COGNITION AND EMOTION
2021, VOL. 35, NO. 2, 385–392
https://doi.org/10.1080/02699931.2020.1823322

BRIEF ARTICLE

Emotional gist: the rapid perception of facial expressions

Elizabeth Gregory [a], James W. Tanaka[b] and Xiaoyi Liu[b]

[a]Department of Psychiatry, University of British Columbia, Vancouver, Canada; [b]Department of Psychology, University of Victoria, Victoria, Canada

ABSTRACT
While it has been established that expression perception is rapid, it is unclear whether early appraisal mechanisms invoke holistic perception. In the current study, we defined gist perception as the appraisal of a stimulus within a single glance (<125 ms). We employed the expression composite task used previously by Tanaka and colleagues in a 2012 study, with several critical modifications: (i) we developed stimuli that eliminated contrast artifacts, (ii) we employed a masking technique to abolish low-level cues, and (iii) all the face stimuli were composite stimuli compared to mix of natural and composite stimuli previously used. Participants were shown a congruent (e.g. top: angry/ bottom: angry) or incongruent (e.g. top: angry/ bottom: happy) expression for 17, 50 or 250 ms and instructed to selectively attend to the cued expression depicted in the top (or bottom) half of the composite face and ignore the uncued portion. Compared to the isolated condition, a facilitation effect was found for congruent angry expressions, as well as an interference effect for incongruent happy and angry expressions at the shortest exposure duration of 17. Together these results provide evidence that the holistic gist perception of expression cannot be overridden by selective attention.

ARTICLE HISTORY
Received 11 February 2020
Revised 4 July 2020
Accepted 9 September 2020

KEYWORDS
Facial expression; holistic gist

The perception of expression in faces is integral to emotional communication (Frith, 2009). Despite the rapid and transitory nature of facial emotions, people are able to detect and recognise expressions in exposure durations as brief as 10 ms (Sweeny et al., 2013). To facilitate their speeded recognition, it has been speculated that facial expressions are perceived holistically, where the separate components of a facial emotion (e.g. furled brow, clinched teeth) are combined into a unified perception of emotion (e.g. anger) (Calder et al., 2000; Calder & Jansen, 2005; Tanaka et al., 2012). According to a holistic view of expression perception, it is difficult to attend to a local feature of a facial emotion (furled brow) while ignoring information (clinched teeth) in other face areas. While face expressions have key regions on which the viewer primarily focuses to "diagnose" their valence (e.g. smiling mouth for happy, furled brow for angry; Smith et al., 2005), there is evidence viewpoint-invariant gaze patterns in facial expression appraisal, demonstrating that holistic mechanisms are present at early perceptual stages (Guo & Shaw, 2015). Furthermore, neurophysiology studies demonstrate that expression encoding is associated with right-lateralised N`170 activity, typically associated with configural, i.e. holistic, processing of faces (Calvo & Beltrán, 2014).

The composite face paradigm has been developed to test the holistic recognition of facial identity (Young et al., 1987). In the original version, a composite face is created by combining the top and bottom halves of two different identities to form a new composite identity. Participants are shown the composite face and asked to identify the identity in either cued top or bottom half while ignoring the identity in the uncued half. The key finding was that participants were less accurate and slower identifying the cued half of the face when top and bottoms were aligned. The poor performance in the aligned condition was attributed to holistic interference caused by the to-be-ignored face half disrupting perception of the attended half.

CONTACT Elizabeth Gregory elizabeth.gregory@ubc.ca
This article has been republished with minor changes. These changes do not impact the academic content of the article.

