Metacognition Fosters Cultural Learning:

Evidence from Individual Differences and Situational Prompts

Michael W. Morris

Columbia University

Krishna Savani

Nanyang Technological University

Katrina Fincher

Columbia University

May 24, 2018

In press, *Journal of Personality and Social Psychology*

Abstract

We investigated the role of metacognition in the process by which people learn new cultural norms from experiential feedback. In a lab paradigm, participants received many trials of simulated interpersonal situations in a new culture, each of which required them to make a choice, and then provided them with evaluative feedback about the accuracy of their choice with regard to local norms. Studies 1 to 3 found that participants higher on an individual difference dimension of metacognitive proclivity learned to adhere to the local norms faster. This relationship held up in simple and complex situations, that is, when the feedback was noisy rather than completely reliable, and it also held up when possibly confounding individual differences were controlled (Study 2). Further evidence suggested that the underlying mechanism is the largely implicit process of error monitoring and reactive error-based updating. A measure of surprise (an indicator of error monitoring) mediated the link between metacognitive proclivity and faster learning (Study 3). In experiments that varied the task so as to afford different kinds of metacognitive processing, participants learned faster with post-error prompts but not with post-accuracy prompts (Study 4). Further, they learned faster with non-directed prompts that merely provided a break for processing rather with prompts that directly instructed them to reason explicitly (Study 5). We discuss the implications of these findings for models of culture, first- and second-culture learning, and for training and selecting people for foreign or intercultural roles.

Keywords: metacognition; intercultural competence; learning; norms

Metacognition Fosters Cultural Learning:

Evidence from Individual Differences and Situational Prompts

Metacognition refers to monitoring and control of one's thought processes (Nelson & Narens, 1994). Developmental and educational theorists have long maintained that metacognitive processes play a critical role in learning (Dewey, 1933; Flavell, 1979). Students perform better in academic classes if they actively monitor their comprehension for errors (Kinnunen & Vauras, 1995) and adjust their assumptions after errors (Hirsh & Inzlicht, 2010). Likewise, adults higher in self-reported metacognitive proclivity learn better from the same training course (Ford, Smith, Weissbein, Gully, & Salas 1998; Schmidt & Ford, 2003).

But does metacognition matter outside of classroom learning? Long before schooling, our species thrived by learning cultural norms and skills (Chudek & Henrich, 2011). Our early ancestors learned their own tribe's norms so as to coordinate in foraging and self-defense (Sterelny, 2012). Some of them learned the norms of outgroups enough to permit intergroup exchange, more than among their Neanderthal rivals (Ambrose, 2010). The ability to learn foreign norms became more and more important as small-scale societies grew into large nation-states. Today, more people than ever before need to learn new cultural norms. The level of migration—by immigrants, expatriates, exchange students and refugees—is at an all-time historic high (The World Bank, 2015). At the same time, an increasing number of people marry individuals of different ethnicities, religions, or national origins, undergo midlife changes in their organization, or even enter occupations that require them to learn new norms (Sam & Berry, 2010). In sum, how people learn cultural norms—and any help that metacognitive processes offer—is increasingly important to understand.

Different kinds of cultural knowledge are likely acquired through different learning processes. Verbal instruction, whether in a classroom or around the dinner table, likely serves well for transmitting declarative knowledge, such as legal codes and religious prescriptions. However, people often learn complex skills through processes of imitating the actions (or sequences thereof) of a role model (Derex, Beugin, Godelle, & Raymond, 2013; Henrich, 2015). A less researched but equally important kind of cultural knowledge is interpersonal norms, the largely tacit dance steps of social life one follows to mesh with others in interactions such as greetings, favor requests, negotiations, and displays of deference (Hall,1966; 1983).

Organizations that seek to bring people up to speed for overseas roles—such as the military, the Peace Corp, and multinational corporations—assume that interpersonal norms are learned through trial and error experience (Bhawuk, 1998; Tung, 1993).
A widely used framework for this experiential learning describes cultural newcomers as starting with “unconscious incompetence” (or blissful ignorance), which then becomes “conscience incompetence” (or metacognitive awareness of mistakes) and sets the stage for learning the new behavioral pattern (Howell, 1982). People come to recognize their mistakes by picking up evaluative feedback from the locals that they interact with. There is much research tracking how emigrants, expatriates and exchange students fare abroad (Harris, 1973; Kealey & Protheroe, 1996; Abbe, Gulick, Herman, 2008; Caligiuri, 2000). However, this research does not measure learning, so it is not clear whether metacognition helps expatriates learn foreign norms.

In order to test whether people can learn the interpersonal norms of a new culture from experiential feedback, and whether this involves metacognitive processing of errors, we developed a laboratory paradigm that simulates the experience an expatriate would have with a particular class of situations in their first months in the new cultural setting. Participants encountered a long series of simulated interpersonal situations that required them to choose a behavior and then provided them with feedback about their accuracy vis-a-vis local norms. This task provided a behavioral measure of learning—participants would likely make more culturally appropriate choices as they learn the local norms across successive trials. We could therefore test whether individual differences in metacognitive proclivity are associated with faster learning, and try to trace the metacognitive processes at work. Finally, we conducted experiments manipulating task factors that afford particular kinds of metacognitive processing in order to test which of them are causally linked to faster learning.

Before developing the hypotheses, it is worth reviewing relevant literatures on cultural learning and metacognition.

Cultural Learning

A recent theme in social psychology is that implicit tendencies patterns are evoked by the social environment. Some prejudices, for instance, against African Americans and overweight people, develop as a result of exposure to television shows in which nonverbal negativity is expressed toward these groups (Weisbuch & Ambady, 2009; Weisbuch, Pauker, & Ambady, 2010). Likewise, Japanese expatriates who move to North America have increased self-esteem, whereas North Americans who move to Japan have decreased self-esteem (Heine & Lehman, 2004). It is possible that such self-esteem changes may reflect imitation; however, other studies find that implicit attentional biases, such as whether people encode objects contextually or decontextually, also differ for expatriates compared to peers who remained at home (Kitayama, Duffy, Kawamura, & Larsen, 2003).

These effects of exposure to the cultural setting can even be simulated in the laboratory. When American participants read about everyday Japanese situations (or vice versa), they came to exhibit some of the psychological patterns associated with the other culture, such as self-criticism versus self-enhancement, or secondary control versus primary control (Kitayama, Markus, Matsumoto, & Norasakkunkit, 1997; Morling, Kitayama, & Miyamoto, 2002). People take on the behavioral choices of another culture gradually with exposure to more and more interpersonal situations from that culture (Savani, Morris, Naidu, Kumar, & Berlia, 2011). However, these studies do not isolate the learning process involved, so we do not know whether people pick up patterns from the environment through reinforcement learning or through other mechanisms, such as imitation, affordance, or priming.

A longstanding literature on immigrant acculturation distinguishes strategies of engagement, such as assimilation versus separatism (Sam & Berry, 2010; Thomas & Znaniecki, 1918). Assimilation means participating primarily in the mainstream host-culture community, whereas separatism means participating primarily in the minority, heritage-culture community. On the surface, assimilation may seem to imply high second-culture learning and separatism its lack. However, separatism often arises in second-generation immigrants precisely because they have a greater understanding of the host culture and its prejudices (Portes & Rumbaut 2001). So engagement is not necessarily a proxy for learning.

Research on expatriates tends to focus on cultural adjustment, measured as subjective comfort with different aspects of the host culture (Bhawuk, Sakuda, & Munusamy, 2008; Black & Stephens, 1989). This is a valuable measure for applied questions, such as which expatriates will complete their assignment as opposed to terminate early, but it does not suffice to test arguments about cultural learning, as these outcomes might be driven by emotional coping processes rather than learning. Countless studies have correlated cultural adjustment with personality traits and aptitudes in search of criteria that can be used to select people for these roles (Harris, 1973; Kealey & Protheroe, 1996). More extroverted, agreeable, and emotionally stable sojourners tend to show better adjustment; more conscientious sojourners show better work/school performance (see Abbe et al., 2008; Caligiuri, 2000). Even though widely used for selection, measures of IQ do not correlate with adjustment (Morris, Savani, Mor & Cho, 2014; Morris, Savani, & Roberts, 2014).

Recently, researchers have sought to develop more specific measures of the qualities that enable intercultural effectiveness, labelling this as “cultural intelligence” (Ang et al., 2007; Leung, Ang, & Tan, 2014). Both of the leading cultural intelligence inventories involve scales measuring cultural metacognition (Ang et al., 2007; Thomas et al., 2008). These are self-reports of whether people are aware of, check, and update their assumptions in intercultural interactions (e.g., from Ang et al., 2007: “I check the accuracy of my cultural knowledge as I interact with people from different cultures” and “I adjust my cultural knowledge as I interact with people from a culture that is unfamiliar to me”; from Thomas et al., 2008: “When interacting with people from other cultures, I check on the accuracy of what I think I know about them” and “When interacting with people from other cultures, I ask myself how I'm feeling"). Theorists have posited that individuals higher in cultural metacognition are more likely to learn from everyday social experiences in the new cultural setting (Ng, Van Dyne, & Ang, 2009). However, researchers have not yet found a way to empirically measure experiential learning. High self-reported metacognitive proclivity on these scales has been found to correlate with exchange student's adjustment in the new country (Klafehn, Li, & Chiu, 2013; c.f. Shu, McAbee, & Ayman, 2017). That said, in a study of expatriate employees in Japan, after personality, demographics, and language ability were controlled, the effect of cultural metacognition on cultural adjustment was no longer significant (Huff, Song, & Gresch, 2014). According to other studies, metacognition predicts intercultural trust (Rockstuhl & Ng, 2008), teamwork (Crotty & Brett, 2012), and shared team values (Adair, Hideg, & Spence, 2013). However, these effects on collaboration appear to be affectively mediated, rather than cognitively mediated, so the role of learning, if any, remains unclear (Chua, Morris & Mor, 2012).

In sum, although many literatures reference cultural learning, none of them directly measure learning, or examine whether learning is fostered by metacognitive processing.

Metacognitive Processes

One way of studying the effects of metacognition is through individual differences (Fleming, Weil, Nagy, Dolan, & Rees, 2010; Molenberghs, Trautwein, Böckler, Singer, & Kanske, 2016; Pintrich, & De Groot, 1990; Schraw & Dennison, 1994). In academic contexts, individual differences in self-reported proclivity toward metacognition predicts how well students learn from classes or trainings (Ford et al., 1998; Schmidt & Ford, 2003). More extreme differences are seen in the case of cognitive disorders; domain-specific metacognitive deficits have been implicated in disorders such as dyslexia, math disability, and problem gambling (Brevers et al., 2014; Garrett, Mazzocco, & Baker, 2006; Mason & Mason, 2005; Job & Klassen, 2012). This research suggests that metacognitive proclivity or sensitivity is at least partly domain-specific (Kelemen, Frost, & Weaver, 2000). For instance, a person with Asperger's syndrome may be able to monitor himself actively for math errors but not for social errors.

