to buy enough food. A few days ago, he had had a job interview that seemed to go very well. When his phone rang, it was the chief of the

interview panel saying that Sam was an impressive candidate but that they had decided to give the job to someone else.

Adam and his wife were getting divorced. It was hard on everyone, particularly their ten-year old son. One afternoon, Adam found his son crying in his bedroom. When he asked what the matter was, the son said that he wanted to live with his mother.

Ian and his girlfriend had been together for a long time, but they had been having a difficult time recently. They had argued a lot but had tried to stick together. They had even been to see a relationship counsellor. The next time Ian saw his girlfriend, she said that she wanted to split up.

Aaron and his wife had always wanted to have children. They had been trying to get pregnant for a couple of years but without success. They had recently been to get their fertility tested. When they visited the clinic to get the results, the doctor told them that Aaron had a fertility problem and that they would not be able to have children.

Steve had become good friends with his neighbour, and they had lived next door to each other for the past two years. Recently, the neighbour had gotten a new job in a different city and Steve had helped him pack up his house. Steve had even held a goodbye party for him. The next morning, Steve was looking out the window when he saw the removal men pull up.

Neil and his partner had been spending less and less time together. They were both often very busy with work, and his partner was spending more time with her friends. When Neil and his partner

would go out together, there were often long silences. The next time they went out together, Neil's partner said that they should see other people.

James decided to enter a crossword contest run by a national newspaper. The crossword was very difficult, and he had to work for several days to solve it. He sent his solution to the newspaper; and a few weeks later, he received a reply. The letter said that he had solved all but two of the clues correctly but that he had not won the contest.

FILLER MATERIALS

DISGUST

Joseph was starting a new job for the summer, and he was keen to get to know his co-workers. At lunchtime, the guy at the next desk invited Joseph down to the canteen with him. They got their trays of food and sat down. Joseph watched as his co-worker ate his entire lunch with his hands, dropping food down his shirt and licking his fingers loudly.

Denise went on a date with a guy she had met at a party. She had taken care to look her best. They went to a nice restaurant and had a meal of three courses and a bottle of wine. Right after finishing his dessert, Denise's date belched loudly without covering his mouth.

Jim worked as a personal shopper in men's clothing in a big department store. One day, an older man came in for his hour-long appointment. Jim noticed that he had dried food all over his front and greasy hair. As the man approached, Jim realised that he smelled very strongly of body odour.

Grace and her sister lived far apart and rarely saw each other. Now, Grace's sister was coming to stay for the weekend and she had planned a fun weekend that included attending a party with all of her friends. At the party, Grace's sister drunk a lot and started swearing loudly. Then on the walk home, she vomited all over Grace's shoes.

Liz had a friend who loved to play practical jokes. She had told him that she did not like practical jokes and did not think they were very funny. Liz's friend invited their group around for dinner one evening. As Liz stirred her soup, she saw a big hairy spider float to the surface.

Claudia and her friends had booked a holiday to America. On the plane, they had to split up and Claudia had ended up sitting next to an older man. He drank throughout the flight and had eaten so messily that he had spilled gravy all over Claudia's bags. Then he fell asleep on Claudia's shoulder, snoring loudly and drooling all over her shirt.

Jasper had worked in a number of restaurants as a chef and was now shortlisted for a job at an exclusive restaurant. He had worked hard to create a new dish that would impress the head chef at his interview. When he brought out his dish for the chef to try, he saw that the man was unsteady on his feet and smelled heavily of alcohol. Then the man sneezed all over Jasper's dish and promptly passed out.

Eric was close to his grandmother, who was getting very old and frail. In fact, she had difficulty walking and so hardly left the house any more. Eric decided to visit her for a couple of weeks over

the summer. When he walked into her house, Eric discovered rubbish piled high all over, including rotting food on the kitchen surfaces.

