Agency and Facial Emotion Judgment in Context

Kenichi Ito
Institute on Asian Consumer Insight

Takahiko Masuda and Liman Man Wai Li
University of Alberta

Author Note
Correspondence concerning this article should be addressed to Kenichi Ito, Institute on Asian Consumer Insight, S3-01c-113, Nanyang Technological University, 50m Nanyang Ave., Singapore, 639798. E-mail: kito@ntu.edu.sg. The project was supported by the standard grant of the Social Science and Humanities Research Council of Canada to the second author (G124130642). We thank research assistants at the Culture and Cognition Lab at the University of Alberta for data collection.
Abstract
Past research showed that East Asians’ belief in holism was expressed as their tendencies to include background facial emotions into the evaluation of target faces more than North Americans. However, the pattern can be interpreted as North Americans’ tendency to downplay background facial emotions due to their conceptualization of facial emotion as volitional expression of internal states. Examining this alternative explanation, we investigated whether different types of contextual information produce varying degrees of effect on face judgment across cultures. In three studies, European Canadians and East Asians rated the intensity of target facial emotions surrounded with either affectively salient landscape sceneries or background facial emotions. The results showed that, although the salient valence of landscapes influenced the judgment of the two cultural groups, European Canadians downplayed the background facial emotions unlike East Asians. The role of agency as differently conceptualized across cultures, and multilayered systems of cultural meanings are discussed. (150 words)
Agency and Facial Emotion Judgment in Context

Contextual information is informative for social judgment. The atmosphere of a fancy restaurant creates a certain mood for romantic couples. Advertisers use ambience and peripheral cues to influence consumers’ decision-making processes. A similar effect has even been observed when people evaluate facial expressions. According to James Russell, a leading researcher on the effect of context on facial emotion judgment, “judgment of emotion in a facial expression is not a simple straightforward registration of the meaning of that face. The face is judged not in an absolute manner but relative to the context of judgment” (1991, p. 150). A plethora of research demonstrates that the salient context strongly influences people’s judgment of a target person’s facial expressions. For example, face perception has been found to be influenced by language (Barrett, Lindquist, & Gendron, 2007), body posture (Meeren, van Heijnsbergen, & de Gelder, 2005), and written descriptions of situations where emotion is occurring (Carroll & Russell, 1996; Matsumoto, Hwang, & Yamada, 2012). Furthermore, researchers have reported that the mere presentation of an anchor facial emotion (Russell, 1991; Russell & Fehr, 1987) or affect-provoking landscape scenery (Hietanen, Klemettilä, Kettunen, & Korpela, 2007) before the judgment of the target face influences categorizations of target facial emotions. The current studies add to this line of research by investigating the effect of contextual information on facial emotion judgment from a cultural psychological perspective. In particular, we present evidence that the cultural meanings prevalent among members of North American cultures attenuate the influence of contextual information on North Americans’ evaluation of facial emotions.

Cross-cultural studies of the context effect on face judgment have shown people’s culturally learned biases (Matsumoto & Hwang, 2010). Masuda and colleagues (Masuda et al.,
2008; Masuda, Wang, Ishii, & Ito, 2012) asked American and Japanese participants to evaluate the intensity of a facial emotion presented against a background of four other facial emotions. Target facial emotions were either congruent or incongruent with the background facial emotions (e.g., happy center with happy others vs. happy center with sad others); a larger discrepancy between the congruent and incongruent condition ratings indicated a larger effect of context. The results showed that American participants were less influenced by background facial emotions than were Japanese participants.

Where does the cultural difference in judgment come from? Many researchers have interpreted the context effect based on their argument of differences in culturally shared meaning systems: holistic thought vs. analytic thought. According to this theoretical framework, East Asians believe that elements in the universe are interconnected and exist in a state of flux (Nisbett, 2003; Nisbett & Masuda, 2003; Nisbett, Peng, Choi, & Norenzayan, 2001). Past research indeed suggests that East Asians are more context sensitive than North Americans (Masuda, Gonzalez, Kwan, & Nisbett, 2008; Masuda & Nisbett, 2001, 2006; Miyamoto, Nisbett, & Masuda, 2006; Wang, Masuda, Ito, & Rashid, 2012). According to this logic, East Asians, who tend to endorse holistic thought, would be more willing to connect background facial emotions with the target facial emotions. In contrast, North Americans, who tend to endorse analytic thought, would be more willing to dissociate the target facial emotions from background facial emotions. As a result, the discrepancy of the ratings between the congruent and the incongruent condition was smaller for North Americans than for East Asian participants (Masuda et al., 2008; Masuda et al., 2012).

Several findings, however, offer an alternative explanation, suggesting that North Americans conceptualize facial emotions as a volitional or intentional expression of internal
states. For example, North Americans tend to attribute social behavior to dispositions or internal factors (Choi, Nisbett, & Norenzayan, 1999). This tendency is observed when North Americans discuss causes for a variety of social behaviors, ranging from winning an Olympic gold medal (Markus, Uchida, Omorige, Townsend, & Kitayama, 2006) to the way a group of fish swims in a lake (Morris & Peng, 1994). Similarly, North Americans tend to regard an individual as a distinct agent whose emotions are disengaged from those of other people (Markus & Kitayama, 1994; Mesquita & Markus, 2004; Mesquita & Leu, 2007). In addition, they tend to create and live in everyday situations that make them feel efficacy and foster active control (Morling, Kitayama, & Miyamoto, 2002).