Critically, when the top and bottom halves were misaligned or if the composite face was inverted, the holistic interference effect is greatly reduced or abolished (Young et al., 1987). The composite face task has been adapted to test the holistic processing of facial expressions. In the composite expression task, the top half of one facial emotion (*e.g.* angry) is fused with the bottom half of another emotion (*e.g.* happy) to form an incongruent expression. Participants are instructed to identify the expression in the cued half (or bottom half) and to ignore the expression in the uncued half. In studies employing this paradigm, the key finding was that the recognition of the cued portion was slower and less accurately recognised in an incongruent expression, indicating that the recognition of facial expressions involves the integration into a holistic percept (Calder et al., 2000; Tanaka et al., 2012; Tobin et al., 2016). Importantly, the composite effect is disrupted when the top and bottom face halves are misaligned, or when the expression is inverted (Calder et al., 2000; Calder & Jansen, 2005; Tanaka et al., 2012). These findings provide compelling evidence that the recognition of facial expressions, like the recognition of facial identities, involves holistic processing cannot be easily overridden by selective attention.

How fast is holistic expression recognition? Tanaka et al. (2012) demonstrated that exposure durations of 20, 60, 100 and 120 ms was enough time to produce congruency effects in the composite paradigm where congruent expressions (e.g. attend top: angry/ ignore bottom: angry) were more accurately identified than the incongruent expressions (e.g. attend top: angry/ ignore bottom: happy). Given that it takes humans at least 120 ms to execute a visual saccade (Kirchner & Thorpe, 2006 Schiller & Kendall, 2004), it can be assumed that in the Tanaka et al. study, whole face expression processing occurred within the time of a single fixation and has been referred to as *holistic gist* (Liu & Tanaka, 2019). Whereas previous research found the evidence of holistic gist for the recognition of facial identity (Liu & Tanaka, 2019), Tanaka et al. (2012) found evidence that holistic gist was implicated in the recognition of facial emotion.

Although Tanaka et al. (2012) reported evidence for holistic gist recognition of facial expressions, methodological limitations of their study bring their findings into question. First, the researchers employed a general Gaussian mask in their procedure to disrupt visual system processing following expression presentation in order to determine the effect of exposure duration on expression appraisal. However, more recent evidence shows that Gaussian masking does not disrupt the visual system; therefore, the effect of exposure duration is confounded by potential ongoing visual processing following stimulus presentation (Étienne et al., 2017). Hence, it is important to replicate the findings of Tanaka et al. (2012) using a masking procedure that will more effectively disrupt the low-level processes of the visual system.

Second, Tanaka et al. (2012) employed open-mouth, teeth-baring expressions in their study. The visibility of teeth creates an attention bias in happy or angry target faces (Hortsmann et al., 2012), which may have significantly impacted trials where open-mouth expressions were featured. In contrast, using closed mouth expressions eliminates low-level contrast information that is present between the mouth and teeth, and reduces the likelihood of non-holistic processing being implemented.

Third, stimuli used by Tanaka et al. (2012) were a mix of composite and natural expressions. That is, incongruent and neutral trials were created by combining the top and bottom halves of different facial expressions, whereas for congruent trials natural face expressions were used. This could introduce an important confound, specifically that perception of the unnatural, or composite faces, may inherently be more challenging regardless of whether the expression in the top or bottom half match. Thus a replication study employing closed-mouth, truly composite stimuli, is necessary to ensure that the findings reported by Tanaka et al. (2012) are not influenced by these potential confounding factors.

Finally, gist perception is defined by face processing that occurs within a single fixation. With this in mind, the cuing procedure should ensure that participant gaze is in the right location prior to stimulus presentation. Tanaka et al. (2012) employed a central fixation cross, which would require participants to saccade at least once in order to fix their gaze to the correct location on the screen.

Based on the previous findings (Tanaka et al., 2012), we made two predictions. First, applying a more stringent procedure, we expected to replicate the main results of Tanaka et al. (2012) in which we would find evidence of a general effect of holistic interference where information of the unattended half of the expression would impair recognition of the attended half of the expression. Second, we expected to find the evidence of holistic gist, wherein participants would show the effects of interference and congruency even at the shortest exposure duration of 17 ms.

Method

Participants

Sixty-one undergraduate students at the University of Victoria (10 men; $M_{age} = 21.30$ years, $SD = 2.65$) with normal or corrected-to-normal vision participated in this experiment for course credit. 10 participants were left-handed.