Recent cognitive neuroscience studies have increasingly focused on spontaneous processes of error monitoring and error-based updating. These refer to implicit metacognitive mechanisms, in contrast to explicit metacognitive processes such as reflective reasoning and planning (Frith, 2012). Error monitoring begins with confidence, the feeling of knowing. Metacognitively active decision makers not only choose an answer or action, but also implicitly estimate how likely it is that they are correct (Charles, Opstal, Marti, & Dehaene, 2013). High confidence makes people less surprised when given the feedback that they were correct but more surprised when given feedback that they were wrong. Error feedback despite high confidence induce reactive corrective processing (Butterfield & Metcalfe, 2001; Fazio & Marsh, 2009; Reder & Schunn, 1996). Although confidence judgments and reactive updating in response to errors occurs largely at an implicit level (Frith, 2012), surprise is a part of the process that learners can consciously access and accurately report (Charles et al., 2013).

The first evidence for a reactive error-based correction or updating process was the finding of post-error slowing—people pause longer after errors than after accurate responses (Gehring, Goss, Coles, Meyer, & Donchin, 1993; Rabbitt, 1966). Post-error slowing correlates with a distinctive brain process localized in the anterior cingulate cortex (ACC), and is reflected in a brain signal called error-related negativity (ERN) (Gehring, et al., 1993; Holroyd, Dien & Coles, 1998; Krigolson & Holroyd, 2006; Wessel, Danielmeier, Morton & Ullserger, 2012; Yeung & Summerfield, 2012). Individuals with stronger ERNs learn more from their errors and consequently exhibit improved task performance (Legault & Inzlicht, 2013); they also exhibit better downstream outcomes, such as higher academic achievement in college (Hirsh & Inzlicht, 2010). Individual differences in post-error ACC activity and ERN are associated with improved accuracy when presented with the same problem again after an initial error (Hester, Madeley, Murphy, & Mattingley 2009; Holroyd & Coles, 2002), suggesting that error-based updating may be metacognitive process that functions to help us avoid repeating mistakes.

While social and cognitive psychology has studied metacognition as an individual difference, and ensuing processes such as surprise and post-error slowing, research in education and human-computer interaction has studied situational metacognitive prompts. Prompts are messages from a teacher, tutor, or tutorial that interrupt the normal rhythm of the lesson to elicit reflective processing (Wirth, 2009). A prompt can be just a pause to allow the learner to fully process some feedback, or it can be directed instructions to reflect on their knowledge or problem-solving strategy (Schraw, 1998; Tanner, 2012). The two basic questions in this research are (1) when to prompt and (2) what to prompt (Crook & Beier, 2010; Thillmann, Künsting, Wirth, & Leutner, 2009). “When” refers to the question of timing; a break in the flow may help some junctures but hinder at others (Thillmann, et al, 2009). “What” refers to content of the message, whether it should be a non-directive break, or an instruction to reflect on some aspect of one's problem-solving (e.g., instructions to analyze where one's reasoning went awry).

Directed and non-directed prompts may be helpful in different learning environments. Directed prompts have been found to stimulate classroom learning (Lin & Lehman, 1999; Bangert-Drowns, Hurley, & Wilkinson, 2004; Davis & Linn, 2000). Non-directed prompts (which some researchers call "generic prompts") have been found helpful in the context of independent individual learning (Davis, 2003; Ifenthaler, 2012). Non-directed prompts are thought to be effective in the context of individual learning because they do not interfere with the learner's autonomous, automatic processing. In sum, directed prompts are useful for eliciting explicit metacognitive processes, whereas less-directed prompts are more useful for eliciting implicit metacognitive processing.

Current Research and Hypotheses

In the current research, we tested for the first time whether metacognition fosters the experiential learning of cultural norms. To this end, we developed a laboratory task that simulates the experiential feedback in an expatriate's initial months abroad. Participants encountered dozens of interpersonal interactions of a given type, each requiring a behavioral decision and then providing accuracy feedback. Our first hypothesis is that individual differences in metacognitive proclivity would be associated with participants' speed of learning the new cultural norms (Hypothesis 1).

The next question is how this individual difference foster learning. Based on past theory, we propose that these individuals engage more in the process of error-monitoring and error-based updating. Because error monitoring involves forming feelings of confidence about one's decisions, it can be gauged by surprise responses to feedback. When a learner feels confident about the right answer in a situation, she would feel less surprise at (reliable) accuracy feedback and more surprised at (spurious) error feedback. Therefore, our second hypothesis is that the effect of metacognitive proclivity on learning would be mediated by this indicator of error monitoring activity—the differential surprise felt toward reliable versus spurious feedback after correct answers (Hypothesis 2).

In addition to measuring individual differences, metacognitive processes can also be prompted situationally. Students who are working through problems and receiving feedback are sometimes benefited from a break in the flow that spurs metacognitive processing (Wirth, 2009). Prompts can be highly-directed instruction to analyze something, or just a pause that enables reflection. The questions of when and what to prompt depends on the metacognitive mechanisms involved. Given that our proposed mechanism for experiential norm learning is error-based updating, it follows that prompts would aid learning when they come after errors but not when they come after accurate responses (Hypothesis 3). Given that error-based updating is an automatic, implicit process, it follows that non-directed prompts would aid learning more than directed prompts (Hypothesis 4).

Overview of Studies

We present five studies that tested the above hypotheses. Study 1 examined whether individual differences in cultural metacognition predict the speed with which people learn foreign norms from experiential feedback. Study 2 replicated this experiment using a different method of simulating interpersonal interactions in new culture which is a more controlled procedure, while controlling for many other individual differences that may be confounded with cultural metacognition. In addition to the cultural learning task, Study 3 added a second task that gauged participants' level of error-monitoring activity through their surprise responses.

The final two studies were experiments. Study 4 varied the timing of metacognitive prompts: post-error pauses, post-accuracy pauses or no pauses at all. Study 5 varied their content: directed prompts that instructed participants to analyze their mistake or non-directed prompts which merely announced an incidental pause. From our proposed metacognitive mechanisms, we derived the hypotheses that prompts should aid learning most when they occur after errors and when their content is non-directed rather than directed. Figure 1 depicts the relationship between the studies. Across all studies, all participants, experimental conditions, and cultural learning tasks are reported.

<Insert Figure 1>

Data Analysis Strategy. All our studies involved a learning task in which participants had to make a decision across many trials (either 50 or 80). Each decision was coded as *correct* or *incorrect* based on the cultural norms that participants had to learn. We first analyzed the data using a hierarchical logistic regression that assesses the change in participants' accuracy on a trial-by-trial level. We ran this analysis using the *xtlogit* command in STATA©. Level 1 slopes were fixed at Level 2 to ensure that the model converges, and a population average model was used. Simply stated, this model runs a logistic regression within each participant to estimate the slope of each participant's learning curve in the log odds unit, and then estimates how participants' slopes vary as a function of their metacognition score (see Raudenbush & Bryk, 2002, Chapter 10, for details). To increase ease of interpretation, we coded the trial order to range from -.5 (first trial) to .5 (last trial), such that the effect of trial order represents the change in accuracy over the entire experiment. The dependent measure was participants' accuracy on each trial, and the predictor variables were trial order, individual difference measures (all mean centered) / condition dummies, and interaction terms.

To assess the robustness of the findings, we conducted an additional linear analysis. To test for linear and curvilinear learning effects, we divided the trials into three sets (early, middle, and late), and then computed participants' average accuracy within each set. A repeated measures ANOVA could not be used because it assumes that the trial sets represent three separate categories, when in reality, they represent one continuum (early, middle, late). We thus analyzed the data using hierarchical linear models (HLM), treating trial sets as nested within participants (see Raudenbush & Bryk, 2002, Chapter 2, for details). We ran this analysis using the *xtmixed* command in the STATA© data analysis software. Level 1 slopes were allowed to vary at Level 2, the covariance structure was unrestricted, and robust standard errors were used. The dependent measure was participants' accuracy in each trial set, and the predictor variables were trial set (coded -.5, 0, .5), individual difference measures (all mean centered) / condition dummies, and interaction terms. As there were no quadratic effects of the trial set in any of the studies, only models estimating linear effects of trial set are reported herein.

Study 1 – Metacognitive Proclivity and Cooperation Norms

Our first study exposed students in India to the everyday influence situations that American college students had encountered. In each situation, participants read a brief description of a person trying to influence another person and had to choose whether to accommodate or not. Afterward, they received feedback about whether their choice was considered appropriate by American cultural standards.

The task was difficult because the frequent types of influence situations and the norm about how to respond to them vary starkly between the US and India. In the US, most influence attempts are self-serving efforts to benefit the influencer. Accordingly, Americans frequently resist others' requests and regard this as appropriate. In India, conversely, most influence attempts are other-serving efforts to help the influencee. Accordingly, Indians are highly likely to accommodate and regard this as appropriate (Savani et al., 2011). As newcomers to the US, Indians have to learn the types of influence situations that they should resist.

We measured cultural learning by assessing Indian participants' change in accuracy over successive trials. Accuracy means that the participants made the decision which was considered appropriate by the American standard. We predicted an interaction between metacognitive proclivity and trial order, such that participants who are higher in metacognitive proclivity would exhibit a greater increase in accuracy over successive trials.

Method

Participants. We decided on a sample size of 40 participants, similar to the sample size of previous studies that used the same stimuli (Savani et al., 2011, Studies 4 and 5). We posted a study seeking to recruit 40 participants over two weeks at a selective university in Bangalore, India. Only 38 undergraduates (14 women, 24 men;

mean age 20.84 years; all ethnic Indians and Indian citizens) completed the study before the end of the semester.

Materials. Participants were exposed to 50 trials. Each trial came with a brief description of an interpersonal influence situation representative of the everyday experiences of American college students. As illustrated in Figure 2, each trial described (1) the relationship between influencer and influencee, (2) the decision options which the influencee faced, (3) the influencee's initial inclination versus the influencer's requested action, and (4) the influencer's motive. For instance: "Suppose you are with a friend and you are deciding between taking a nap or going to a lecture. You originally preferred to take a nap but your friend is influencing you to go to lecture so that the two of you could sit together."

The stimulus descriptions were standardized abstractions of incidents sampled from US students in a previous study (Savani et al., 2011, Study 2). All local references were transformed into more culturally general ones (e.g., *Christmas* was changed to *a religious holiday*). In a previous study, a group of American participants had rated whether accommodation would have positive or negative consequences in each influence situation (Savani et al., 2011, Study 4). Based on whether their mean rating was above or below the scale midpoint, we categorized each situation as being accommodation-rewarding or not. Savani et al. (2011, Study 4) found that 92% of India-sourced influence situations, but only 48% of US-sourced situations, were accommodation-rewarding.

Procedure. The task was programmed using the DirectRT© software (Jarvis, 2010). Each Indian participant saw the 50 US-sourced influence situations in a different random order. Participants were explicitly informed that the situations were sourced from the US and that their goal was to learn to act like Americans. The descriptions stayed on the screen for a minimum duration, depending upon the paragraph length (400 milliseconds per word). Thereafter, participants had to make a binary decision about whether to accommodate, “to do what the other person wants you to do,” or to resist, “to NOT do what the other person wants you to do.”

Immediately after their response, participants received feedback based on whether their decision matched the decision endorsed by a majority of Americans. Specifically, participants saw either “CORRECT + 100 points” or “WRONG -100 points” displayed on the screen in green and red color font, respectively, for two seconds. In order to increase engagement with the task, participants were told to try to accumulate as many points as possible. The most accurate participant received a bonus of Indian Rupees 500, which amounted to approximately US$10 at the time of the study. As the language of instruction at the university is English, the entire study was run in English. Figure 2 illustrates the study procedure.