Given these North American ways of being a person, it is possible to assume that North Americans would expect people to have strong control over the production of facial emotions which represent their internal states of mind, and to express unambiguous information about the self without being affected by other people. Therefore, when North Americans evaluate the intensity of the target’s facial expressions, they would downplay the salient facial emotions of background people, regarding them as distinct from those of the target person. That is, people may share a meaning system in which the target facial emotions and surrounding others’ facial emotions are independent of or impervious to each other. Also, everyone is entitled to produce their own facial emotions regardless of target or background, and their internal states, which are expressed through their facial emotions, should be respected as a part of who they are.

In fact, evidence suggests that North Americans are sensitive to context when there is no agency information (Hietanen et al., 2007; Ito, Masuda, & Hioki, 2012). For example, in Ito et al.’s recent research (2012), European Canadian participants were asked to categorize, as quickly and accurately as possible, facial emotions presented against a background of landscape scenery.
Similar to Masuda and colleagues’ research (Masuda et al., 2008; Masuda et al., 2012), the emotional undertones of the landscape scenery were either congruent or incongruent with the target facial emotions (e.g., the target’s happy facial expression against a beautiful beach vs. the same facial expression shown with a dirty toilet). Following the argument that North Americans, with their low levels of holism, dissociate the target from its context, one might expect North American participants to be unaffected by changes in backgrounds. On the contrary, Ito et al. found the participants’ response latency to be significantly influenced by the congruency of the target facial emotions with the emotional undertones of landscape sceneries. Such results give credence to the meaning-based interpretation of the context effect on North Americans’ face evaluation, and pose a challenge to the holism interpretation of the issue of face recognition.

Many researchers (e.g., Nisbett et al., 2001) maintain that the holism explanation and the agency explanation are complementary, and that amalgamating the multiple effects of meaning systems provides a better way of comprehensively understanding people’s behavior in a given cultural context. However, it is also advisable to decompose, degrade, and particularize the observed effect to articulate the effect of each cultural meaning system. This endeavor is hindered by the failure of past studies to provide evidence of different context effects, by using different types of contextual information. Because most researchers have used context with either agency or non-agency, it has been difficult to articulate whether, in addition to the holism explanation, the agency explanation is necessary to explain the judgment patterns of targets’ facial expression. Current research contributes to the literature by showing the boundary condition between the two explanations.

Overview and Hypotheses of the Studies
To advance our understanding of the context effect on face evaluation and cultural meaning systems, we examined whether two types of salient context information (the non-agency information of landscape sceneries vs. the agency information of others’ facial emotions) differently influence North Americans’ intensity ratings of target facial emotions compared to those of East Asians. We compared participants’ intensity ratings on facial emotions rather than categorical judgment or response speed to capture subtle yet significant influence of contexts in our daily life. We posit that the agency in facial expressions fostered North Americans’ tendency to downplay the contextual effect of background facial emotions.

In three studies, European Canadians and East Asians were asked to rate the intensity of happy or sad facial emotions. The affect-provoking landscape sceneries used in Study 1 contained no agent; yet, as previous research has demonstrated (e.g., Ito et al., 2012), the affective nature of the sceneries was salient enough to influence both North Americans’ and East Asians’ categorical judgment. We therefore hypothesized that, in this condition, salient landscape information would influence both European Canadians’ and East Asians’ intensity ratings of the target facial emotions. In Study 2, the target facial emotions were surrounded by facial emotions of others. Similar to Masuda and colleagues’ research (Masuda et al., 2008; Masuda et al., 2012), the context contained agents different from the target. We therefore hypothesized that, in this condition only, North Americans would downplay salient agent information in the context, because they believe that the target’s facial emotions reflect the target’s private state of mind. Thus the evaluation of the target’s facial emotions would remain the same, regardless of the congruency between facial emotions of the target and background others, because one’s facial emotions are considered immune to the facial emotions of others. In Study 3, we attempted to replicate the results of Studies 1 and 2 by combining the procedures of
these studies while targeting only North Americans. Using a within-subject design, we examined whether North Americans’ evaluation of the target facial emotions was influenced more when the background was non-agentic than when it was agentic. In addition, we included our original questionnaires as well as pre-existing cultural scales (e.g., the holism scale, the interdependent vs. independent scale) to examine whether North Americans’ beliefs in agency of facial emotions were in fact associated with the differences in their intensity ratings for the two types of context.

**Study 1**

**Method**

**Participants.** Forty-six European Canadian students (27 females and 19 males) and 41 East Asian international students (24 females and 17 males; 20 Chinese, 11 Japanese, 8 Koreans, and 2 others) at the University of Alberta participated in the experiment for part of a course credit. One European Canadian student and eight East Asian international students were graduate students in departments other than psychology.

**Stimuli.** We selected 16 target facial portraits that were used in Ito et al.’s (2012) study. The target portraits showed either happy or sad facial emotion based on the criteria of Ekman and Friesen (1975). The target portraits varied equally in terms of models’ gender (male vs. female) and ethnicity (European American vs. Japanese). European Canadian and Japanese participants in a pre-test reported the equivalent intensity of happiness on four happy European American faces and four happy Asian faces, $ts < 1$. Similarly for sad faces, the two cultural groups reported the equivalent intensity of sadness, $ts < 1.25$ $p = ns$. We selected four positive and four negative images of landscape scenery from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2005). Both European Canadian and Japanese participants in a pre-test reported that the mean rating of the feelings elicited by four positive landscapes was
significantly more positive than the mean rating of feelings elicited by four negative landscapes $t_{CND}(27) = 22.07, p < .001$, and $t_{JPN}(19) = 10.26, p < .001$, respectively.\(^1\) We combined landscape images or a blank white background (control group) with each target facial portrait to create 48 experimental stimuli (see Figure 1).