Materials

The stimuli set were four constructed female identities which were modified images from the NimStim Face Set database (Tottenham et al., 2009). In contrast to the Tanaka et al. (2012) study, which simply combined different expressions from the same identity, the stimuli used in the current study were truly composite in that the features in a given face were each taken from a separate identity. The same template face was used for all four composite faces to retain uniformity in hair and face shape across identities. To ensure the stimuli were of high quality, a pilot group of 15 participants were asked to judge the facial expressions in the happy and angry condition for each identity; accuracy was at or above 85% in each case. Participants were also asked to rate the intensity of the expressions on a scale of 1–5 (1 for no expression, 5 for highest intensity). On average, happy expressions were given a rating of 2.96/5, while angry expressions were given a rating of 3.42/5.

Each identity had seven face conditions that combined happy and angry expressions in the top or bottom half of a face with another face half: congruent (e.g. happy top + happy bottom), incongruent (e.g. angry top + happy bottom), neutral (e.g. neutral top + happy bottom) and isolated (e.g. no top + happy bottom). As expressions tend to vary in the weight given to each feature of the face (Smith et al., 2005), expression combinations with diagnostic features present (DFP), and diagnostic features absent (DFA), were possible in the stimuli set: Happy is a "bottom" emotion, and thus trials including the smiling mouth were classified as DFP. Similarly, angry is a "top" expression, therefore trials including angry eyes were classified as DFP (Calvo & Beltrán, 2014). Examples of DFP stimuli are shown in Figure 1(A). In contrast, stimuli with neither diagnostic feature present (i.e. no angry eyes and/or smiling mouth) were classified as DFA.

Design

Replicating the experimental design used by Tanaka et al. (2012), a within-subjects design was used with the factors of condition (congruent, incongruent, isolated, neutral), expression (happy, angry) and exposure duration (17, 50, 250 ms). Stimuli were separated into two blocks, where participants were asked to attend solely to the bottom half of the face in one block and the top half of the face in the other. The block, and thus target region, order was randomised across participants. Each block contained 96 trials, for a total of 192 unique trials. The blocks were repeated twice, giving 384 trials per participant. Because we were interested in how the selective attention to a given expression is influenced by the unattended information in the rest of the face, only DFP expression trials were analysed. Further, neutral expressions can convey distinct information and are therefore not an effective control condition (Suess et al., 2015). DFA and neutral trial types were seen as catch trials and not included in the statistical analyses. Thus 144 trials in total were analysed for each participant (with 72 of those trials being unique).

Procedure

The participant sat in a darkened room at approximately 60 cm from a 15-inch LG Flatron F700P monitor on which the faces were presented by a PC computer and MATLAB software. Participants were asked to label the expression shown in the target top or bottom half of the face. Trials were blocked by top and bottom cued faces. At the beginning of a block, participants were told which half of the face to attend to for the remainder of the block.

In a given trial, participants were shown a fixation arrow for 500 ms which directed participant gaze to the relevant half of the screen and ensure judgments were made within a single fixation. The face stimulus was presented for 17, 100 or 250 ms. A diffeomorphic mask was then presented for 500 ms (Stojanoski & Cusack, 2014). The diffeomorphic mask is superior compared to the traditional Gaussian noise mask (used in Tanaka et al., 2012) in disrupting the perception of high-level visual details, while preserving low-level properties.