SURPRISE:

Jennifer and her ex-flatmate had separated on bad terms after having a big row about rent payments. Jennifer had moved to a new place and had not seen her ex-flatmate in four years. One night, Jennifer and her friends went to a birthday party. She realized that there were only six guests and one of them was her ex-flatmate.

Harriet had stayed late at her job one night to finish off some work. Her co-workers had all gone a couple of hours before. When she took a break, she decided to turn on some loud music and sing along. As she was dancing around, she turned around to see her boss standing in the doorway.

Julia had arrived at the city's theatre as she was an understudy for a new play. She had gotten there very early so she could do her makeup, do some warm-up exercises, and practice her monologue a few times. A few minutes after she had arrived, the director's assistant found her backstage. He said that the lineup had changed and that she would be going on in five minutes.

Henrietta was out in the middle of town one Saturday when a reporter for the local newspaper approached her. He asked if he could interview her about some of the planning controversies going on in the town centre. Henrietta told him what she thought and even let her picture be taken. When the newspaper arrived at her house the next week, she saw that her picture was on the front page, with one of her quotes even used as the main headline.

Marnie and her friends decided to plan an amazing night out after their exams had finished. They had had a stressful few weeks, but now that the night had arrived, Marnie took her time to look her best. She was running very late, so she made took a taxi to the pub rather than take the bus. As she looked around the pub, Marnie realised that her friends were not there.

Anna was at home one Saturday just having a lazy day. Suddenly, her doorbell rang. When Anna answered the door, she saw a tall woman with long dark hair. The woman apologised for bothering Anna and then said she was researching her family background and discovered that Anna was her long-lost younger sister.

Sadie's best friend had recently left a long relationship. At Sadie's urging, the friend had gone out and had started to date a few guys. Sadie was now going to meet the friend's new boyfriend. When Sadie arrived at the friend's house, the friend announced that she had eloped was now married.

Jolene was riding the bus to work early one morning. The bus was very crowded, but most people were by themselves and were very quiet. At the next stop, several people got on and a tall man had to stand next to Jolene. He turned to her and then suddenly burst into an operatic song.

Paul's girlfriend had always been shy and meek. She always opened up with Paul, but other people thought she was mousy. He knew she was going through a difficult time and he tried to be supportive. The next time Paul went to her flat, he saw that she had gotten a large tattoo of a dragon across her back.

Seb enjoyed football but he had never had much talent for it. Even so, he had joined the University's team as a trainer and assistant. He enjoyed being close to the action and helping out his teammates with water and stretches during the match. At the next match, the manager told him that so many of the players had the flu that Seb would have to play in order to make up a full team.

Christopher was out pulling some weeds in his garden on a hot day. He had his headphones on and was singing and dancing as he worked. He turned around to put some more weeds in his wheelbarrow. It was then that he saw his neighbour peering over the fence, watching him and laughing.

Joel had always been very close to his parents, particularly as he was an only child. He had recently moved out to a flat of his own, but he saw his parents regularly. One Sunday when he visited for lunch, his parents asked him to sit down. They apologised for not discussing it before and then told Joel they loved him but that he was adopted.

Aidan's parents had split up when he was a baby, and he had not seen his father since he was a toddler. Eventually, his mother had remarried and Aidan was very close to his stepfather. One day, the doorbell rang and an older man was standing on the step. Aidan realised that the man was his father.

Alex was an only child and didn't have many living relatives. One day, when he was visiting his parents, they got a visit from a lawyer. The lawyer explained that she was executing the will of

someone who turned out to be a distant aunt who had amassed a large fortune. The lawyer then said that the aunt was leaving Alex and his parents ten million pounds.

Nathan's flatmate had been away visiting his family for a week. Although Nathan was working late a lot, he'd made good use of having the flat to himself – listening to music loudly, watching whatever he wanted on TV, and the like. One evening when he returned from work, he noticed that a light was on in the flat. When he unlocked the door, he saw his flatmate was back and had invited over dozens of people, who were all dancing, drinking, and generally trashing the flat.