**Procedure.** Upon arrival, participants were asked to read the instructions, which stated that the task was to evaluate the intensity of positive and negative emotions expressed by the target person. Participants were instructed to rate positive intensity ranging from 0 (not at all) to 9 (extremely), followed by negative intensity ranging from 0 (not at all) to 9 (extremely), for each stimulus. Stimuli remained on the monitor until participants finish the two intensity ratings. The instructions were presented in the participants’ native language. The first author translated the English version of the instructions into Japanese, and a bilingual research assistant back-translated the instructions. Any discrepancies in the back translation were resolved through discussions. Experimental stimuli were presented semirandomly using E-prime version 1.2 (Schneider, Eschman, & Zuccolotto, 2002) on a 17-inch monitor at a resolution of 870 × 459 pixels. We first randomly presented target faces combined with affectively salient background landscapes, and subsequently presented the target faces with white backgrounds (i.e., control condition). Participants sat in front of a computer at a viewing distance of approximately 60 cm. They were given six practice trials and subsequently completed 48 experimental trials. Participants’ intensity ratings for each stimulus were recorded. After answering a demographic information questionnaire at the end of the experiment, participants were fully debriefed.

**Results and Discussion**

We conducted a series of 2 (Culture: European Canadian vs. Japanese) × 3 (Congruency: congruent vs. incongruent vs. control) ANOVA, with participants’ intensity ratings being the
dependent variable. Culture was the between-subjects variable, and Congruency was the within-subject variable.\(^2\)

For the negative intensity ratings on sad target face, we found a main effect of Congruency, \(F(2, 170) = 11.35, p < .001, \eta_p^2 = .12\); but neither main effect of Culture nor Culture \(\times\) Congruency interaction was significant, \(F_s < 1\). Supporting our hypothesis, further analysis within each cultural group showed that both European Canadian and East Asian participants reported that a sad face surrounded with negative landscapes appeared more negative compared to the same faces surrounded with positive landscapes, \(F(1, 170) = 45.12, p < .001\) and \(F(1, 170) = 18.29, p < .001\), respectively (see Figure 2).\(^3\) For positive intensity ratings on happy target, we used the Geisser and Greenhouse (1958) adjusted degrees of freedom to test \(F\) value because the data did not meet the assumption of sphericity, \(\chi^2(2) = 13.11, p < .001\). The results indicated a significant main effect of Congruency \(F(1.77, 148.54) = 28.16, p < .001, \eta_p^2 = .25\); but neither a main effect of Culture, \(F(1, 85) = 1.56, p < .22\), nor a Culture \(\times\) Congruency interaction, \(F < 1\), was significant. Supporting our hypothesis, simple effect analysis within each cultural group showed that both European Canadian and East Asian participants reported that a happy target face surrounded with positive landscapes appeared more positive than the same faces surrounded with negative landscapes, \(F(1, 170) = 14.05, p < .001\) and \(F(1, 170) = 9.21, p < .001\), respectively (see Figure 3).\(^4\)

The results of Study 1 supported our hypothesis that, similar to Japanese, European Canadians were influenced by the congruency between target facial emotions and the salient valence of contexts in the absence of agents in the contexts. In Study 2, we tested the second hypothesis that the presence of agents in contexts would alter the impact of congruency on European Canadians but not on East Asians.
Study 2

Method

Participants. Ninety European Canadians (47 females and 43 males) at the University of Alberta and 113 Japanese (61 females and 52 males) at Kobe University participated in the experiment. In exchange for their participation, European Canadian participants received a course credit, and Japanese participants received a gift card worth about $5.

Stimuli. We used the same 16 target facial portraits (happy vs. sad) used in Study 1. We also selected pairs of background facial portraits from Ito et al.’s (2012) study; models showed either happy or sad facial emotions based on the same criteria as for the target portraits. A paired-samples t-test confirmed that the mean of happy background portraits were significantly different from the mean of sad background portraits among European Canadian participants, $t(17) = 17.26, p < .001$, and among Japanese participants, $t(19) = 25.20, p < .001$. Each background pair consisted of one male and one female who belonged to the same ethnic group as each other. We created two sets of 48 experimental stimuli by combining target facial portraits with a happy background pair, a sad background pair, or a white blank background. The ethnicity of target portraits and background portraits was the same in one set but was different in the other set (see Figure 4).

Procedure. Procedure was similar to that of Study 1. Participants evaluated the intensity of positive and negative expression of the target person (i.e., target was the center person if there were three people in an image, or the person in an image if there was only one person). Participants were instructed to rate positive intensity ranging from 0 (not at all) to 9 (extremely) followed by negative intensity ranging from 0 (not at all) to 9 (extremely). Stimuli remained on the monitor until participants finish the two intensity ratings. We first randomly presented target
faces combined with background faces (i.e., three-person image) and subsequently presented the target faces without background faces (i.e., control condition). Within the block of three-person image, we created two blocks of 16 trials to avoid presenting identical faces consecutively. The order of blocks was counterbalanced between participants. The experimental settings, and the size and the number of stimuli, were identical to those of Study 1. After answering a demographic information questionnaire at the end of the experiment, participants were fully debriefed.