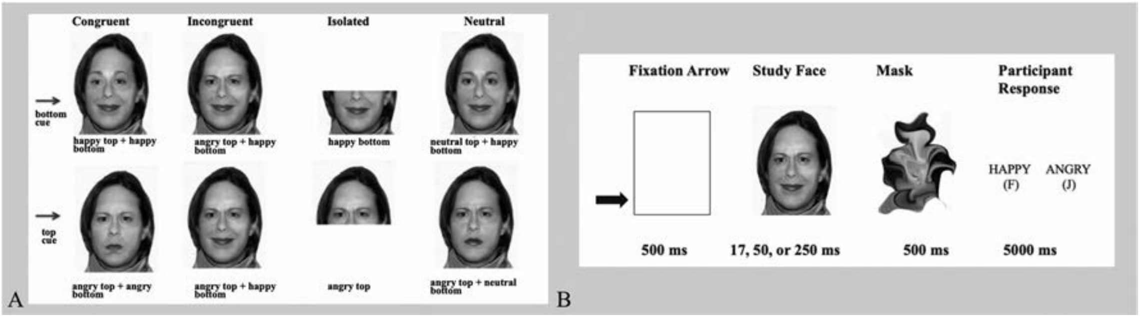

Figure 1. (A) Example of DFP stimuli presented in the task. Each type of stimuli, i.e. congruent, incongruent and isolated, were shown with happy as the bottom cue, or angry as the top cue. (B) Example of an incongruent trial. Participants were first shown a fixation arrow for 500 ms, followed by the study face for a variable SOA, which was then masked using a diffeomorphic mask for 500 ms. Finally, participants were asked to indicate which expression they saw in the targeted region of the face, using either the "F" key to indicate happy or the "J" key to indicate angry.

Following the masking, a response screen was shown for 5000 ms, or until a response was given. Participants responded on a standard QWERTY keyboard by pressing either the "J" key for angry, or the "f" key for happy. The trial procedure is demonstrated in Figure 1(B).

Statistical analyses

All analyses were performed using R, version 3.5.1. A within-subjects ANOVA (package ez, version 4.4.0; Lawrence, 2016) was performed for the dependent variables of accuracy and response time. Independent variables loaded into the ANOVA were expression (happy and angry), exposure duration (17, 50, 250 ms) and condition (congruent, incongruent, isolated) as factors. Sum of squares was set to "3". Post-hoc tests, using Holm–Bonferroni correction (package emmeans, version 1.4.0), were performed for significant interactions.

Results

Accuracy

The main effect of expression ($F(1, 60) = 37.82$, $p < .001$, $\eta^2 = 0.39$) was significant, demonstrating that happy expressions ($M = 0.90$, $SD = 0.29$) were more accurately recognised than angry expressions ($M = 0.78$, $SD = 0.41$), $p < .001$. The effect of condition was significant ($F(2, 120) = 160.77$, $p < .001$, $\eta^2 = 0.73$), with congruent trials being most accurate ($M = 0.94$, $SD = 0.24$), followed by isolated trials ($M = 0.85$, $SD = 0.36$), with incongruent trials showing the lowest accuracy ($M = 0.72$, $SD = 0.45$). Even at the gist exposure durations (17 and 50 ms), holistic facilitation (advantage in congruent compared to isolated trials, $p < .001$ for both exposure durations) and holistic interference (disadvantage in incongruent compared to isolated trials, $p < .001$ for both exposure durations) were demonstrated.

Across all conditions, longer exposure durations benefited performance ($F(1, 60) = 52.78$, $p < .001$, $\eta^2 = 0.47$), with participants showing the lowest accuracy for 17 ms trials ($M = 0.81$, $SD = 0.38$), higher accuracy on the 50 ms trials ($M = 0.83$, $SD = 0.36$, $p = .05$) and 250 ms trials ($M = 0.87$, $SD = 0.33$) being most accurate ($p < .001$ in both cases).

There was a reliable two-way interaction between expression and condition ($F(120, 2) = 17.19$, $p < .001$, $\eta^2 = 0.22$). For angry expression trials, congruent condition trials ($M = 0.93$, $SD = 0.25$) were significantly better, and incongruent condition trials significantly worse ($M = 0.64$, $SD = 0.48$), than the isolated trials ($M = 0.76$, $SD = 0.42$), $p < .001$. Happy expression trials, however, showed no difference between the congruent condition trials ($M = 0.95$, $SD = 0.22$) and the isolated ($M = 0.94$, $SD = 0.024$) condition trials, $p > .999$, although similar to angry, incongruent trials ($M = 0.80$, $SD = 0.40$) were worse than isolated trials, $p < .001$.