Jonathon and his wife had wanted children for a long time, and finally they were pregnant. They didn't have much money, but they planned to give their child the best life they could. They went to the hospital to have the first scan. After a few minutes, the technician told them that they would be having triplets.

Re-analysis of Experiments 1 and 2 to discover any effect of lexical priming

Introduction

One possible explanation for the results is that the effects are being driven by a subset of vignettes that contain individual words that prime one or the other of the emotional response options. This would mean that participants were effectively responding to a lexical prime in some vignettes rather than to the subsequently presented facial emotion. To address this alternative explanation, we re-ran our analyses for both experiments in which we removed data that was derived from vignettes that contained possible cue words for our emotional response options.

To identify relevant possible lexical cue words for "angry," "fearful," "happy," and "sad," we utilised the Small World of Words free association English database (SWOW; S. De Deyne, personal communication, 15 January 2018; see De Deyne, Navarro, & Storms, 2013, for the procedure of constructing the SWOW Dutch database). The SWOW database offers the advantage over similar databases (e.g., the University of South Florida Free Association Norms; Nelson, McEvoy, & Schreiber, 1998) that the associations have been more recently collected and are, therefore, more up-to-date. Second, these data were collected from English speakers globally, meaning that the associations are less reflective of any particular dialect.

For the SWOW database, four samples of one hundred participants each were asked to generate three responses that they freely associated with "angry," "fearful, "happy," or "sad." Because the item words that SWOW participants are asked to respond to are generated randomly, it is possible that some participants may have appeared in more than one of these samples. Thus, the SWOW database gave us an initial pool of three hundred lexical associations for each emotion response option. For each pool, we calculated frequencies for each cue (regardless of whether the cues were first, second, or third associations). We then discarded all cues that had a frequency of one as these may have reflected idiosyncratic associations. Thus, all cues that we then used had at

least two SWOW participants associate that cue with the relevant emotional response option. We then identified the vignettes that contained any of these cue words in order to exclude them from the re-analysis. If a vignette contained (for example) an adjectival or adverbial form of a noun cue, we followed a rule that that vignette should also be excluded from the reanalysis. The searched cue words and the vignettes in which any were found (as identified by the main character's name) can be found in Tables 1 and 2. Four vignettes were excluded from the re-analysis of Experiment 1 (all fearful) and five from Experiment 2 (three happy and two sad).

(Table 1 about here)

(Table 2 about here)

Results for re-analysis of Experiment 1

Initially, we utilised arousal as a covariate but the pattern of results were the same: the models with and without arousal were not significantly different from each other and the key interaction remained significant in the model using arousal. Therefore, here we fully report the re-analysis without arousal as a covariate. In our LMM analysis for the RT data, the fixed effects were Vignette Emotion (Fear, Anger), Facial Emotion (Fear, Anger), and the interaction between these factors. We used deviation coding for each of the experimental factors. Our model contained crossed random effects for participants, vignettes, and faces. The model with the most maximal effects structure that converged included random intercepts and additive slopes for both fixed factors by participants, and by vignettes, and random intercepts and slopes for the Facial Emotion factor by faces. Restricted maximum likelihood estimation was used when reporting the linear mixed model parameters (see Table 3 for parameter estimates). The model revealed an interaction between Vignette Emotion and Facial Emotion that was significant at <.001 alpha level (estimated by approximating to the z-distribution).

(Table 3 about here)

The interaction was explored with pairwise comparisons performed using the emmeans package in R (Lenth, Love & Hervé, 2017) with degrees of freedom approximated using the Kenward-Roger method. The pairwise comparisons were interpreted using a Bonferroni-corrected alpha level of .025. Table 4 contains the marginal means and standard errors calculated using the emmeans package. They show that RTs were slower for identifying a facial emotion when it mismatched the vignette emotion. Angry faces were recognised faster after angry vignettes than after fearful vignettes, $t(21.73) = 3.491$, $p = .002$, while fearful faces were recognised faster after fearful vignettes than after angry vignettes, $t(25.76) = 4.407$ $p = .002$.