**Results and Discussion**

We conducted a series of 2 (Culture: European Canadian vs. Japanese) × 3 (Congruency: congruent vs. incongruent vs. control) ANOVA, with participants’ intensity ratings being the dependent variable. Culture was the between-subjects variable, and Congruency was the within-subject variable. Because the data did not meet the assumption of sphericity for negative ratings on sad target, $\chi^2(2) = 13.11, p < .01$, and for positive ratings on happy target, $\chi^2(2) = 33.36, p < .001$, we used the Geisser and Greenhouse (1958) adjusted degrees of freedom to test $F$ value.

For the negative intensity ratings on sad target face, we found a main effect of Culture, $F(1, 201) = 22.30, p < .001, \eta^2_p = .10$, and Congruency, $F(1.88, 378.02) = 9.41, p < .001, \eta^2_p = .05$, and more important, the significant Culture × Congruency interaction, $F(1.88, 378.02) = 3.93, p < .05, \eta^2_p = .02$. Although a main effect of Culture was unexpected, this might reflect cultural differences in avoidance orientation toward potentially harmful stimuli that were prevalent (Hamamura, Meijer, Heine, Kamaya, & Hori, 2009). Supporting our hypothesis, within-subject ANOVA for each cultural group showed that the congruency between a target facial emotion and salient valence of surrounding others’ facial emotions significantly influenced Japanese participants’ ratings, $F(1.88, 378.02) = 12.65, p < .001$, but not Canadians’ ratings,
Simple effect analysis showed that Japanese participants reported that a sad target face surrounded with happy others (i.e., incongruent condition) were less negative than the sad face surrounded with sad others (i.e., congruent condition), $F(1, 402) = 12.88, p < .001$, or a sad face alone (i.e., the control condition), $F(1, 402) = 23.49, p < .001$ (see Figure 5).

For the positive intensity ratings of happy face, a main effect of congruency was significant, $F(1.73, 348.46) = 18.67, p < .001, \eta_p^2 = .09$; and more importantly, we found a significant Culture $\times$ Congruency interaction, $F(1.73, 348.46) = 3.71, p < .05, \eta_p^2 = .02$.

Supporting our hypothesis, within-subject ANOVA for each cultural group showed that the congruency between a target facial emotion and salient valence of surrounding others’ facial emotions significantly influenced Japanese participants’ ratings, $F(1.73, 348.46) = 21.99, p < .001$, but not Canadians’ ratings, $F(1.73, 348.46) = 2.58, p = .11$. Simple effect analysis showed that Japanese participants reported that a happy face surrounded with sad others (i.e., incongruent condition) was significantly less positive than the happy face surrounded with happy others (i.e., congruent condition), $F(1, 402) = 24.69, p < .001$, or a happy face alone (i.e., the control condition), $F(1, 402) = 34.55, p < .001$ (see Figure 6).

As expected, the results of Studies 1 and 2 showed that, unlike East Asians, European Canadians’ evaluations of facial emotions were less influenced by contexts with identifiable agents (i.e., pairs of people) compared to contexts without identifiable agents (i.e., landscape sceneries). According to our interpretation, this effect of context types results from the fact that, in general, North Americans consider facial emotions to be expressions of individuals’ private states of mind, and these expressions ought to be unaffected by other people’s expressions.
Therefore, in their evaluation of target facial emotions surrounded by other agents, North Americans successfully downplay the agency context.

### Study 3

The findings of Studies 1 and 2 suggest that East Asians’ judgment of target emotions is highly influenced by context information regardless of its types, whereas the context effect influences European Canadians’ judgment only when the context contains no identifiable agents. If that is the case, do European Canadians selectively downplay the agency contexts when presented with both agency and non-agency contexts? To answer this question, and to test the validity of the first two studies, we conducted a within-subject study combining the procedures of Studies 1 and 2, targeting only European Canadians.

In Study 3, we improved our method in order to rule out alternative explanations of the results. We pre-tested not only the valence but also the arousal levels of contextual images, verifying that images of landscapes and of pairs of people were equivalent in their intensity of emotional undertone. In addition, we examined whether response speeds differed when participants evaluated facial emotions in different types of context. Differences in exposure times could be a confounding variable when comparing the effect of background people with the effect of landscape sceneries. Our European Canadian participants might have spent more time evaluating target facial emotions with landscape backgrounds than those with people in the background. Thus, we measured the reaction times for participants’ evaluations to directly test whether exposure to agency vs. non-agency contexts affected the reaction times.

We also tested whether there is a direct association between North Americans’ cultural meanings and their distinct ways of evaluating facial emotions in different contexts with an original questionnaire. This questionnaire is designed to measure the agency belief of facial
emotions; that is, individuals’ endorsement of the belief that people control and are responsible for their facial emotions because facial emotions represent their internal state of mind and express unambiguous information about the self, without being affected by other people. We posited that people who endorse the agency belief of facial emotions are inclined to downplay contexts with identifiable agents more than contexts without identifiable agents.

Finally, we also added existing self-report measures of cultural orientations to test whether the measures were correlated with the intensity ratings of the targets’ facial emotions. Although several researchers have suggested that the correlation between some self-report measures and actual behavioral data is weak (Heine, Lehman, Peng, & Greenholtz, 2002; Kitayama & Uskul, 2011), self-report measures are often used to search for the meditation effect among variables.