The two-way interaction between expression and exposure duration was significant ($F(60,1) = 4.93$, $p = .004$, $\eta^2 = 0.08$). Multiple comparisons for the interaction revealed that at 17, 50 and 250 ms exposure durations, angry expression trials showed significantly worse accuracy compared to happy expression trials ($p < .001$ for all cases).

The two-way interactions of expression and exposure duration, and expression and condition, were qualified by the three-way interaction between exposure duration, expression, and condition ($F(120, 2) = 13.79$, $p < .001$, $\eta^2 = 0.19$), which was primarily driven by the difference in accuracy across exposure duration for incongruent angry expressions. At both 17 and 50 ms conditions, angry incongruent trials (17 ms: $M = 0.57$, $SD = 0.50$; 50 ms: $M = 0.64$, $SD = 0.48$) were significantly worse than happy incongruent trials (17 ms: $M = 0.80$, $SD = 0.40$; 50 ms: $M = 0.80$, $SD = 0.40$), $p < .001$. However, at 250 ms, incongruent happy ($M = 0.80$, $SD = 0.40$) and incongruent angry ($M = 0.71$, $SD = 0.45$) expression trial performance was no different, $p = .39$.

Response time

The main effect of expression ($F(60, 1) = 8.15$, $p = .006$, $\eta^2 = 0.12$) was significant, whereas the main effects of condition ($F(120,2) = 0.61$, $p = .55$, $\eta^2 = 0.01$) and exposure duration ($F(120, 2) = 0.003$, $p = .95$, $\eta^2 = 0.001$) were not reliable. Response time for happy expression trials ($M = 293.39$, $SD = 206.60$) was significantly faster than for angry expression trials ($M = 338.50$, $SD = 224.57$), $p < .001$. A two-way interaction was found between expression and condition ($F(120, 2) = 3.55$, $p = .03$, $\eta^2 = 0.05$), with differences in response time found between expressions for isolated trials: angry expression trial ($M = 346.05$, $SD = 243.33$) response times were significantly longer compared to happy expression trials ($M = 271.50$, $SD = 132.38$), $p = .02$.

Discussion

Gist mechanisms of holistic expression perception

The current study investigated the implication of holistic perception in expression recognition at gist exposure durations. Applying the Composite Expression task, participants were instructed to judge either the cued top or bottom half, while ignoring the uncued half, of facial expressions that were presented for 17, 50 and 250 ms. The composite effect was clearly demonstrated for angry expression, where information in the unattended bottom half of the face facilitated recognition in the congruent, and interfered with recognition in the incongruent, condition relative to the isolated condition. Although angry expressions are diagnosed primarily using the top half information of a face (Smith et al., 2005), these results suggest that information from the bottom face contributes to the identification of angry expressions, and confirm the findings reported by Tanaka et al. (2012).

In contrast, happy expression identification was not enhanced by the presence of congruent happy eyes relative to the isolated condition, suggesting that the smiling mouth of a happy expression is a highly salient signal for happiness (Calvo & Beltrán, 2014). However, when the cued smiling mouth was accompanied with uncued angry eyes, holistic interference disrupted recognition accuracy, replicating the findings reported by Tanaka et al. (2012). Thus when there is conflicting information in the eyes and mouth regions of a face, an interference effect occurs that impairs the recognition of angry and happy expression. For the angry expression, when there is congruent information in the eyes and mouth, recognition of anger is facilitated relative to the isolated condition whereas congruent eyes and mouth information did not facilitate the recognition of happy. In sum, the effect of condition reported by Tanaka et al. was replicated in the current paper; collectively, our results suggest that holistic expression recognition is dependent on both emotional expression and congruency of eye and mouth information.