<Insert Figure 2>

After the situation decision task, participants completed the cultural metacognition measure from Thomas et al.’s (2008) cultural intelligence instrument (e.g., “When interacting with people from other cultures, I select and organize the cultural knowledge I need to use”; scale: 1 = “Strongly disagree” to 7 = “Strongly agree”; α = .57).

Results

Hierarchical logistic model. We first analyzed the data using a hierarchical logistic regression that assesses the change in participants’ accuracy on a trial-by-trial level. To increase ease of interpretation, we coded the trial order to range from -.5 (first trial) to .5 (last trial). The dependent measure was participants' accuracy on each trial, and the predictor variables were trial order, cultural metacognition (centered), and the interaction. The main effect of cultural metacognition was nonsignificant, $B = -0.011$, 95% CI [-0.12, 0.10], $SE = .057$, *odds ratio* = 0.99, $z = 0.19$, $p = .85$. The main effect of trial order was also nonsignificant, $B = -0.016$, 95% CI [-0.33, 0.29], $SE = .16$, *odds ratio* = 0.98, $z = 0.10$, $p = .92$. However, there was a significant trial order X metacognition interaction, $B = 0.47$, 95% CI [0.068, 0.87], $SE = .20$, *odds ratio* = 1.60, $z = 2.29$, $p = .022$. The odds ratio of 1.60 indicates that a one-unit increase in cultural metacognition (on a 7-point scale) was associated with a 60% increase in participants' odds of being accurate on the last trial relative to the first trial.

Hierarchical linear model. We next conducted hierarchical linear analyses. We divided the 50 trials into three sets of 17, 17, and 16 trials, respectively, and computed participants' accuracy in each of the three sets. We ran a hierarchical linear regression treating trial sets as nested within participants. The dependent measure was participants' accuracy in each trial set, and the predictor variables were trial set (coded -.5, 0, .5), participants' cultural metacognition score (centered), and the trial set X metacognition interaction. The main effect of trial set was nonsignificant, $B = 0.0050$, 95% CI [-0.046, 0.056], $SE = .026$, $z = 0.19$, $p = .85$. The main effect of cultural metacognition was nonsignificant, $B = -0.0016$, 95% CI [-0.024, 0.021], $SE = .011$, $z = 0.14$, $p = .89$. However, there was a significant trial set X metacognition interaction, $B = 0.082$, 95% CI [0.021, 0.14], $SE = .031$, $z = 2.63$, $p = .009$. Thus, a one-unit increase in cultural metacognition (on a 7-point scale) increased participants' learning (i.e., gain in accuracy from the first trial set to the last) by 8.2%.

Next, we conducted simple slopes analyses to assess learning at high vs. low levels of metacognition. At one standard deviation below the mean on metacognition, the main effect of trial set was nonsignificant, B = -0.060, 95% CI [-0.14, 0.017], SE = .039, z = 1.52, p = .13. However, at one standard deviation above the mean on metacognition, there was a significant positive effect of trial set, B = 0.070, 95% CI [-0.0069, 0.13], SE = .032, z = 2.17, p = .030. Thus, the accuracy of participants high in cultural metacognition increased by 7% from the first to the last trial set (see Figure 3).

<Insert Figure 3>

Discussion

Study 1 supported the first hypothesis, that participants' metacognitive proficiency is associated with the speed with which they learn foreign norms from experiential feedback. Indians who reported that they habitually evaluate and correct their intercultural assumptions were able to learn American norms faster when they made decisions in a series of influence attempts that had occurred in the US.

An important question that arises from these findings is whether this laboratory simulation of learning cultural norms captures the same learning processes that play out over a longer time course as, for instance, those in the case of an expatriate's first few months in a foreign setting. A study that involved Singaporean undergraduates studying abroad, primarily in Europe and North America, gave us an opportunity to test this assumption (Savani, Phua, Morris, & Hong, 2015). Participants completed the Study 1 experiential learning task in a lab session. Two months into their foreign stay, they completed an online survey. The survey measured their self-reports of how much they had learned the local cultural norms in 20 different domains (adapted from Ward & Kennedy, 1999). Participants who performed better in our cultural learning task also reported that they learned more local norms during their first two months abroad. This finding indicated that the lab learning task taps into the same capabilities required for real experiential learning of another culture's norms, despite the fact that the latter plays out over a longer period of time.

Undoubtedly, the validity of our expatriate simulation task partially came from the fact that Indian participants had to learn an actual cultural norm. They encountered a representative sample of American influence situations and tried to learn the tacit American standard about when to go along with requests and when to say *no*. Even American participants themselves would probably have difficulty articulating the rule so that it fully captures the consensual answers about the 50 situations in this study. Although this aspect makes the task realistic, the highly verbal nature of the task — participants were presented with 50 long paragraphs in succession — is an artificial aspect. The fatiguing nature of the task makes it possible for some third variable, such as personality or intelligence, to drive both metacognitive proclivity and learning. Therefore, in our next study, we switched to a less taxing and more pictorial way of simulating interpersonal encounters in a new cultural setting.

Study 2 – Metacognitive Proclivity and Greeting Norms

The goal of Study 2 was to re-test the first hypothesis with a different participant population. In addition, we used a different method for simulating interpersonal interactions and included more control variables. Specifically, in an online survey, we first measured participants' cultural metacognition, along with a host of potentially related constructs. At least one week after they completed the online survey, we brought the participants to the lab and administered a multi-trial norm learning task.

Instead of presenting participants with written descriptions of interpersonal interactions, we showed them faces of individuals hypothetically encountered in a new culture. We also gave participants some information about the setting in which they encountered the foreign individuals, for instance, daytime or nighttime, indoors or outdoors, sitting or standing. Similar to the written descriptions in Study 1, the trials varied on a large number of dimensions, including visual details of the photograph (e.g., gender, age, attractiveness, and dominance of the person) as well as verbal details about the context of the interaction. Through trial and error, participants had to figure out which cue to use and which to ignore. In each encounter, participants had to decide whether to greet the other person with a wave or a nod. We also provided participants with feedback about whether their decision matched an underlying cultural rule. We added noise to the feedback in order to mimic the fact that social interactions do not always provide clear and accurate feedback.

Participants were randomly assigned to two conditions that varied the complexity of the information about the context in which participants encountered the foreigner. In the less complex version, we gave participants only one verbal cue. In the more complex version, participants received three verbal cues. In both cases, only one of these cues, and none of the salient visual cues in the photograph, was relevant to the cultural greeting rule. The manipulation explored whether the benefits of metacognitive proclivity for learning is primarily evident in easier tasks, more complex tasks, or both types of tasks alike.

Method

Participants. As there were two conditions in this study, we decided on a sample size of 100 participants. To account for attrition, we posted a survey on the Columbia Business School subject pool seeking 120 European Americans for a two-part study. In response, 111 participants completed the first part, which was an online survey. After one week, these 111 participants were invited to the lab to complete the second part. Only 65 participants responded to the invitation. Of these, two participants reported being non-European American and thus were excluded. The final sample consisted of 63 European American participants (43 women, 20 men; mean age 23.60 years). Participants were randomly assigned to the less complex or the more complex condition.

Individual difference measures. In the online survey, participants were asked to complete a series of individual difference measures. The survey was administered at least a week before the cultural learning task. This was done in order to prevent participants' answers to these subjective questions from being biased by their performance on the learning task, or vice versa. The survey included Ang and Van Dyne's cultural intelligence instrument (Ang et al., 2007, which includes the key measure of cultural metacognition), big-five personality (Gosling, Rentfrow, & Swann, 2003), ethnocentrism (Neuliep & McCroskey, 1997), and negative emotions during intercultural communications (Spencer-Rodgers & McGovern, 2002). These scales were measured on 7-point response scales ranging from *strongly disagree* to *strongly agree*. Participants also completed the willingness to engage in intercultural communications scale (Kassing, 1997), in which they indicated the percentage of times they would be willing to interact with someone different from them. Finally, participants completed the 18-item short form of the Raven's progressive matrices (Bilker et al., 2012).

To limit the length of the online survey, a few measures that involved objective biographical questions were administered during the laboratory session. These included the multicultural experiences scale (Leung & Chiu, 2010), which is a collection of questions using response formats such as percentages, binary responses, and counts ranging from 1 to 5. Finally, we included a measure of the time participants had lived outside the US (Lu et al., 2017; Maddux & Galinsky, 2009). See Table S1 in the Supplementary Materials for the mean, standard deviations and reliabilities of all the individual difference measures, along with their intercorrelations.

Intercultural learning task. We increased the number of trials from 50 to 80 as the task was much easier compared to that used in Study 1. This time, participants did not have to read multiple paragraphs describing each interpersonal interaction but instead, had to respond to faces with only one line of accompanying textual information. This task was programmed using the Inquisit® software and was administered in the lab. Participants received the following instructions:

> "In this study, we want to see how people learn to interact with individuals from other cultures. Imagine that you move to Kyrgyzstan, a Central Asian country, and interact with many people there (80 different individuals in all). Your job is to figure out the appropriate ways to greet people in Kyrgyzstan. You know that they sometimes wave and sometimes nod to greet each other. For each individual you interact with, you will see their picture as well as a description of the circumstances in which you meet them. All of this information may be relevant to the appropriate greeting, so please attend to it carefully. Every time you encounter someone, you need to choose whether to wave or nod. Based on your greeting, you will see either Correct or Wrong. Your job is to learn how to appropriately greet people in Kyrgyzstan."

Afterward, participants were presented with 80 trials, which presented portrait photographs of Asian young adults, 40 men and 40 women. The photographs were taken from the Asian Emotion database (Wong & Cho, 2007a, 2007b), along with information about the context of the meeting. We separately randomized all trials for each participant. In the less complex condition, a statement below the photograph informed participants about the person's position ("This person is standing" or "This person is sitting"), the person's location ("This person is inside" or "This person is outside"), or the time of the day ("You see this person during the day" or "You see this person during the night"). Participants in this condition were randomly assigned to the position, location, or time of day sub-condition and received only one of the three cues (i.e., either position or location or time of day). In the more complex condition, participants received all three cues simultaneously, e.g., "This person is outside during the day and standing." We counterbalanced the order of the three cues across three sub-conditions, to which participants were randomly assigned. After two seconds, we asked participants, "How would you greet this person? Press W to Wave and N to Nod." Immediately after participants pressed *W* or *N*, they received feedback, "Correct" or "Wrong," based on whether their response matched the underlying rule that had to be learned.

In the less complex, single verbal cue condition, the learning rule was based on the cue displayed, e.g., "Wave to people outside but nod to people inside." In the more complex multiple verbal cues condition, the learning rule was based on the first cue displayed, while the other two verbal cues were irrelevant. In 90% of the trials, participants received the "Correct" feedback if they chose the greeting that was consistent with the rule, and “Wrong” if they chose the greeting that was inconsistent with the rule. To increase the difficulty of the learning task, we added noise in 10% of the trials, such that if participants responded consistently with the rule, they received negative feedback, but if they responded inconsistently with the rule, they received positive feedback.

Results

See Tables S1 and S2 in Supplementary Materials for descriptive statistics, and for results of all regression analyses.

Hierarchical logistic model. We first analyzed the data using a hierarchical logistic regression, treating trials as nested within participants. The dependent measure was whether on each trial, participants chose the greeting that was consistent with the underlying rule (correct = 1, incorrect = 0). The predictor variables were the experimental condition (less complex = -.5, more complex = .5), trial order (first trial = -.5, last trial = .5), cultural metacognition (centered), and all interactions.