(Table 4 about here)

In our GLMM analysis for the accuracy data, the fixed effects were Vignette Emotion (Fear, Anger), Facial Emotion (Fear, Anger), and the interaction between these factors. Our model contained crossed random effects of participants and vignettes. The random effect of faces was dropped in order to arrive at a model that converged. The model with the most maximal effects structure that converged included only random intercepts by participants and by vignettes (see Table 5 for parameter estimates). The model revealed an interaction between Vignette Emotion and Facial Emotion that was significant at < .001 alpha level (based on the z-distribution).

(Table 5 about here)

The interaction was explored with contrasts performed using the emmeans package in R on the log odds ratio scale. The pairwise comparisons were interpreted using a Bonferroni-corrected alpha level of .025. Table 4 contains the marginal means and standard errors calculated using the emmeans package. Accuracy decreased when identifying facial emotions that mismatched the preceding vignette emotions. Angry faces were responded to with higher accuracy after angry vignettes than after fearful vignettes, $z = 4.101$, $p < .001$, while fearful faces were responded to with higher accuracy after fearful vignettes than after angry vignettes, $z = 2.797$, $p = .0052$.

Results for re-analysis of Experiment 2

A LMM was fitted to the RT data, following the same procedure used the re-analysis of Experiment 1. The fixed effects were Vignette Emotion (Happy, Sad), Facial Emotion (Happy, Sad), and the interaction between these factors. We used deviation coding for each of the experimental factors. Our model contained crossed random effects for participants, vignettes, and faces. The model with the most maximal effects structure that converged included random intercepts and additive slopes for both fixed effects by participants, by vignettes, and by faces. Restricted maximum likelihood estimation was used when reporting the LMM parameters (see Table 6 for parameter estimates). The model revealed an interaction between Vignette Emotion and Facial Emotion that was significant at <.001 alpha level (estimated by approximating to the *z*-distribution).

(Table 6 about here)

As with Experiment 1, the interaction was explored with pairwise comparisons performed using the emmeans package in R (Lenth et al., 2017) with degrees of freedom approximated using the Kenward-Roger method. The pairwise comparisons were interpreted using a Bonferroni corrected alpha level of .025. Table 4 contains the marginal means and standard errors calculated using the emmeans package. Happy faces were recognised faster after happy vignettes than after sad vignettes, t (14.25) = 5.595, p < .001, while sad faces were recognised faster after sad vignettes than after happy vignettes, t (14.78) = 4.849, p < .001.

In order to calculate the Bayes Factor to compare our RT Model 1 (i.e., with additive effects of Vignette Emotion and Face Emotion and with the interaction between them) to RT Model 2 (i.e., with additive effects of Vignette Emotion and Face Emotion, but with no interaction between them) we had to simplify the random effects structure; for both models we had random intercepts and slopes involving additive effects of Vignette Emotion and Face Emotion for participants and vignettes, and faces. This produced a Bayes Factor (BF12) of 3,269,017 in support of Model 1 over Model 2.

In our GLMM analysis for the accuracy data, the fixed effects were Vignette Emotion (Happy, Sad), Facial Emotion (Happy, Sad), and the interaction between these factors. Our model contained crossed random effects of participants and vignettes. The random effect of faces was dropped in order to arrive at a model that converged. The model with the most maximal effects structure that converged included only random intercepts by participants and by vignettes. The model revealed a significant interaction between Vignette Emotion and Face Emotion, p = .0443 (see Table 7 for parameter estimates).

(Table 7 about here)

The interaction was explored with contrasts performed using the emmeans package in R (Lenth et al., 2017) on the log odds ratio scale. The pairwise comparisons were interpreted using a Bonferroni-corrected alpha level of .025. Table 4 contains the marginal means and standard errors calculated using the emmeans package. Neither comparison reached statistical significance, suggesting that the effect observed in this re-analysis is very weak. Happy faces were not responded to with higher accuracy after happy vignettes than after sad vignettes, z = 1.453, p = .1462, and sad faces were not responded to with higher accuracy after sad vignettes than after happy vignettes, z = 1.337, p = .1811.