The current study, however, reports the effects of exposure duration contrasting with Tanaka et al. (2012). As Figure 2 shows, holistic processing occurred within the gist range of processing (Kirchner & Thorpe, 2006; Schiller & Kendall, 2004), where the exposure duration of 17 ms permitted only a single eye fixation but was sufficiently long to elicit holistic interference for happy and angry expressions. For the angry expressions, accuracy improved with increased exposure duration indicating that a longer presentation allowed participants to filter out the distracting smiling mouth and focus their attention to the target eyes. However, exposure duration had little effect on the perception of incongruent happy expressions. Even at the longest exposure duration of 250 ms, the conflicting information in the incongruent condition drove performance below the baseline isolated condition, indicating that the interfering angry eyes were so compelling that participants were not able to filter out the conflicting signal. It was striking that when participants viewed the same mouth in isolation, accuracy improved and approached near ceiling levels of performance. We also found holistic facilitation within the gist time window of 120 ms for

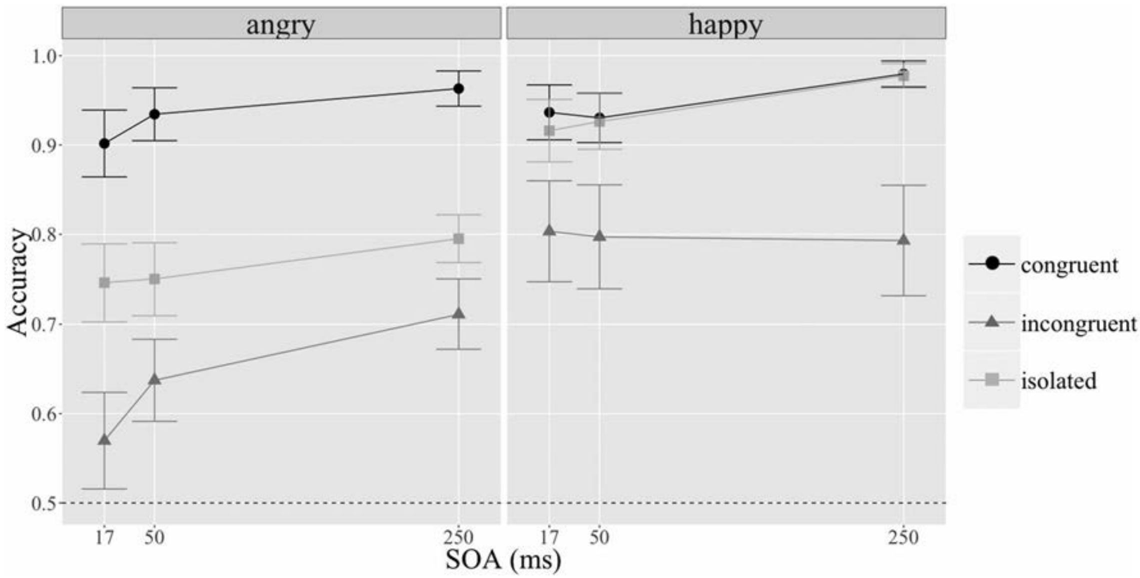

Figure 2. Mean accuracies across levels of expression (happy versus angry), exposure duration (17, 50 or 250 ms) and condition (congruent, incongruent and isolated). Error bars represent standard error. Accuracy is indicated as the proportion of correct trials.

congruent angry expressions: At the 17 and 50 ms exposure durations, recognition of the cued angry eyes was facilitated by uncued angry mouth.

Conversely, Tanaka et al. (2012) found only holistic interference, and no holistic facilitation, effect for angry expressions, and no holistic interference or facilitation effect for happy expressions, at the shortest exposure duration (20 ms), although all holistic processing effects (except holistic facilitation in happy expressions) were reported at longer exposure durations. The findings reported in the current study therefore suggest that holistic processing occurs instantaneously, rather than being implicated when participants are given longer to process a face. We also report that congruent expression recognition was not affected by exposure duration; performance was just as good at the shortest exposure duration of 17 ms compared to the longest exposure duration of 250 ms. Tanaka et al. (2012), however, report that congruent expression recognition improved with longer exposure duration.