A main effect of condition indicated that participants’ accuracy was lower in the more complex condition than in the less complex condition, B = -1.14, 95% CI [-1.74, -0.55], SE = .30, *odds ratio* = 0.32, z = 3.76, p < .001. A main effect of trial order indicated that participants’ responses became more accurate with subsequent trials, B = 1.59, 95% CI [1.23, 1.95], SE = .18, *odds ratio* = 4.92, z = 8.64, p < .001. A main effect of metacognition indicated that participants who scored higher on cultural metacognition were more accurate overall, B = 0.33, 95% CI [0.023, 0.64], SE = .16, *odds ratio* = 1.40, z = 2.11, p = .035. A metacognition X condition interaction indicated that the effect of metacognition on the higher overall accuracy was weaker in the more complex condition, *B* = -0.66, 95% *CI* [-1.28, -0.036], *SE* = .32, *odds ratio* = 0.52, *z* = 2.07, *p* = .038. A condition X trial order interaction indicated that participants' speed of learning was slower in the more complex condition, *B* = -1.14, 95% *CI* [-1.86, -0.41], *SE* = .37, *odds ratio* = 0.32, *z* = 3.08, *p* = .002. More importantly, a significant metacognition X trial order interaction indicated that more metacognitively inclined participants learned the underlying greeting norms more quickly, *B* = 0.68, 95% *CI* [0.34, 1.03], *SE* = .18, *odds ratio* = 1.98, *z* = 3.87, *p* < .001. Thus, a one-unit increase in cultural metacognition (on a 7-point scale) doubled participants' learning (i.e., gain in odds of being accurate from the first trial to the last). The three-way condition X metacognition X trial order interaction was nonsignificant, *B* = -0.28, 95% *CI* [-0.97, 0.41], *SE* = .35, *odds ratio* = 0.76, *z* = .78, *p* = .43. This indicates that metacognition predicted participants' speed of learning to a similar degree in both the less complex and more complex conditions.

Next, we conducted simple slopes analyses to assess learning at high vs. low levels of metacognition. At one standard deviation below the mean on metacognition, we observed a significant positive effect of trial order on accuracy, *B* = 1.02, 95% *CI* [0.68, 1.37], *SE* = .18, *odds ratio* = 2.78, *z* = 5.77, *p* < .001. This indicated that even participants low on cultural metacognition became more accurate across successive trials. At one standard deviation above the mean on metacognition, the positive effect of trial order on accuracy was much stronger, *B* = 2.16, 95% *CI* [1.61, 2.72], *SE* = .28, *odds ratio* = 8.69, *z* = 7.66, *p* < .001, as indicated by an odds ratio that was nearly four times as large as that for participants who were one standard deviation below the mean on metacognition.

Additional analyses. We further tested whether only the metacognition dimension of Ang et al.'s (2007) cultural intelligence scale predicts learning, not the other dimensions (knowledge, motivation, and behaviour). To conserve degrees of freedom, as metacognition similarly predicted the slope of participants' learning curves across the less complex condition and the more complex condition, we did not include interactions between condition and the individual difference measures. We entered the four dimensions of the cultural intelligence scale (centered) in the model, along with each of their interactions with trial order. Again, there is a main effect of condition, indicating that participants were less accurate in the more complex condition than the less complex condition, $B = -0.99$, 95% CI [-1.56, -0.43], $SE = .29$, *odds ratio* = 0.37, $z = 3.45$, $p = .001$. There is also a main effect of order, such that participants' accuracy increased with the number of trials, $B = 1.54$, 95% CI [1.22, 1.87], $SE = .16$, *odds ratio* = 4.68, $z = 9.40$, $p < .001$. A condition X trial order interaction indicated that participants' speed of learning was slower in the more complex condition, $B = -0.93$, 95% CI [-1.56, -0.30], $SE = .32$, *odds ratio* = 0.40, $z = 2.89$, $p = .004$. There is no main effect of metacognition, $B = 0.30$, 95% CI [-0.086, 0.68], $SE = .20$, *odds ratio* = 1.34, $z = 1.52$, $p = .13$. There are no main effects of cultural knowledge, $B = 0.011$, 95% CI [-0.30, 0.32], $SE = .16$, *odds ratio* = 1.01, $z = 0.07$, $p = .95$; cultural motivation, $B = 0.21$, 95% CI [-0.17, 0.59], $SE = .20$, *odds ratio* = 1.23, $z = 1.08$, $p = .28$; and cultural behavior, $B = -0.23$, 95% CI [-0.54, 0.081], $SE = .16$, *odds ratio* = 0.79, $z = 1.45$, $p = .15$. The interactions of these dimensions with trial order are also nonsignificant, cultural knowledge X trial order, $B = 0.20$, 95% CI [-0.13, -0.54], $SE = .17$, *odds ratio* = 1.23, $z = 1.20$, $p = .23$; cultural motivation X trial order, $B = 0.10$, 95% CI [-0.30, 0.50], $SE = .20$, *odds ratio* = 1.11, z = 0.50, p = .62; cultural behavior X trial order, B = -0.11, 95% *CI* [-0.45, 0.23], *SE* = .17, *odds ratio* = 0.90, z = 0.61, p = .54. However, we found that the metacognition dimension, in particular, predicted participants' speed of learning. The interaction between metacognition and trial order was significant, B = 0.51, 95% *CI* [0.094, 0.92], *SE* = .21, *odds ratio* = 1.66, z = 2.41, p = .016.

Additional analyses controlling for all the alternative variables measured are reported in the Supplementary Materials. They were tested one at a time in order to ensure sufficient power in the analyses. The metacognition X trial order interaction remained statistically significant in all analyses.

Hierarchical linear model. We next conducted hierarchical linear analyses. We first divided the 80 trials into three sets of 27, 27, and 26 trials, respectively, and computed participants' average accuracy in each of the three sets. We found that in the third trial set, the most frequent accuracy rate was the extreme right end of the distribution—26% of the participants had 100% accuracy. To approximate the normal distribution, the peak frequency must be in the middle of the distribution. Thus, it was not advisable to analyze the data using a linear model. In all other studies, the mode of the dependent measure was not at an end-point of the distribution.

Discussion

Study 2 found that participants with higher metacognitive proclivity were faster at learning foreign norms. This study conceptually replicated the finding in Study 1; this time, however, we used a different cultural metacognition scale, a different domain of behavior, as well as a different way of simulating interpersonal interactions: we gave participants visual images of the people they encounter rather than verbal descriptions.

The experimental manipulation explored how the complexity of detail in the learning task affected learning and the role of metacognition. Not surprisingly, the manipulation dramatically changed the overall speed and the amount of learning. However, the effect of metacognition was the same across conditions, as seen in the nonsignificant condition X metacognition X trial order interaction.

The relationship between higher metacognition and faster learning remained despite the fact that cultural metacognition was measured more than a week prior to the learning task. Importantly, the relationship remained even when controlling for a wide array of alternative individual difference constructs, including other dimensions of cultural intelligence, big-five personality traits, IQ, ethnocentrism, attitudes about intercultural interactions, and current and past intercultural experiences.

Study 3 – Surprise as a Mediator

We propose that metacognitive proclivity fosters learning through processes of error monitoring and error-based updating. Error monitoring involves forming confidence judgments about one’s accuracy, which result in stronger feelings of surprise upon receiving negative feedback and, in particular, at *spurious* negative feedback. People register surprise on a conscious level even when they formed confidence judgments implicitly. Hence, surprise can be measured with self-reports (Charles et al., 2013). We hypothesized that metacognitively inclined learners would feel more surprise at spurious error feedback, and thus surprise responses should mediate the relationship between metacognitive proclivity and speed of learning (H2). We measured surprise and learning in two separate tasks to obtain independent measures of both variables.

Method

Participants. As we wanted to test the role of a mediator, which could only be measured indirectly, we targeted a larger sample size of 200 participants. We posted surveys on Amazon Mechanical Turk (MTurk) seeking 200 US residents, in two different waves. In response, 213 participants completed the study (72 women, 141 men; mean age 37.37 years; 154 European Americans, 21 African Americans, 7 Latin Americans, 2 Native Americans, 14 Asian Americans, 3 Other races, 12 multiracial; 206 US citizens, 6 citizens of other countries, and 1 did not report their citizenship). Of these, 200 participants submitted the HIT on MTurk; 13 participants completed the study but did not submit the HIT on MTurk. All completed responses were unique based on their MTurk ID.

Learning task. Participants first completed the cultural learning task, which was programmed using the Qualtrics© software. The task was similar to the one in Study 2. The trials showed faces of the 80 Asian adults from Study 2; this time, however, we did not include any verbal cues. On each trial, we asked participants, “Will you shake hands or bow when you meet this person?”, showed them the face, and then showed two radio buttons that were labeled “Shake hands” and “Bow.” Immediately after participants made a response, they received feedback, either “Correct!” (in green color front) or “Incorrect!” (in red color font), for one second. The underlying greeting rule was based on the gender of the target: shaking hands with men and bowing to women.

Just as in Study 2, we included noise in the feedback. Shaking hands with men and bowing to women was followed with “correct” feedback in 80% of the trials and with “incorrect” feedback in the remaining 20% of the trials. We stepped up the degree of noise in order to make the task more challenging. We told participants that the person who accumulated the most "correct" responses on this task would receive a bonus of $10.

Surprise task. After participants completed the bow-handshake learning trials, we used the Inquisit© software to run another learning task, which involved a different aspect of greeting. Participants were informed that, in Japan, people use two forms of *you*, the informal *Soji* and the formal *Koji*. The words *Soji* and *Koji* were fictitious, although the Japanese language does have different forms of *you*. Participants' task was to learn when to address Japanese people using the informal *Soji* or the formal *Koji*. Next, we showed participants the pictures of 80 Asian faces, as in the previous task, one at a time. Each face appeared for two seconds. Thereafter, participants saw a prompt: "Press S for Soji, K for Koji," and made a response. Immediately after the response, participants received feedback, "Correct" or "Wrong," for one second. Afterward, they were asked to rate how surprised they were with the feedback on a 7-point scale ranging from *Not at all surprised* to *Extremely surprised*.

The task was designed in such a way that participants received deterministic feedback in the first 20 trials: *Soji* was the correct response when the stimulus person was a woman and *Koji* was the correct response when the stimulus person was a man. This deterministic feedback ensured that nearly all participants would learn the rule in the first 20 trials, and thus our subsequent measure of surprise would not be affected by individual differences in learning. In the remaining 60 trials, we inserted spurious feedback on every fifth trial, telling participants "Wrong" if their response was correct, and "Correct" if their response was wrong. In general, individuals who engage in error monitoring and form confident judgments when their experience points clearly to an answer are likely to feel greater surprise at the spurious error feedback. This individual difference should mediate the effect of the cultural metacognition scale on participants' speed of learning in the initial learning task.

Cultural intelligence. Finally, participants completed all four subscales of the cultural intelligence scale, measured on a 7-point response scale ranging from *Strongly disagree* to *Strongly agree* (Ang et al., 2007; α's > .85). Although the other dimensions did not matter in Study 2, some studies have found that they highly correlate with each other (Klafehn et al., 2013), suggesting the need to control for other dimensions when assessing the effects of cultural metacognition.