The Bayes Factor comparing Accuracy Model 1 (i.e., with additive effects of Vignette Emotion and Face Emotion and with the interaction between them) to Accuracy Model 2 (i.e., with additive effects of Vignette Emotion and Face Emotion, but with no interaction between them) was calculated to be 2.23 (BF21) in support of Model 2 over Model 1.

Summary of re-analysis

The RT and accuracy data were re-analysed after excluding vignettes that contained possible cue words for our emotional response options in order to investigate the possible explanation that a subset of vignettes were driving the effects through lexical priming. The pattern of results was largely the same. Furthermore, the Bayes Factors for the Experiment 2 re-analysis demonstrated positive support for a congruency effect in RTs (Model 1) and nearly equivocal evidence for a *lack* of a congruency effect in accuracy, which was the same overall pattern as was found in the main analysis. Thus, the pattern of these Bayes Factors continues to support the extended language-as-context account (which would be supported by evidence for RT Model 1 regardless of whether Accuracy Model 1 or 2 is supported) over the limited situational dominance account (which would be supported by evidence for RT Model 2 and subsequently by Accuracy Model 2). It should be noted in particular that the effects demonstrated by the pairwise comparisons for Experiment 2 became even *stronger* (rather than weaker) following the exclusion of the relevant vignettes. Thus, the re-analysis does not support the explanation that the observed effects were being driven by lexical priming rather than congruency effects between the vignettes and the faces. This conclusion is consistent with evidence from the literature that lexical priming effects are secondary to discourse level situation models and, furthermore, that lexical-level priming is influential only when the discourse-level model is unreliable due to incoherence or inadequate information (Camblin, Gordon, & Swaab, 2007; Ledoux, Camblin, Swaab, & Gordon, 2006). Indeed, in their investigation of whether readers utilise general knowledge of events to understand a discourse, Metusalem et al. (2012) re-analysed their data along similar lines to our re-analyses to explore whether lexical priming could provide an alternative explanation for their effects. After removing items from the analysis that were most likely to produce the strongest lexical priming effects, the same pattern of effects emerged. Metusalem et al. concluded that any influence of lexical priming was weak and that the discourse-level model exerted a much stronger influence. Altogether, such evidence from the literature alongside the evidence from our main analyses and the re-analyses presented here

strongly supports the predicted context congruency effects of the vignettes on facial emotion

processing as explained by Barrett and colleagues' extended language-as-context account.

References

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence

and lexical association during sentence processing: Evidence from ERPs and eye tracking.

Journal of Memory and Language, 56, 103-128. doi: 10.1016/j.jml.2006.07.005

De Deyne, S., Navarro, D. J., & Storms, G. (2013). Better explanations of lexical and

semantic cognition using networks derived from continued rather than single-word

associations. *Behavior Research Methods, 45,* 480-498. doi: 10.3758/s13428-012-0260-7

Ledoux, K., Camblin, C. C., Swaab, T. Y., & Gordon, P. C. (2006). Reading words in

discourse: The modulation of lexical priming effects by message-level context. *Behavioral*

and Cognitive Neuroscience Reviews, 5, 107-127. doi: 10.1177/1534582306289573

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012).

Generalized event knowledge activation during online sentence comprehension. *Journal of*

Memory and Language, 66, 545-567. doi: 10.1016.j.jml.2012.01.001

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida

word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/

Table 1: Cues for "Angry" and "Fearful" (Experiment 1) and vignettes excluded from re-analysis

"Angry" vignettes

Cue	Vignette	Cue	Vignette	Cue	Vignette
Anger		Annoyed		Birds	
Cross		Emotion		Enraged	
Face		Fight		Frown	
Frustrated		Frustration		Furious	
Fury		Hate		Hateful	
Irate		Irked		Mad	
Me		Mean		Mood	
Pissed		Pissed off		Rage	
Red		Red faced		Sad	
Scream		Shout		Temper	
Unhappy		Upset		Violent	
Wrath		Yell			