These differences in SOA effects are surprising, especially given that the current study applied a diffeomorphic masking procedure which effectively disrupts primary visual cortex processing (Stojanoski & Cusack, 2014), allowing for more confidence in effects exposure duration compared to traditional masking procedures (e.g. Gaussian masking). However, Tanaka et al. (2012) simply used a fixation cross in the centre of the screen, while we presented an arrow pointing to the location of the relevant feature, prior to stimulus presentation. Without the arrow cuing participants to attend to a particular location on the screen, upon stimulus presentation participants would need to direct their gaze prior to stimulus appraisal. 20 ms, which was the shortest SOA used by Tanaka et al. (2012), is not enough time for more than 1 eye fixation (Hsiao & Cottrell, 2008). Therefore, the lower level of accuracy at shorter SOA previously reported may be artificial because participants were not able to fixate on the relevant face half *and then* appraise the stimulus within the given exposure duration. The results of the current study are thus promising evidence of the gist nature of expression recognition. That is, a single eye fixation is "good enough" to accurately appraise an expression.

Limitations

There are limitations in the current study that must be addressed. The composite expression task used in this study is not emblematic of expression perception in everyday life. Differentiating between happy and angry expressions is not a difficult task, largely because their diagnostic features show little overlap

(Calvo & Beltrán, 2014). Furthermore, we only analysed accuracy on the "key" diagnostic features for each expression, being the mouth for happy, and the eyes for angry. For target expressions that involve more whole face information (e.g. sadness) rather than an isolated diagnostic feature, it is predicted that holistic processing would be more strongly recruited. Similarly in their more nuanced versions (a slight grin or an irritated look), it is likely that the recognition of both happy and angry expressions would benefit from holistic processing than the more extreme versions of the expressions employed in this study.

Second, the participants in the study were sampled from a university student subject pool. Seeing as facial expression processing varies across the lifespan, age may have an effect on emotional gist (Meinhardt-Injac & Hildebrandt, 2016). In addition, the sample was mostly comprised of female participants, whereas holistic perceptual mechanisms may be influenced by sex (Rennels & Cummings, 2013). As such, these results can be considered to reflect emotional gist processing in healthy, young female adults.

Finally, this study is only able to address holistic processing in a forced-choice condition wherein a participant differentiates between two emotions. The other universal emotions, namely fear, surprise, sadness and disgust, were not included in this paradigm. Given that everyday differentiation of emotions involves (i) more nuanced expressions of emotion and (ii) differentiating between all emotions, future studies using a wider range of expressions are warranted.

Conclusion

The results of the current study impress upon the gist mechanisms of expression appraisal, where we find a high level of accuracy even at the 17 ms exposure duration, demonstrating that we only need a single glance to appraise a facial expression. Our results demonstrate that within the gist time window of 120 ms, holistic interference was found for the recognition of angry and happy expressions and holistic facilitation was found for the recognition of angry expressions. In summary, the present results provide promising evidence of the "holistic gist" nature of expression recognition.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by a grant to James W. Tanaka from the Natural Sciences and Engineering Research Council of Canada (NSERC).

ORCID

Elizabeth Gregory http://orcid.org/0000-0003-1301-6565

References

Calder, A. J., & Jansen, J. (2005). Configural coding of facial expressions: The impact of inversion and photographic negative. *Visual Cognition*, *12*(3), 495–518. https://doi.org/10.1080/13506280444000418

Calder, A. J., Young, A. W., Keane, J., & Dean, M. (2000). Configural information in facial expression perception. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 527–551. https://doi.org/10.1037/0096-1523.26.2.527

Calvo, M. G., & Beltrán, D. (2014). Brain lateralization of holistic versus analytic processing of emotional facial expressions. *NeuroImage*, *92*, 237–247. https://doi.org/10.1016/j.neuroimage.2014.01.048