Results

Metacognition and Learning. The descriptive statistics and correlations among all the variables are shown in Table 1. Table 2 and 3 present the results of all regression analyses. We tested whether the speed of learning was predicted by the metacognitive dimension, rather than by the other dimensions of cultural intelligence—behavioral, motivational, and knowledge.

*Hierarchical logistic model***.** We first analyzed the data using a hierarchical logistic regression. The dependent measure was accuracy on each trial. The predictor variables were the same as above but instead of trial sets, we used trial order (coded to range from -.5, indicating the first trial, to .5, indicating the last trial). There is no significant effect of trial order, $B = 0.048$, 95% *CI* [-0.063, 0.16], $SE = .057$, *odds ratio* = 1.05, $z = 0.85$, $p = .40$. We found a main effect of metacognition, $B = 0.18$, 95% *CI* [0.041, 0.32], $SE = .071$, *odds ratio* = 1.20, $z = 2.54$, $p = .011$, which indicated that more metacognitively inclined participants were more accurate. Furthermore, as predicted, we found a trial order X metacognition interaction, $B = 0.27$, 95% *CI* [0.14, 0.40], $SE = .068$, *odds ratio* = 1.31, $z = 3.98$, $p < .001$. Thus, more metacognitively oriented participants were faster at learning. There are no significant main effects of the other dimensions, cultural knowledge, $B = -0.011$, 95% *CI* [-0.13, 0.11], $SE = .061$, *odds ratio* = 0.99, $z = 0.19$, $p = .85$; cultural motivation, $B = -0.070$, 95% *CI* [-0.20, 0.060], $SE = .067$, *odds ratio* = 0.93, $z = 1.06$, $p = .29$; cultural behavior, $B = 0.0020$, 95% *CI* [-0.13, 0.13], $SE = .067$, *odds ratio* = 1.00, $z = 0.03$, $p = .98$. Once again, we found the unpredicted negative interaction between trial order and behavioral cultural intelligence, $B = -0.24$, 95% *CI* [-0.37, -0.12], $SE = .064$, *odds ratio* = 0.78, $z = 3.83$, $p < .001$. Both the cultural knowledge X trial order interaction, $B = -0.031$, 95% *CI* [-0.14, 0.083], $SE = .058$, *odds ratio* = 0.97, $z = 0.53$, $p = .60$, and cultural motivation X trial order interaction, $B = 0.025$, 95% *CI* [-0.099, 0.15], $SE = .063$, *odds ratio* = 1.03, $z = 0.40$, $p = .69$, are nonsignificant.

Hierarchical linear model. We next analyzed the data using hierarchical linear analyses. We divided the 80 trials into three sets of 27, 27, and 26 trials, respectively, and computed participants' accuracy in each of the three sets. We analyzed the data using a hierarchical linear regression treating trial sets as nested within participants. The dependent measure was participants' accuracy in each trial set, and the predictor variables were trial set (coded -.5, 0, .5), the four dimensions of cultural intelligence (knowledge, metacognition, motivation, behavior; all centered), and interactions between trial set and the four cultural intelligence dimensions. The main effect of trial set was not significant, $p = .83$. We found a main effect of metacognition, $p = .009$, which indicated that more metacognitively inclined participants were more accurate. Furthermore, as predicted, we found a trial order X metacognition interaction, $p < .001$.

More metacognitively oriented participants were faster at learning. We also found an unpredicted negative interaction between trial order and X behavioral cultural intelligence, $p = .006$. This indicated that greater behavioral plasticity or flexibility was associated with slower learning from feedback. However, we did not see this effect in Study 2, so its replicability remains unclear. Importantly, the metacognitive dimension was the only intercultural competence that helped learning.

To investigate the pattern of the interaction, we ran simple slopes analyses. At one standard deviation below the mean on metacognition, there was a significant negative effect of trial order on accuracy, $B = -0.020$, 95% CI [-0.032, -0.0068], $SE = .0065$, $z = 3.00$, $p = .003$. This finding indicated that participants who were low on cultural metacognition became less accurate across successive trials (see Figure 4). See Supplementary Materials for additional analyses to understand the source of this unexpected negative simple slope effect. At one standard deviation above the mean on metacognition, there was a significant positive effect of trial order on accuracy, $B = 0.018$, 95% CI [0.0037, 0.032], $SE = .0071$, $z = 2.47$, $p = .013$. The finding indicated that as predicted, participants who were high on cultural metacognition became more accurate across successive trials (see Figure 4).

<Insert Table 1>

<Insert Table 2>

<Insert Table 3>

<Insert Figure 4>

Metacognition and Surprise. We next tested whether individuals who were higher on the metacognition dimension showed stronger surprise responses to spurious negative feedback, which was used to measure error monitoring. In the Soji-Koji greeting task, we analyzed data from trials 21 to 80. During these trials, participants received spurious feedback 20% of the time. We indexed participants' degree of error monitoring by computing the difference between their average surprise rating after a correct answer that received (valid) positive feedback and after one that received (spurious) negative feedback. Because we only analyzed data from trials on which participants responded accurately as per the rule that they had learned in the first 20 trials, only 84% of the trials were included in the analysis.

We ran a linear regression with participants as the unit of analysis, rather than a hierarchical regression with trials as the unit of analysis, because we needed a participant level score for surprise differential, indicating how much the participant actively engages in error monitoring through metacognitive confidence in their decision. Participants' surprise differential was the dependent measure, and the four dimensions of cultural intelligence were the predictors (see Table 3). As hypothesized, only the metacognition dimension was a significant predictor, $p = .049$. Participants who scored higher on cultural metacognition were more surprised when they received (spurious) negative feedback compared to when they received (valid) positive feedback. This is consistent with the hypothesis that metacognitively inclined participants felt more confidence in their decisions.

Surprise and Learning. Next, we examined whether participants' surprise response, which was measured by task 2, predicted their cultural learning, measured by task 1. This analysis follows the structure of the *metacognition and learning* hierarchical linear analysis; however, it uses the surprise differential score rather than the cultural metacognition scale, as the predictor. We analyzed participants' responses on the cultural learning task (the bow-handshake task) using a hierarchical linear regression (see Table 2, Model 2). The dependent measure was participants' average accuracy in each trial set, and predictor variables were the trial set (-.5, 0, .5), surprise (centered), and trial set X surprise. We found a main effect of surprise, $p < .001$, which indicated that participants with a higher surprise differential were on average more accurate on the bow-handshake learning task. More importantly, we found a significant trial set X surprise score interaction, $p < .001$. Participants with a higher surprise differential (indicating active error monitoring in task 2) were faster to learn the culturally appropriate pattern of greetings (in task 1).

To investigate the pattern of the interaction, we ran simple slopes analyses. At 1 standard deviation below the mean on surprise difference, we found a significant negative effect of trial order on accuracy, $B = -0.019$, 95% CI [-0.031, -0.0070], $SE = .0062$, $z = 3.08$, $p = .002$. This finding indicated that participants who were low on surprise difference became less accurate across successive trials (see Figure 5). At one standard deviation above the mean on surprise difference, we found a significant positive effect of trial order on accuracy, $B = 0.017$, 95% CI [0.0064, 0.028], $SE = .0055$, $z = 3.12$, $p = .002$. This finding indicated that participants who were high on surprise differential became more accurate across successive trials (see Figure 5).

<Insert Figure 5>

Mediation. Finally, we analyzed participants' performance on the cultural learning task (task 1) using a hierarchical linear regression to test whether surprise differential mediates the relationship between metacognition and culture learning. Predictors were trial set (-.5, 0, .5), the four dimensions of cultural intelligence, interactions between trial order and cultural intelligence factors, surprise, and an interaction between trial order and surprise (see Table 2, Model 3). The metacognition X trial set was significant, $p = .002$, as was the surprise X trial order interaction, $p < .001$. The partial mediation effect was significant using Selig and Preacher's (2008) Monte Carlo simulator with the default settings: the 95% confidence interval for indirect effect with 20,000 repetitions did not include zero, [0.000016, 0.0060].

Discussion

It was found again in Study 3 that metacognitively inclined individuals were faster at learning foreign norms from noisy feedback. Further, they had a higher surprise differential (when they answered accurately, they were less surprised at positive feedback and more surprised at spurious negative feedback), a finding that is consistent with the error-monitoring process of implicit confidence judgments. Higher surprise differential scores were associated with faster learning, and partially mediated the relationship between the cultural metacognition scale and learning. In sum, results of Study 3 were consistent with the proposed metacognitive process of error monitoring through constructing implicit confidence judgments.

Study 3 also found an unpredicted negative simple slope for low metacognition participants, meaning that their performance actually declined across the many trials of the learning task. We did not observe this negative simple slope in Study 2, which used a similar greeting paradigm. However, the feedback was much less noisy: Study 2 featured spurious feedback on only 10% of the trials, whereas Study 3 featured spurious feedback on 25% of the trials, thus making learning from feedback much more difficult. The additional analyses (reported in the Supplementary Materials) explored this negative simple slope by examining an indicator of whether participants “gave up” and stopped trying to be accurate: high metacognition participants varied between the two response options (bow vs. shake hands) to the same extent across early, middle, and late trial sets; however, low metacognition participants began to not vary their answer as they moved from early to middle to late trial sets. This explains why low metacognition participants’ accuracy declined toward chance levels in the later trials.

Study 4 – Metacognitive Prompts and the Question of Timing

Studies 1 to 3 provided correlational evidence for a link between metacognitive proclivity and norm learning, as well as preliminary evidence suggesting that the mediating metacognitive mechanism involves implicit error-monitoring and error-correction processes. Studies 4 and 5 sought to establish the causal link between the metacognitive mechanism and norm learning. As different kinds of prompting elicit different kinds of metacognitive processing, our experiments manipulated the factors of when to prompt and what to prompt (Thillmann et al., 2009). More specifically, the tested whether norm learning is fostered by the prompting designed to enable the implicit process of reactive error-based updating. Namely, it should be fostered by prompts after errors and not necessarily by prompts after accuracy (H3). The prompts were non-directed, just a pause before the next trial. To reduce any demand characteristic for greater learning, the pauses were explained as computer processing delays.

Method

Participants. We targeted the same sample size of 200 participants as in Study

3. We posted a survey on Amazon Mechanical Turk (MTurk) seeking 200 US residents. In response, 192 participants completed the study (92 women, 93 men, 7 of unreported gender; mean age 36.01 years; 143 European Americans, 14 African Americans, 8 Latin Americans, 1 Native Americans, 5 Asian Americans, 14 multiracial, 7 did not report their race; 182 US citizens, 3 citizens of other countries, 7 did not report their citizenship). An additional 8 participants submitted the HIT on MTurk but their data were not recorded by the Inquisit© software, which was used to run the study. All completed responses were unique based on their MTurk ID. Participants were randomly assigned to the no-pause, post-accuracy pause, or the post-error pause conditions.

Procedure. The learning task was similar to the bow-handshake greeting task from Study 3; however, it was programmed using the Inquisit© software (Inquisit 4.01, 2014). We asked participants to imagine that they were visiting a hypothetical Central Asian country called Tavanistan and that they had to learn how to appropriately greet people. Afterward, we presented participants with the same 80 faces, which we used in the past two studies. Initially, each face appeared on the screen for two seconds. Thereafter, participants saw a prompt, “Press B to bow, S to shake hands,” and were asked to respond. The correct greeting was shaking hands with men and bowing to women. As in Study 3, we introduced noisy feedback in 20% of the trials, such that participants received feedback inconsistent with the greeting rule.