Method

Participants. Twenty-nine European Canadian undergraduate students (14 females) at the University of Alberta participated in the experiment for partial fulfillment of a course credit.

Stimuli. When comparing the effect of different types of context, it is crucial that the emotional undertone of the two context types be equivalent. Thus, we conducted a pre-test to verify the equivalence of our contextual stimuli. Twenty-three European Canadian undergraduate students were presented with two types of contextual image—landscape and pairs of people—in random order. They were asked to evaluate on 9-point scale the intensity of valence and intensity of arousal for each image. A paired sample t test comparing their responses showed that participants reported equivalent arousal levels between the two types of contextual images, ts < 1. They also reported equivalent negative valence and positive valence, ts < 1.11, ps = ns, between the two types of contextual images. To create stimuli for Study 3, the same 16 target
facial emotions used in Studies 1 and 2 were combined with these images of landscapes and pairs of people. Unlike Studies 1 and 2, we excluded the control condition (i.e., white background). In total, we prepared 128 stimuli that differed in terms of type of context (landscape vs. surrounding people), valence of context and target (positive vs. negative), and gender and ethnicity of models.

**Questionnaires.** To measure situated beliefs in agency, we asked participants to report how they felt when the targets’ facial emotion was the same as the emotional tone of the background landscape or background people. On bipolar 9-point scales, participants indicated (a) their perceived discomfort, (b) difficulty of judgment, and (c) obscurity of target facial emotion. Based on the definition of agency belief in the current study, we created relative scores for situated agency for background people versus landscape sceneries (i.e., the average ratings when context was pairs of people minus the average ratings when context was landscape sceneries). Higher numbers indicate that participants experienced more discomfort, difficulty of evaluation, and mitigation of facial emotions when they evaluated target faces with background people than with landscape sceneries. Reliability of these three questions was sufficiently high, Cronbach’s α = .72 (Stangor, 1998), and inspection of eigenvalues and the scree plot suggested a one-factor model; thus, we averaged the scores for each participant.

We also included several questionnaires that might uncover an explanatory variable for European Canadians’ tendency to ignore contexts with identifiable agents. These questionnaires were chosen as they pertain to North Americans’ perception, emotional experience, or belief in agency. The Analysis-Holism Scale (AHS; Choi, Koo, & Choi, 2007) measures one’s endorsement of holism based on four subcomponents: causality, attitudes towards contradiction, perception of change, and locus of attention. The Implicit Social Orientation Questionnaire (ISOQ; Kitayama & Park, 2007) measures the intensity of experiencing socially engaged
emotions (e.g., calm or guilt) and socially disengaged emotions (e.g., pride or frustration) in 10 everyday situations. The scale of general belief in agency (Ryff, cited in Miyamoto & Ji, 2011) measures one’s tendency to actively change or manipulate the social environment.

**Procedure.** As in Studies 1 and 2, participants evaluated the intensity of the target person’s positive and negative expression. They were instructed to rate positive intensity, ranging from 0 (*not at all*) to 9 (*extremely*), followed by negative intensity, ranging from 0 (*not at all*) to 9 (*extremely*). Stimuli remained on the monitor until participants finished the two intensity ratings. In order to avoid the effect of participants’ fatigue, we created two equal sets of 64 stimuli that were counterbalanced between participants. Within each set, there were an equal number of target faces with landscape sceneries and with pairs of people; the order of presenting the two context types was also counterbalanced between participants. That is, half of the participants first rated target facial emotions against landscape backgrounds, and then rated target facial emotions with pairs of people in the background. The order was reversed for the other participants. At the end of the experimental sessions, participants answered a battery of questionnaires.

**Results**

**Response speed for evaluation.** We conducted a 2 (Context types: landscape vs. people) × 2 (Congruency: congruent vs. incongruent) ANOVA on participants’ response latency to evaluate negative intensity for sad target faces, and on their response latency to evaluate positive intensity for happy target faces. For evaluations of both happy target faces and sad target faces, we found a main effect of Congruency for happy target faces, $F(1, 28) = 4.56, p < .05$, and sad target faces, $F(1, 28) = 5.41, p < .05$. That is, participants responded faster to the congruent condition than to the incongruent condition. However, the effect of Context types and the
interaction effect were not significant, $F$s < 1. Participants evaluated target faces with a pair of people as quickly as target faces with landscape sceneries. Thus we eliminated participants’ exposure to stimuli as a cause of differential effect of context types on their facial emotion evaluation, and will not discuss this any further.

**Intensity evaluation.** We conducted 2 (Context types: landscape vs. people) × 2 (Congruency: congruent vs. incongruent) ANOVA on participants’ negative intensity evaluation of sad target faces and on their positive intensity evaluation of happy target faces.\(^{10}\) For the evaluations of sad target faces, we found a main effect of Congruency, $F(1, 28) = 13.47, p < .01$, and a marginally significant effect of Context types, $F(1, 28) = 3.21, p = .08$. These main effects were qualified by a significant interaction between the Context types and the Congruency, $F(1, 28) = 4.37, p < .05$. That is, the effect of the congruency between target facial emotion and emotional undertone of the background depended on whether the target faces were presented against landscape sceneries or against pairs of people.\(^{11}\) For the evaluations of happy target faces, we also found a main effect of valence, $F(1, 28) = 15.71, p < .01$. The interaction effect was marginally significant, $F(1, 28) = 2.92, p = .10$. Simple effect analyses revealed that European Canadian participants’ evaluations of sad target faces were influenced by the emotional undertone of landscape sceneries significantly more than that of background people, $t(28) = 2.23, p < .05$. Their evaluations of happy target faces were also more influenced by the emotional undertone of landscape sceneries than that of background people, although the effect was marginal, $t(28) = 1.71, p = .10$ (see Figure 7).\(^{12}\)