Étienne, C. J., Arleo, A., & Allard, R. (2017). Maximizing noise energy for noise-masking studies. *Behavior Research Methods*, *49*(4), 1278–1290. https://doi.org/10.3758/s13428-016-0786-1

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society of London*, *364* (1535), 3453–3458. https://doi.org/10.1098/rstb.2009.0142

Guo, K., & Shaw, H. (2015). Face in profile view reduces perceived facial expression intensity: An eye-tracking study. *Acta Psychologica*, *155*, 19–28. https://doi.org/10.1016/j.actpsy.2014.12.001

Hortsmann, G., Lipp, O. V., & Becker, S. I. (2012). Of toothy grins and angry snarls – Open mouth displays contribute to efficiency gains in search for emotional faces. *Journal of Vision*, *12*, 5. https://doi.org/10.1167/12.5.7

Hsiao, J. H., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, *19*(10), 998–1006. https://doi.org/10.1111/j.1467-9280.2008.02191.x

Kirchner, H., & Thorpe, S. J. (2006). Ultra-rapid object detection with saccadic eye movements: Visual processing speed revisited. *Vision Research*, *46*(11), 1762–1776. https://doi.org/10.1016/j.visres.2005.10.002

Lawrence, M.A. (2016). ez: Easy analysis and visualization of factorial experiments. R package version 4.4-0. https://CRAN.R-project.org/package=ez

Liu, X., & Tanaka, J. W. (2019). Holistic perception of faces in 17 milliseconds: Evidence from three measures. *Journal of Vision*, *19*(10), 92. https://doi.org/10.1167/19.10.92

Meinhardt-Injac, B., & Hildebrandt, A. (2016). Editorial: Face perception across the life-span. Frontiers in Psychology, 7, Article 1338.

Rennels, J. L., & Cummings, A. J. (2013). Sex differences in facial Scanning: Similarities and dissimilarities between infants and adults. *International Journal of Behavioral Development*, *37*(2), 111–117. https://doi.org/10.1177/0165025412472411

Schiller, P. H., & Kendall, J. (2004). Temporal factors in target selection with saccadic eye movements. *Experimental Brain Research*, *154* (2), 154–159. https://doi.org/10.1007/s00221-003-1653-8

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological Science*, *16*(3), 184–189. https://doi.org/10.1111/j.0956-7976.2005.00801.x

Stojanoski, B., & Cusack, R. (2014). Time to wave good-bye to phase scrambling: Creating controlled scrambled images using diffeomorphic transformations. *Journal of Vision*, *14* (12), 12. https://doi.org/10.1167/14.12.6

Suess, F., Rabovsky, M., & Rahman, R. A. (2015). Perceiving emotions in neutral faces: Expression processing is biased by affective person knowledge. *Social Cogntive and Affective Neuroscience*, *10*(4), 531–536. https://doi.org/10.1093/scan/nsu088

Sweeny, T. D., Suzuki, S., Grabowecky, M., & Paller, K. A. (2013). Detecting and categorizing fleeting emotions in faces. *Emotion*, *13*(1), 76–91. https://doi.org/10.1037/a0029193

Tanaka, J. W., Kaiser, M. D., Butler, S., & Le Grand, R. (2012). Mixed emotions: Holistic and analytic perception of facial expressions. *Cognition and Emotion*, *26*(6), 961–977. https://doi.org/10.1080/02699931.2011.630933

Tobin, A., Favelle, S., & Palermo, R. (2016). Dynamic facial expressions are processed holistically, but not more holistically than static facial expressions. *Cognition and Emotion*, *30*(6), 1208–1221. https://doi.org/10.1080/02699931.2015.1049936

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., Marcus, D. J., Westerlund, A., Casey, B. J., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242–249. https://doi.org/10.1016/j.psychres.2008.05.006

Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*(6), 747–759. https://doi.org/10.1068/p160747