After participants made a response (shake hands or bow), we presented them with the feedback “Correct” or “Wrong.” The feedback remained on the screen for one second. In the pause conditions, the feedback was sometimes followed by a blank screen with the message “Loading . . .” for two seconds below the Inquisit logo, which mimicked an Inquisit processing delay message. In the post-accuracy pause condition, this message appeared after participants received “Wrong” feedback, whereas in the post-error pause condition it appeared after participants received “Correct” feedback. In the no-pause condition, this message did not appear after any of the trials. As in the previous studies, participants were told that the participant who accumulated the most number of “Correct” responses would receive a bonus of $10.

Results

Hierarchical logistic regression. We first analyzed the data using a hierarchical logistic regression. To obtain a qualitative understanding of the results, we first ran separate regressions within each condition. The dependent measure was participants’ accuracy on each trial, and the predictor variable was trial order (coded to range from -.5, indicating the first trial, to .5, indicating the last trial). We did not find any effect of trial order on accuracy in the no-pause control condition, B = -0.028, 95% *CI* [-0.21, 0.16], SE = .095, *odds ratio* = 0.97, z = 0.30, p = .77, or in the post-accuracy pause condition, B = 0.092, 95% *CI* [-0.10, 0.29], SE = .10, *odds ratio* = 1.10, z = 0.92, p = .36. However, we found that participants became more accurate with increasing trial order in the post-error pause condition, B = 0.39, 95% *CI* [0.20, 0.57], SE = .094, *odds ratio* = 1.47, z = 4.12, p < .001.

Next, we tested the full model across all three conditions. Specifically, the predictor variables were trial order, a dummy variable indicating the post-error pause condition, a dummy variable indicating the post-accuracy pause condition, and interactions between these dummy variables and trial order. Thus, the no-pause control condition was treated as the dropped baseline. There was no main effect of the post-accuracy pause dummy, *B* = -0.010, 95% *CI* [-0.27, 0.25], *SE* = .13*, odds ratio* = 0.99, *z* = 0.08, *p* = .94, nor of the post-error pause dummy, *B* = 0.14, 95% *CI* [-0.11, 0.38], *SE* = .13, odds *ratio* = 1.15, *z* = 1.07, *p* = .28. The simple effect of trial order, which represented the slope of the learning curve in the no-pause control condition (the dropped baseline), was nonsignificant, *B* = -0.028, 95% *CI* [-0.22, 0.16], *SE* = .096*, odds ratio* = 0.97, *z* = 0.30, *p* = .77. The post-accuracy pause dummy X trial order interaction was nonsignificant, *B* = 0.12, 95% *CI* [-0.15, 0.39], *SE* = .14*, odds ratio* = 1.13, *z* = 0.87, *p* = .38. This indicated that the slope of the learning curve was about the same in the post-accuracy pause condition and the control condition. The post-error pause dummy X trial order interaction was significant, *B* = 0.42, 95% *CI* [0.15, 0.68], *SE* = .13, *odds ratio* = 1.52, *z* = 3.10, *p* = .002. This indicated that the slope of the learning curve was significantly steeper in the post-error pause condition than in the no-pause control condition.

Hierarchical linear regression. We next analyzed the data using hierarchical linear analyses. We divided the 80 trials into three sets of 27, 27, and 26 trials, respectively, and computed participants' accuracy in each of the three sets. First, we examined the pattern of results within each condition. The dependent measure was participants' accuracy in each trial set, and the predictor variables were trial set (coded -.5, 0, .5). We did not find any effect of trial order on accuracy in the no-pause control condition, *B* = -0.0031, 95% *CI* [-0.023, 0.017], *SE* = .010*, z* = 0.30, *p* = .77, or in the post-accuracy pause condition, *B* = 0.0021, 95% *CI* [-0.017, 0.021], *SE* = .0096*, z* = 0.21, *p* = .83. However, we found that participants became more accurate with increasing trial order in the post-error pause condition, *B* = 0.027, 95% *CI* [0.010, 0.043], *SE* = .0085, *z* = 3.12, *p* = .002 (see Figure 6).

Next, we ran the full model by including the same condition dummies as in the hierarchical logistic regression above. There was no main effect of the post-accuracy pause dummy, B = -0.0024, 95% CI [-0.061, 0.056], SE = .030, z = .08, p = .94, nor of the post-error pause dummy, B = 0.030, 95% CI [-0.027, 0.086], SE = .029, z = 1.03, p = .30. The main effect of trial order, which represented the slope of the learning curve in the no-pause control condition (the dropped baseline condition), was nonsignificant, B = -0.0031, 95% CI [-0.023, 0.017], SE = .010, z = 0.30, p = .76. The post-accuracy pause dummy X trial order interaction was nonsignificant, B = 0.0051, 95% CI [-0.022, 0.033], SE = .014, z = 0.37, p = .71. However, the post-error pause dummy X trial order interaction was significant, B = 0.030, 95% CI [0.0037, 0.056], SE = .013, z = 2.24, p = .025. This indicated that the slope of the learning curve was significantly steeper in the post-error pause condition than in the no-pause control condition.

<Insert Figure 6>

Discussion

Study 4 found that pauses after errors, compared to a no-pause control condition, helped participants learn a foreign norm from experiential feedback. Consistent with the mechanism of error-based updating, pauses after errors helped but not pauses after accurate responses. However, it is possible that explicit metacognitive processes such as reasoning about the mistake in one's problem-solving strategy may also be triggered by errors. If so, then we cannot rule out a metacognitive mechanism of explicit reflective reasoning. However, we can incisively test between the two accounts—implicit reactive updating and explicit reflecting reasoning—in our final experiment by manipulating the content of prompts.

Study 5 – Metacognitive Prompts and the Question of Directedness

Whereas directed prompts instruct learners to engage in explicit reflective thinking, non-directed prompts merely enable the kinds of implicit processing that occur spontaneously. While directed prompts have been found effective in the context of structured class lessons, less-directed prompts have been found more effective in independent learning because they do not impede the person's spontaneous learning processes (Ifenthaler, 2012). Our final experiment held constant the timing of prompts after errors and varied whether the content of prompts was non-directed or directed. The non-directed prompts were just pauses disguised as computer processing delays (as in Study 4). The directed prompts instructed participants to reason about how they arrived at their error. We predicted that non-directed prompts would foster learning whereas directed prompts would not (H4). Alternatively, if explicit reflective reasoning is the metacognitive mechanism that helps learning, then the directed prompts should foster learning. Hence the manipulation allows a test between these two possible metacognitive mechanisms.

Method

Participants. Given that we were testing a subtle difference between two different post-error prompts, we increased our target sample size to 600 participants. We posted surveys on Amazon Mechanical Turk (MTurk) seeking 600 US residents, in two different waves. We received 616 completed responses, out of which 4 responses were duplicated based on their MTurk ID. Thus, the final sample size was 612 participants (268 women, 329 men, 15 of unreported gender; mean age 36.47 years; 461 European Americans, 36 African Americans, 21 Latin Americans, 5 Native

Americans, 34 Asian Americans, 1 Middle Eastern American, 9 other races, 32 multiracial, 13 did not report their race; 590 US citizens, 9 citizens of other countries, 13 did not report their citizenship). Of these, 584 submitted the HIT on Murk and 28 completed the study but did not submit the HIT on MTurk. Participants were randomly assigned to the no-prompt control, directed prompt, or non-directed prompt conditions.

Procedure. Participants were asked to imagine that they were visiting a hypothetical Central Asian country called Tavanistan and had to learn how to appropriately greet people. They were then presented with the same 80 faces used in the past two studies. Initially, each face appeared on the screen for two seconds. Thereafter, participants saw a prompt, “Press B to bow, S to shake hands,” and made a response. As in Study 4, the correct greeting for men was to shake hands and that for women was to bow in 80% of the trials.

The no-prompt control condition was identical to that in Study 4. In the other two conditions, two-second pauses occurred after participants received a “Wrong” feedback, as illustrated in Figure 7. The accompanying message varied across the two prompt conditions: “Please wait – Image Loading Process” (non-directed prompt condition) versus “Please think – Analyze the feedback” (directed prompt condition). The non-directed prompt was similar to that used in Study 4, whereas the directed prompt differed in that it instructs the participant to engage in explicit reasoning.

<Insert Figure 7>

Results

We coded participants’ accuracy on each trial based on whether their responses were consistent with the underlying contingency—shaking hands with men and bowing to women.

Hierarchical logistic regression. We first analyzed the data using a hierarchical logistic regression. To obtain a qualitative understanding of the results, we first ran separate regressions within each condition. The dependent measure was participants' accuracy on each trial, and the predictor variable was trial order (coded to range from -.5, indicating the first trial, to .5, indicating the last trial). There was no effect of trial order on accuracy in the no-prompt control condition, *B* = 0.10, 95% *CI* [-0.000049, 0.21], *SE* = .053, *odds ratio* = 1.11, *z* = 1.96, *p* = .050, and in the directed prompt condition, *B* = -0.016, 95% *CI* [-0.13, 0.096], *SE* = .057, *odds ratio* = 0.98, *z* = 0.28, *p* = .78. However, we found that participants became more accurate with increasing trial order in the non-directed prompt condition, *B* = 0.30, 95% *CI* [0.20, 0.40], *SE* = .051, *odds ratio* = 1.35, *z* = 5.82, *p* < .001.

Next, we tested the full model across all three conditions. Specifically, the predictor variables were trial order, a dummy variable that indicated the directed prompt condition, a dummy variable that indicated the non-directed prompt condition, and interactions between these dummy variables and trial order. The no-prompt control condition was treated as the dropped baseline. There was no main effect of the directed prompt dummy, *B* = 0.14, 95% *CI* [-0.011, 0.29], *SE* = .076, *odds ratio* = 1.15, *z* = 1.81, *p* = .070, nor of the non-directed prompt dummy, *B* = 0.015, 95% *CI* [-0.12, 0.15], *SE* = .071, *odds ratio* = 1.01, *z* = .21, *p* = .84. The simple effect of trial order, which represented the slope of the learning curve in the no-prompt control condition, was significant, *B* = 0.10, 95% *CI* [0.00025, 0.21], *SE* = .053, *odds ratio* = 1.11, *z* = 1.96, *p* = .049. The directed prompt dummy X trial order interaction was nonsignificant, *B* = -

0.12, 95% *CI* [-0.27, 0.033], *SE* = .078, *odds ratio* = .89, $z = 1.54$, $p = .12$. This indicates that the slope of the learning curve in the directed prompt condition was not different from that in the control condition. By contrast, the non-directed prompt dummy X trial order interaction was significant, $B = 0.19$, 95% *CI* [0.050, 0.34], *SE* = .073, *odds ratio* = 1.21, $z = 2.63$, $p = .008$. This indicates that the slope of the learning curve was significantly steeper in the non-directed prompt condition than in the control condition.