**Explanatory factors for the effect of context types.** We observed that although European Canadians took contexts without identifiable agents into account, they tended to ignore contexts with identifiable agents. To test the underlying factors, we first calculated the ratings for
the congruent condition minus the ratings for the incongruent condition for each context type, to represent the effect of congruency of affective information. A larger number indicates the perception of more intense target emotions when the emotional undertone of the background was congruent rather than incongruent with the target face. Then, we subtracted this intensity rating for background people from the rating for background landscapes. The resulting intensity ratings represented the relative impact of context types on participants’ evaluation of sad or happy facial emotion. Finally, we correlated the two intensity ratings with the four measures of possible underlying factors. As shown in Table 1, the situated agency scale was the only measure that was significantly correlated with the relative impacts of context types on sad target faces, $\beta = .69$, $t(28) = 4.13$, $p < .01$, and happy target faces $\beta = .42$, $t(28) = 2.10$, $p < .05$. That is, European Canadians who reported higher discomfort, higher difficulty of evaluation, and higher obscurity for evaluating target faces surrounded by other faces, were more likely to ignore the emotional undertone of surrounding others and to take into account that of landscape sceneries.

**General Discussion**

As previous findings have suggested, the results of the current studies indicate that contextual information influences facial emotion judgment. However, the condition in which the context is informative is highly dependent on the meaning of facial emotions shared by people in a given society. Both European Canadians and East Asians reported that the intensity of target facial emotions in affectively congruent contexts was higher than that in affectively incongruent contexts when the contextual information lacked identifiable agents (Study 1). However, when the target facial emotions were surrounded by agents (Study 2), European Canadians were not affected by such a manipulation, suggesting that they were able to sort out where such affective sources derived from. In support of our hypothesis, European Canadians dissociated background
facial emotions from the target person’s internal state of mind, and thus the background facial emotions were perceived as less important information for understanding the target facial emotions. However, European Canadians did incorporate the affective information of landscape sceneries in the evaluation of target facial emotions. The Study 3 results showed that this tendency was especially strong for those European Canadians who considered facial emotions as volitional and intentional actions. Several implications of the current studies are discussed below.

**Cultural Variations in Perception of Agency**

In the past, the idea that humans are attentive to agency produced a variety of social psychological research in Western societies, in areas such as dispositional attributions (Jones & Harris, 1967); moral judgment based on the justice principle (Kohlberg, 1976); and the effect of personal choice on motivation (Bandura, 1982). Recent findings in cultural psychology, however, have shown that findings regarding the perception of individuals as free agents might be specific to Western populations (see Henrich, Heine, & Norenzayan, 2010, for review). For example, compared to their Western counterparts, Koreans and Japanese were less likely to attribute the cause of one’s actions to personal disposition (Choi et al., 1999); Indians were more likely to perceive a conflict as the violation of interpersonal expectations rather than violation of justice (Miller & Bersoff, 1992); and Asian American children were more motivated to perform a task chosen by their parents, rather than a task chosen by the children themselves (Iyengar & Lepper, 1999). The current findings are given credence by these studies.

In what situation, then, do East Asians perceive agency? Evidence from other research indicates that East Asians perceive agency in the social groups to which people belong, rather than in individuals. For example, newspaper articles in Japan were likely to attribute causes of corporate scandals to organizational characteristics rather than to traits of individual workers
(Menon, Morris, Chiu, & Hong, 1999); and Chinese commercial arbitrators were likely to attribute organizations’ violations of contract terms to internal, rather than external, causes (Friedman, Liu, Chen, & Chi, 2007). In addition, as a result of emphasizing agency of the social group, East Asians were more likely than their European American counterparts to consider members of the social group as sharing similar characteristics (Spencer-Rogers, Williams, Hamilton, Peng, & Wang, 2007). The current findings shed light on this aspect again, and provide researchers with directions for future research to further elucidate the relationship between culture and agency.

**Amalgamation vs. Particularization of Meaning Systems**

For more than two decades, cultural psychologists have meticulously depicted cultural variations in psychological processes by using dichotomous explanations of culturally shared beliefs, such as individualism vs. collectivism, independent vs. interdependent self-construals, and holistic vs. analytic thoughts. Morling and Masuda (2012) recently reviewed these works and maintained that some psychological phenomena could be interpreted as the amalgamation of more than one belief. For example, many researchers reported that the contrast between holistic and analytic thoughts, together with the contrast between interdependent and independent self-construals, can complementarily help explain cultural phenomena such as fundamental attribution errors (Choi et al., 1999); self-awareness (Cohen & Hoshino-Browne, 2005); and social influence and power (Miyamoto & Ji, 2011; Morling et al., 2002). In general, these researchers discussed the influence of two dimensions: (a) interdependence and holistic perception, and (b) independence and analytic perception. Thus, “if one perceives oneself as embedded within a large context of which one is an interdependent part, it is likely that other objects or events will be perceived in a similar way” (Markus & Kitayama, 1991, p. 246); we call
such phenomena the *amalgamation of meaning systems*. Attention to context also resonates with a holistic understanding of the world, in which everything in the world is related to each other and no single element can be logically viewed as acting independently of anything else (Nisbett, 2003).