"Fear" vignettes

Cue	Vignette	Cue	Vignette	Cue	Vignette
Afraid		Alone		Anxiety	
Anxious		Child		Cower	
Danger	Henry (dangerous)	Dark	Mike Lindsey Liliana	Face	
Fear		Fire		Fright	
Frightened		Frightening		Ghost	
Horror		Mouse		Nervous	
Phobia		Powerless		Scared	
Scary		Small		Spider	
Spiders		Tearful		Terrified	
Terror		Timid		Upset	
Weak		Worried			

Table 2: Cues for "Happy" and "Sad" (Experiment 2) and vignettes excluded from re-analysis

"Happy" vignettes					
Cue	Vignette	Cue	Vignette	Cue	Vignette
Birthday	Ethan	Bright		Cheerful	
Content		Crappy		Day*	
Days*		Dog		Dwarf	
Ecstatic		Elated		Excited	
Face		Fun		Gay	
Glad		Glee		Gleeful	
Go lucky		Good	Danielle	Grin	
Happiness		Joy		Joyful	
Joyous		Love	Ethan	Lucky	
Pleasant		Pleased		Sad	
Satisfied		Smile		Smiles	
Smiling		Sun		Sunny	
Teeth		Very	Becky	Yellow	
Yes					

"Sad" vignettes					
Alone		Bad	Joanna (badly)	Blue	
Bored		Clown		Cry	
Crying	Adam	Day*		Death	
Depressed		Depression		Down	
Emotion		Emotional		Face	
Feeling		Frown		Gloom	
Gloomy		Glum		Grief	
Happy		Lonely		Mad	
Melancholy		Miserable		Morose	
Mournful		Panda		Rain	
Sack		Sorrow		Tear	
Tears		Unhappy		Upset	
Weak					

* There were vignettes that contained "day" but as this is a cue for both "happy" and "sad," these vignettes were included in the re-analysis.

Table 3: Parameter estimates following the linear mixed effects model reaction time re-analysis for

Experiment 1.

	b	SE	t
Intercept	2352.69	212.54	11.070
Vignette Emotion	-27.64	139.89	-0.198
Facial Emotion	-99.78	168.74	-0.591
Interaction	-1436.65	242.46	-5.925*

Table 4: Marginal means and SEs for Experiments 1 and 2 re-analyses

	RT mean	RT SE	Accuracy mean	Accuracy SE
Anger Anger	1930	236	0.985	0.009
Fear Anger	2676	259	0.860	0.035
Fear Fear	2057	216	0.981	0.009
Anger Fear	2748	275	0.927	0.023
Happy Happy	1255	90	0.992	0.006
Sad Happy	1665	104	0.977	0.014
Sad Sad	1348	95	0.998	0.002
Happy Sad	1683	96	0.990	0.007

Table 5: Parameter estimates following the generalized linear mixed effects model accuracy re-analysis for Experiment 1.

	b	SE	z
Intercept	3.11	0.28	11.020
Vignette Emotion	0.49	0.39	1.237
Facial Emotion	-0.23	0.36	-0.649
Interaction	3.74	0.72	5.159*

Table 6: Parameter estimates following the linear mixed effects model reaction time re-analysis for Experiment 2.

	b	SE	t
Intercept	1487.90	85.33	17.437
Vignette Emotion	-37.23	53.43	-0.697
Facial Emotion	-55.07	53.77	-1.024
Interaction	-746.04	91.57	-8.147*

Table 7: Parameter estimates following the generalized linear mixed effects model accuracy re-analysis for Experiment 2.

	b	SE	z
Intercept	4.81	0.59	8.196
Vignette Emotion	-0.26	0.72	-0.363
Facial Emotion	-1.13	0.64	-1.747
Interaction	2.58	1.28	2.011*