Hierarchical linear regression. We divided the 80 trials into three sets of 27, 27, and 26 trials, respectively, and computed participants' accuracy in each of the three sets. We analyzed the data using a hierarchical linear regression treating trial sets as nested within participants. First, we examined the pattern of results within each condition. The dependent measure was participants' accuracy in each trial set, and the predictor variables were the same as in the logistic regression except that we used trial set (coded -.5, 0, .5) instead of trial order. We first ran separate models within each condition with trial set as the sole predictor. Participants' accuracy did not significantly increase with trial set in the no-prompt condition, $B = 0.012$, 95% *CI* [-0.012, 0.036], *SE* = .012, $z = 1.01$, $p = .31$, and in the directed prompt condition, $B = -0.00050$, 95% *CI* [-0.022, 0.021], *SE* = .011, $z = 0.05$, $p = .96$. However, participants became more accurate with increasing trial set in the non-directed prompt condition, $B = 0.045$, 95% *CI* [0.024, 0.067], *SE* = .011, $z = 4.12$, $p < .001$ (see Figure 8).

Next, we ran the full model, including the condition dummies. There was no main effect of the non-directed prompt dummy, $B = 0.0033$, 95% *CI* [-0.027, 0.034], *SE* = .016, $z = 0.21$, $p = .83$, nor of the directed prompt dummy, $B = 0.030$, 95% *CI* [-0.0029, 0.064], *SE* = .017, $z = 1.79$, $p = .074$. The simple effect of trial order, which represented the slope of the learning curve in the control no-prompt condition, was also nonsignificant, *B* = 0.012, 95% *CI* [-0.012, 0.036], *SE* = .012, *z* = 1.01, *p* = .31. The directed prompt dummy X trial order interaction was nonsignificant, *B* = -0.013, 95% *CI* [-0.045, 0.019], *SE* = .016, *z* = 0.78, *p* = .44. However, the non-directed prompt dummy X trial order interaction was significant, *B* = 0.033, 95% *CI* [0.00090, 0.065], *SE* = .016, *z* = 2.01, *p* = .044. This finding indicated that the slope of the learning curve was significantly steeper in the non-directed prompt condition than in the no-prompt control condition.

<Insert Figure 8>

Discussion

Whereas Study 4 supported our prediction about when prompts are effective, Study 5 supported our prediction about what content is effective. Consistent with an implicit mechanism of error-based updating, non-directed prompts spurred learning more than directed prompts. When participants received the prompt “Please think – Analyze the feedback,” they failed to learn from feedback. This result indicates that the effects of pauses and prompts are not responses to experimental demand, since the condition with most demand did not produce any learning.

The experimental manipulation used in this study sought to vary implicit vs. explicit metacognition in a face-valid manner. As we were trying to manipulate implicit processes, there was no clear way to run a manipulation check. We believe it is reasonable to assume that participants consciously analyzed their errors to a greater extent when explicitly instructed to do so than when not instructed to do so. However, this is an assumption that has not been verified with data.

Our directed prompt likely encouraged participants to engage in explicit metacognition. However, participants' speed of learning was similar across the directed prompt condition and the no prompt control condition). Nevertheless, it would be a mistake to conclude that explicit metacognition is not helpful in learning any aspect of culture. For instance, learning law-like patterns, such as religious doctrine, would lend itself to explicit rule-based reasoning. Yet, the norms of interpersonal interaction present fuzzier patterns. Whereas one would receive blunt and clear feedback upon violating religious rules, such as wearing shoes into a mosque, the feedback after interpersonal interactions is noisy, which we modeled with spurious feedback on 20% of the trials. This makes it challenging to learn by testing rules explicitly against the evidence of experience. Spurious feedback signifies the occurrence of counterexamples for every rule. However, associative learning can pick up the probabilistic patterns in feedback despite noise (Foerde & Shohamy, 2011). Our results suggest that prompts for explicit reasoning do not help learning from the noisy feedback of interpersonal interactions.

General Discussion

The current research tested a proposal about the role of metacognitive processes in experiential learning of cultural norms. The findings supported hypotheses from the proposal. Individuals higher in (self-reported) proclivity to cultural metacognition were faster at learning to act appropriately in a new culture (Study 1), and this effect held when the self-report is separated in time from the learning task, and after controlling for many potentially confounding individual differences (Study 2). The link between higher cultural metacognition proclivity and faster learning was partly mediated by surprise responses, which are indicators of confidence judgments involved in implicit error monitoring (Study 3).

The two final studies were experiments that tested whether prompts for specific metacognitive processes aided experiential learning of cultural norms. Consistent with the implicit process of reactive post error processing, Study 4 found that post-error pauses, but not post-accuracy pauses, increased participants' speed of learning new cultural norms. Study 5 found that that generic, non-directed prompts ("Please wait -- Image loading process") fostered experiential learning but explicit, directed prompts ("Please think – Analyze the feedback") did not. This is consistent with the proposal that experiential learning occurs through an implicit updating process and not through explicit reasoning about the feedback, or in other words, through implicit metacognition rather than explicit metacognition.

Implications for theory

Meso-level analysis. The influence of macro-level patterns on individual behavior is a key issue in research on culture (Kitayama et al., 1997; Kitayama, Mesquita, & Karasawa, 2006), as well as in research on influences of social class (Kohn & Schooler, 1969), communities (Riger & Lavrakas, 1981), and political parties (Huddy, 2001). Theorists have proposed that interpersonal interactions are the nexus through which macro patterns shape individual mentalities and vice versa (Coleman, 1990; House, 1981; James, 1990). The interpersonal level is what most clearly distinguishes social psychology from sociology one side and from experimental psychology on the other; nonetheless, most psychological research on culture either looks at associations between country-level variables (Schwartz, 1986) or links between different constructs at the individual level (Oyserman & Lee, 2008). Only a few research programs have identified interpersonal situations that instill and maintain the group-level tendency in an individual's behavior, such as Kitayama and colleagues' work on situational affordances (Kitayama et al., 1997; Morling et al., 2002) and Weisbuch and colleagues' work on nonverbal cues (Weisbuch, Lamer, Treinen, & Pauker, 2017). Our research documents another: experiencing interactions and evaluative feedback reinforces norm-compliant behavior. We have modeled the experience of expatriates learning norms of new country, but this interpersonal process likely also operates to shape the behavior of grunts at boot camp, first-generation students' behavior at college, and new employees at strong-culture corporations, such Disney, Facebook, or Bridgewater Associates.

Second-culture learning. The current findings inform theories of how people learn new cultures. Two influential models of "cultural intelligence" agree in that they both include metacognition dimension (Ang et al., 2007; Thomas et al., 2008). Cultural metacognition was recently termed a "new frontier in cross-cultural competence research" (Chiu, Lonner, Matsumoto, & Ward, 2013, p. 846). Theorists have proposed that higher proclivity toward metacognition makes an expatriate more likely to learn more from everyday experiences (Ng et al., 2009). However, past research has not found a way to measure the learning of cultural norms from experience. Through our laboratory simulation of experiencing interpersonal situations in a new culture, we found that the speed of learning was associated with individual differences in metacognitive proclivity, supporting the proposal of cultural intelligence theorists.

That said, the current findings point to different intervening metacognitive processes than those described in the cultural intelligence literature. Cognitive and developmental psychologists have always posited *both* implicit and explicit processes of monitoring and control (Flavell, 1979). However, cultural intelligence theory places the emphasis on explicit processes. While Ang et al.'s (2007) have remained agnostic on this point, Thomas et al. (2008) assert that cultural metacognition processes are "deliberate, planful, intentional, goal-directed, and future-oriented." (p. 131). They even try to re-define the term in ways that Flavell would not recognize: "The term metacognitive [should] be reserved for 'conscious' and 'deliberate' thoughts that have as their object other thoughts" (p. 132). This narrowed definition is quite out of step with cognitive psychology, which studies implicit error monitoring and control processes using neuroscience methods, such as evoked potentials (Fleming, Dolan, & Frith, 2012; Frith, 2012; McCurdy, Maniscalco, Metcalfe, Liu, de Lange, & Lau, 2013).

Although our process evidence is less direct than neural measures, it clearly points to the widely studied implicit processes of error monitoring and reactive updating.[1] The measure of surprise differential indicated that the error monitoring process of confidence judgments (Maniscalco & Lau, 2012) partly mediated the effect of metacognitive proclivity on experiential learning. The signature of implicit reactive updating was seen in our experimental findings that breaks in the learning task that allow thinking helped more when they came after errors than after accuracy (Study 4) and helped more when they were non-directed as opposed to instructing explicit analysis of one's error (Study 5). If the relevant metacognitive processes involved explicit reasoning, then learning would have been helped by the instructions to explicitly think about the feedback.

That is not to say, however, that explicit metacognitive processing is never helpful to intercultural effectiveness – just not in the stage of learning norms from experiential feedback. Effective *use* of cultural knowledge, once acquired, often requires the use of explicit thinking to override automatic associations. For instance, for an expatriate to bow at the sight of his Japanese superordinate is appropriate in the Tokyo office; however, bowing to the same person when she visits the New York office may not be appropriate. Conscious metacognitive processes—conscious reasoning and planning—may be important in figuring out when to inhibit the habitual responses of adhering to a norm.

First-culture learning. The mechanism we identify—experiential learning from interactions—is also likely important to how people acquire and maintain their native culture competency. While classical work on this topic focused on vertical transmission from the older generation through socialization procedures, recent work emphasizes horizontal transmission between peers. It has emphasized that people can pick up patterns of behavior through imitating others (Tomasello, 2009). Specifically, there is evidence for use of social-learning heuristics, such as imitating the majority, or imitating those with prestige (Richerson & Boyd, 2005; Chudek & Henrich, 2011). Our research points to a different kind of horizontal transmission, one that has been overlooked in recent theorizing. Henrich (2015) argues that the skills in a culture such as toolmaking procedures (which empowered our ancestors' ascent) were acquired through "social learning" involving imitation rather than "asocial learning" involving first-hand trial-and-error interaction with the environment.

However, our findings about the transmission of interpersonal norms suggests the need for a third category. Learning from interpersonal feedback is learning from first-hand interaction with the environment, but it can hardly be called "asocial learning." It's not learning from imitation, but it is inherently social as it involves information from other people. Learning from social reinforcement may substitute for imitative learning in cases of where it is limited by deficits such as autism. Reinforcement training has proved useful for instilling social skills such as appropriate greetings and conversational turn-taking (e.g., Hwang & Hughes, 2000; Pierce & Schreibman, 1995).

Practical Implications

Training. A longstanding area of applied research and practice has the goal of supporting expatriates' learning of the new culture. While such workshops traditionally occurred pre-departure workshops, such as lectures about individualism versus collectivism, many organizations have reallocated training budgets from pre-departure lectures to post-arrival coaching. Our results on metacognitive prompts suggest that, in order to promote learning of interpersonal norms, coaches should not interrupt people after they've had a success. Instead, coaches should create non-directed pauses after errors to make participants feel surprised and engage in corrective updating. Highly directed prompts are likely less effective for learning interpersonal norms because they engage explicit processing, but they may useful within the context of classroom instruction about well-structured domains such as currency systems, legal codes, etc.

A comparative study of training methods found that learning to avoid judgment biases, such as anchoring, is helped more using game-like experiential feedback than a lecture (Morewedge et al., 2015). The US Army has developed video games in order to teach weaponry skills, where students receive positive feedback for correct moves. But the Army's cultural training still retains traces of the lecture format. They issue wallet-ready "smart cards" listing verbal rules of behavior (*Dos and Don'ts*, such as "Don't point the sole of your foot at an Arab.") In a cultural training game, players "explore an Iraqi village, hear the sounds, speak to locals, and make gestures" (Lane, 2007, p 3). Feedback comes organically from evaluative responses of virtual interactants (what they say, tone of voice, and facial expressions). At the same time, a text window provides verbal hints and directed prompts for reflection. Our Study 6 suggests that for gaining competency in interpersonal norms, game-like feedback is more useful than verbal instruction.