However, in the current studies, we examined the issue of mixed cultural beliefs by applying an alternative methodology to them and by articulating which of the cultural beliefs played a crucial role in creating cultural variation in psychological processes. In fact, previous research into Japanese people’s tendency to attend to surrounding others (e.g., Masuda et al., 2008; Masuda et al., 2012) did not clearly differentiate the explanation of interdependent self-construals from that of holistic attention. Thus, the question of whether cultural variation in context sensitivity can be interpreted as simply a matter of degree, contingent on one’s level of holistic tendency—or whether it should be considered a qualitative difference in the perception of agency, based on variations in cultural beliefs regarding self-construal—has not been addressed previously. To solve this problem, we conducted a series of cross-cultural and monocultural studies in which we presented salient backgrounds so that all participants would be easily influenced by them. We thus succeeded in focusing only on the discussion of cultural variation in construal—that is, the issue of whether the perception of agency influenced one’s pattern of attention. We maintain that this experimental design can be called the *particularization of meaning systems*, as opposed to the amalgamation of meaning systems. We believe that this research design can be applied to a variety of research topics to further scrutinize a variety of unresolved questions in the field of cultural psychology, thus allowing researchers to further unpack the relationship between cultural meaning systems and psychological processes.

**Automatic vs. Deliberate Processes of Judging Facial Emotions in Context**
The current studies attempted to scrutinize cultural similarities and differences in the process of facial emotion perception by asking participants to engage in the evaluation tasks under the experimental manipulation. However, there is an apparent discrepancy between the current findings and the findings of our previous studies. Ito et al.’s (2012) previous findings indicated that the emotional undertone of landscape backgrounds and facial emotion backgrounds equally influenced European Canadian and Japanese participants’ response speed to categorize facial emotions. The results of the current studies, by contrast, suggest that such culturally similar processes were attenuated when people engage in more time-consuming processes of facial emotion evaluation, such as intensity ratings of target facial emotions. We speculate that in general, the salient background information influences the judgment of both Canadians and Japanese. However, in the intensity rating task, the time lag between the stimulus onset and the judgment allowed participants to allocate their cognitive capacity to interpreting the meaning of the stimuli, which would in turn allow them to deliberately access the culturally dominant meaning system in their society. That is, as a default pattern of facial emotion judgment, both European North Americans’ and East Asians’ categorical judgments were equally influenced by changes in extremely salient context. However, European North Americans, but not East Asians, suppress the activated information of surrounding others’ facial emotions, because they believe that one’s facial expression derives mainly from internal factors rather than others’ emotions. Although such speculations are beyond the scope of the current studies, the findings shed light on the processes of automatic vs. deliberate information processing during the facial emotion judgment.

**Limitations and Future Directions**
The current research has several limitations. First, the East Asian groups who were the comparison groups differed between Study 1 and Study 2. However, the focus of this study is North Americans’ belief in agency of facial emotions, and we maintain that Asian international students in Canada and Japanese students in Japan both provided a valid comparison to show the distinctiveness of European Canadians compared to people from other cultural backgrounds. Second, we were unable to control for the size of the context stimuli. Although the intensities of affective information in the two types of contexts were equally salient (see Study 3), the background faces were significantly smaller than the landscape sceneries. Therefore, it is possible that North Americans failed to notice the changes in facial emotions of background people. To avoid such failure, we asked participants whether they noticed that the target facial emotions and contexts (i.e., background people or landscape sceneries) differed in affective information at the end of the experiment. All participants reported that they noticed the differences. Therefore, we argue that European Canadian participants in Study 3 intentionally downplayed background facial emotions in their target face evaluation. However, because the size of context stimuli co-varied with the two context types, further research is needed to test the effect with non-agent context stimuli that is the same size as a face. Finally, as other research suggests (Kitayama & Uskel, 2011), the correlation between the existing self-report measures and the intensity judgments of the target facial emotions was very weak. Although the self-report measures were often used as the meditational variables of cultural variation in human behaviors, the discrepancies between the abstract questions in the measures and the concrete behavior patterns seem to be substantial. This observation strongly resonates with the discussion of attitude–behavior discrepancy (e.g., Insko, Blake, Cialdini, & Mulaik, 1970). However, recent neuroscientific evidence shows a positive correlation between self-report measures and neural
data (Na, Grossmann, Varnum, Kitayama, & Nisbett, 2010). Our findings suggest that there must be unknown confounding variables which may constrain participants’ behavioral patterns (e.g., judgement, reaction time). In future studies, the relationship between attitude, behavior, and neural responses must be comprehensively investigated while controlling potential confounding variables associated with each measurement.
Footnote

1 Please see Ito et al. (2012) for the detailed procedure of the pre-test.

2 Ratings that were 3 SD away from the condition mean (2.14%) were replaced with the condition mean.

3 For the positive rating on sad target faces, we found a main effect of Congruency, $F(1.70, 144.05) = 12.28, p < .001, \eta_p^2 = .13$. Participants from both cultural groups reported higher positive intensity in the incongruent condition ($M = 1.23$) than in the congruent ($M = 0.88$) or the control conditions ($M = 0.93$).

4 For the negative rating on happy target faces, we found a marginally significant main effect of Culture, $F(1, 85) = 3.62, p < .06, \eta_p^2 = .04$, and Congruency, $F(1.30, 110.77) = 29.43, p < .001, \eta_p^2 = .26$. East Asian international students reported higher negative intensity than Canadian participants ($M_{Int\,EA} = .95$ vs. $M_{CND} = .68$), and participants from the two cultural groups reported higher negative intensity in the incongruent condition ($M = 1.13$) than the congruent ($M = 0.69$) or the control conditions ($M = 0.63$).

5 Testing biases due to the ethnicity of models, we conducted a $2 \times 2 \times 3$ ANOVA, with participants’ intensity ratings being the dependent variable. Culture and ethnicity of models were the between-subjects variable, and congruency was the within-subject variable. We found neither significant main effects nor interactions involving ethnicity of models, $Fs < 1.79, p = \text{ns}$; thus, we excluded the ethnicity of models from the following analyses and will not discuss it further.

6 Ratings that were 3 SD away from the condition mean (0.63%) were replaced with the condition mean.
For the positive intensity rating on sad target faces, we found a main effect of Culture, $F(1, 201) = 6.21, p < .05, \eta_p^2 = .03$, and Congruency, $F(1.90, 380.99) = 15.09, p < .001, \eta_p^2 = .07$. Japanese participants reported higher positive intensity than Canadian participants ($M_{JPN} = 1.64$ vs. $M_{CND} = 1.36$); participants from the two cultural groups reported higher positive intensity in the incongruent condition ($M = 1.67$) than in the congruent ($M = 1.42$) or control conditions ($M = 1.42$).

For the negative intensity rating on happy target faces, we found a main effect of Culture, $F(1, 201) = 26.90, p < .001 \eta_p^2 = .12$, and Congruency, $F(1.63, 328.24) = 27.47, p < .001, \eta_p^2 = .12$. Japanese participants reported higher negative intensity than Canadian participants ($M_{JPN} = 1.40$ vs. $M_{CND} = .89$), and participants from the two cultural groups reported higher negative intensity in the congruent condition ($M = 1.36$) than in the incongruent ($M = 1.11$) or control conditions ($M = 0.97$).

Responses that were 3 SD away from the condition mean (1.87%) were replaced with the condition mean.

Ratings that were 3 SD away from the condition mean (2.35%) were replaced with the condition mean.

For the positive intensity rating on sad target faces, we found a main effect of Congruency, $F(1, 28) = 7.40, p < .01, \eta_p^2 = .21$. Participants reported higher positive intensity in the congruent condition ($M = 1.72$) than in the incongruent condition ($M = 1.34$).

For the negative intensity rating on happy target faces, we found a main effect of Congruency, $F(1, 28) = 10.00, p < .01, \eta_p^2 = .26$. Participants reported higher negative intensity in the incongruent condition ($M = 1.25$) than in the congruent condition ($M = 0.80$).
References


Table 1 Association between the relative effect of contexts (landscapes minus pairs of faces) on sad or happy face evaluation and four relevant measures in Study 3.

<table>
<thead>
<tr>
<th>Evaluation Target</th>
<th>Situated Agency</th>
<th>General Agency</th>
<th>AHS</th>
<th>ISOQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sad Face</td>
<td>.69</td>
<td>.16</td>
<td>.08</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>4.13**</td>
<td>.89</td>
<td>.65</td>
<td>.19</td>
</tr>
<tr>
<td>Happy Face</td>
<td>.42</td>
<td>.12</td>
<td>.16</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>2.10*</td>
<td>.54</td>
<td>.81</td>
<td>.93</td>
</tr>
</tbody>
</table>

** <.01 * <.05
Figure 1. Stimuli used in Study 1 differed in terms of the valence of facial expression in the foreground and of landscape scenery in the background. Four possible combinations were (a) sad foreground face–negative landscape scenery, (b) happy foreground face–negative landscape scenery, (c) sad foreground face–positive landscape scenery, (d) happy foreground face–positive landscape scenery, (e) sad foreground face–no landscape scenery, and (f) happy foreground face–no landscape scenery.
Figure 2. Negative intensity rating on sad target face between European Canadians and East Asians in Study 1. Error bars represent standard error of the mean. * $p < .05$. ** $p < .01$. *** $p < .001$. 
Figure 3. Positive intensity rating on happy target face between European Canadians and East Asians in Study 1. Error bars represent standard error of the mean. * $p < .05$. ** $p < .01$. *** $p < .001$. 
Figure 4. Stimuli used in Study 2 differed in terms of the valence of facial expression in the foreground and of facial expression in the background. Four possible combinations were (a) sad foreground face–sad background faces, (b) happy foreground face–sad background faces, (c) sad foreground face–happy background faces, (d) happy foreground face–happy background faces, (e) sad foreground face–no background faces, and (f) happy foreground face–no background faces.
Figure 5. Negative intensity rating on sad target face between European Canadians and the Japanese in Study 2. Error bars represent standard error of the mean. * $p < .05$, ** $p < .01$, *** $p < .001$. 
Figure 6. Positive intensity rating on happy target face between European Canadians and the Japanese in Study 2. Error bars represent standard error of the mean. * $p < .05$. ** $p < .01$. *** $p < .001$. 
Figure 7: Differences in the effect of congruency manipulation on participants’ sad target evaluation and happy target evaluation with landscape sceneries or background faces as context. Congruency effect was calculated by subtracting intensity ratings for incongruent conditions from intensity ratings for congruent conditions. Error bars represent 1 standard error away from the condition means. * $p < .05$ † $p < .10$