Selection. The current research has implications for practices of selecting people for overseas or intercultural roles. While many organizations such as the State Department rely heavily on measures of explicit aptitude such as IQ, our results showed that IQ is not associated with faster experiential learning but metacognitive proclivity was. Of course, these organizations are interested in many kinds of learning so it may be wise to keep IQ in the mix but they should also add measures of metacognitive proclivity. There are limitations of using a self-report test for assessment, but innovations in question delivery and response format have brought largely "fake-proof" self-report assessments, which are being applied to measures of cultural intelligence (Morris, Savani, & Roberts, 2014).

Issues for Future Research

Assessing metacognitive proclivity. Our studies followed the cultural intelligence literature in assessing metacognitive proclivity with questions specific to the intercultural domain. This has the advantage of allowing us to compare effects of metacognitive proclivity in this domain to other strengths in this domain captured by other dimensions of the cultural intelligence scales such as knowledgeability, motivation, and behavioral flexibility. These inventories ask people whether they are aware of their assumptions and check their assumption (e.g. “I am conscious of the cultural knowledge I apply to cross-cultural interactions.”) This leaves open the question of to what extent metacognitive proclivity is domain-general as opposed to domain-specific. Disorders such as autism and dyslexia suggest that people can be selectively low in some domains. Recent work has found that metacognition in perceptual and memory tasks are associated with separate brain structures and can be selectively impaired with different lesions (McCurdy et al., 2013; Fleming, Ryu, Golfinos, & Blackmon, 2014).

Measuring domain-general proclivity simply by broadening the language of the items in an inventory (e.g. “I am conscious of the knowledge I apply in many domains of my life”) seems unlikely to work, as generality makes the question difficult to reliably answer. Another approach might be to adapt methods used to measure metacognitive sensitivity, the calibration of confidence with accuracy (Maniscalco & Lau, 2012). However, metacognitive sensitivity may diverge from metacognitive proclivity (see Fleming & Lau, 2014, for an analysis of awareness in relation to different sensitivity measures).

Another question concerns self-report scales. It may seem odd at first glance that we measure metacognitive proclivity using self-report scales even though we propose that the metacognitive mechanism operates largely outside of awareness. A self-report measure can assess proclivity toward an implicit process so long as products of that process are introspectively available to the person. For example, error monitoring through forming confidence judgments occurs implicitly, but it gives rise to surprise responses that people are conscious of. So, someone who habitually forms confidence judgments and experiences surprise responses would likely see themselves as concerned with the accuracy of their cultural assumptions even though they lack introspective access to their process of estimating confidence.

Findings from research on other kinds of competencies supports the usefulness of self-report assessments. Self-reported emotional intelligence measures incrementally predict job performance over and above ability-based tests of emotional intelligence (Joseph & Newman 2010). Similarly, self-reported cultural intelligence predicts task performance in multicultural teams over and above an ability-based measure derived from a situational judgment test (Rockstuhl, Ang, Ng, Lievens & Van Dyne, 2013). Besides, this is consistent with how implicit theories are measured-- via agreements with explicit statements (Dweck, 2006). People are aware of holding such beliefs but not aware of the ways that the beliefs affect their judgments

Underlying learning processes. The current studies cannot ascertain whether the whole learning process occurs unconsciously and procedurally. In some domains, people can pick up response patterns to stimuli without awareness of the feedback or of the patterns that they have learned (Pessiglione, Petrovic, Daunizeau, Palminteri, Dolan, & Frith, 2008; Sternberg & McClelland, 2012). The current studies found evidence for implicit metacognitive processes; however, these could operate in conjunction with some explicit conscious processing. Future research can break down the learning task into components to assess the sub-components that are learned implicitly versus explicitly.

Negative effects of high metacognition. Finally, although the current findings point to the upsides of metacognition, few psychological predispositions represent unmixed blessings. For example, at the extreme of proclivity toward metacognition, individuals with obsessive-compulsive disorder torture themselves by constantly worrying and checking whether they have done tasks correctly. Excessive concern about making interpersonal errors can cause crippling social anxieties, which then create the interpersonal problems which were feared in the first place. Furthermore, just as chronic individual predispositions toward metacognition have downsides, so do situations that prompt metacognition (Belenky & Nokes, 2009). Consistent with the principle that more difficult learning tasks induce long-term learning (Bjork, 1994), it has been found that prompts, even in primate studies, enhance short-term but not long-term performance (Kornell & Terrace, 2007).

Conclusion

The current studies provide the first evidence for the role of metacognition in cultural learning, specifically in learning interpersonal norms from experiential feedback. Our findings suggest that implicit metacognitive mechanisms are at work, rather than explicit mechanisms. These implicit metacognitive mechanisms offer fresh theoretical insights about acculturation, enculturation, cultural evolution, and the interpersonal level of analysis as nexus of macro structures and micro processes. Moreover, they suggest practical insights relevant to selecting and training people for interactions with people of other cultures.

References

Abbe, A., Gulick, L. M., & Herman, J. L. (2008). *Cross-cultural competence in Army leaders: A conceptual and empirical foundation* (2008-01). Arlington, VA: Army Research Institute.

Adair, W. L., Hideg, I., & Spence, J. R. (2013). The culturally intelligent team: The impact of team cultural intelligence and cultural heterogeneity on team shared values. *Journal of Cross-Cultural Psychology*, *44*, 941-962.

Ambrose, S. H. (2010). Coevolution of composite-tool technology, constructive memory, and language: implications for the evolution of modern human behavior. *Current Anthropology*, *51*, S135-S147.

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural intelligence: its measurement and effects on cultural adaptation and task performance. *Management and Organization Review, 3,* 335-371.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, 74, 29-58.

Belenky, D. M., & Nokes, T. J. (2009). Examining the role of manipulatives and metacognition on engagement, learning, and transfer. *The Journal of Problem Solving*, *2*, 102-129.

Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*, 564-577.

Bhawuk, D. P. (1998). The role of culture theory in cross-cultural training a multimethod study of culture-specific, culture-general, and culture theory-based assimilators. *Journal of Cross-Cultural Psychology*, *29*, 630-655.

Bhawuk, D. P., Sakuda, K. H., & Munusamy, V. P. (2008). Intercultural competence development and triple-loop cultural learning. *Handbook of cultural intelligence: Theory, method, and applications*, 342-355.

Black, J. S., & Stephens, G. K. (1989). The influence of the spouse on American expatriate adjustment and intent to stay in Pacific Rim overseas assignments. *Journal of Management*, *15*, 529-544.

Bilker, W. B., Hansen, J. A., Brensinger, C. M., Richard, J., Gur, R. E., & Gur, R. C. (2012). Development of abbreviated nine-item forms of the Raven's Standard Progressive Matrices Test. *Assessment*, *19*, 354-369.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185-206). Cambridge, MA: MIT Press.

Brevers, D., Cleeremans, A., Bechara, A., Greisen, M., Kornreich, C., Verbanck, P., & Noël, X. (2014). Impaired metacognitive capacities in individuals with problem gambling. *Journal of Gambling Studies*, *30*, 141-152.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1491-1494.

Caligiuri, P. M. (2000). The big five personality characteristics as predictors of expatriate's desire to terminate the assignment and supervisor-rated performance. *Personnel Psychology*, *53*, 67-88.

Charles, L., Van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *Neuroimage, 73,* 80-94.

Chiu, C. Y., Lonner, W. J., Matsumoto, D., & Ward, C. (2013). Cross-cultural competence theory, research, and application. *Journal of Cross-Cultural Psychology*, *44*, 843-848.

Chua, R. Y., Morris, M. W., & Mor, S. (2012). Collaborating across cultures: Cultural metacognition and affect-based trust in creative collaboration. *Organizational Behavior and Human Decision Processes*, *118*, 116-131.

Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences, 15,* 218-226.

Coleman, J. S. (1990). *Foundations of social theory.* Cambridge, MA: Harvard University Press.

Crook, A. E., & Beier, M. E. (2010). When training with a partner is inferior to training alone: The importance of dyad type and interaction quality. *Journal of Experimental Psychology: Applied*, *16*, 335-348.

Crotty, S. K., & Brett, J. M. (2012). Fusing creativity: Cultural metacognition and teamwork in multicultural teams. *Negotiation and Conflict Management Research*, *5*, 210-234.

Davis, E. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences, 12*, 91-142.

Davis, E. A., & Linn, M. C. (2000). Scaffolding students' knowledge integration: prompts for reflection in KIE. *International Journal of Science Education, 22*, 819–837.

Derex, M., Beugin, M. P., Godelle, B., & Raymond, M. (2013). Experimental evidence for the influence of group size on cultural complexity. *Nature, 503,* 389-391.

Dewey J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process.* Boston, MA: Heath.

Dweck, C. (2006). *Mindset: The new psychology of success.* New York: Random House.

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*, 267-285.

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*, 88-92.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34,* 906-911.

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*, 443. doi: 10.3389/fnhum.2014.00443

Fleming, S. M., Dolan, R. J., & Frith, C. D. (2012). Metacognition: Computation, biology and function. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 1280-1286.

Fleming, S. M., Ryu, J., Golfinos, J. G., & Blackmon, K. E. (2014). Domain-specific impairment in metacognitive accuracy following anterior prefrontal lesions. *Brain, 137*, 2811-2822.

Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J., & Rees, G. (2010). Relating introspective accuracy to individual differences in brain structure. *Science, 329,* 1541-1543.

Foerde, K., & Shohamy, D. (2011). Feedback timing modulates brain systems for learning in humans. *The Journal of Neuroscience, 31,* 13157-13167.

Ford, J. K., Smith, E. M., Weissbein, D. A., Gully, S. M., & Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology*, *83*, 218-233.

Frith, C. D. (2012). The role of metacognition in human social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367,* 2213-2223.

Garrett, A. J., Mazzocco, M. M., & Baker, L. (2006). Development of the metacognitive skills of prediction and evaluation in children with or without math disability. *Learning Disabilities Research & Practice*, *21*, 77-88.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385-390.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*, 504-528.

Hall, E. T., (1966). *The hidden dimension*. New York, NY: Doubleday.

Hall, E. T. (1983) *The dance of life: the other dimension of time*. New York, NY: Doubleday.

Harris, J. G. (1973). A science of the South Pacific: Analysis of the character structure of the Peace Corps volunteer. *American Psychologist*, *28*, 232-247.

Heine, S. J., & Lehman, D. R. (2004). Move the body, change the self: Acculturative effects on the self-concept. In M. Schaller & C. Crandall (Eds.), *Psychological Foundations of Culture*. Mahwah, NJ: Erlbaum.

Henrich J. (2015). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton, NJ: Princeton Univ. Press.

Hester, R., Madeley, J., Murphy, K., & Mattingley, J. B. (2009). Learning from errors: Error-related neural activity predicts improvements in future inhibitory control performance. *The Journal of Neuroscience*, *29*, 7158-7165.

Hirsh, J. B., & Inzlicht, M. (2010). Error-related negativity predicts academic performance. *Psychophysiology, 47*, 192-196.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679-709.

Holroyd, C. B., Dien, J., & Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neuroscience Letters*, *242*, 65-68.

House, J. S. (1981). Social structure and personality. In M. Rosenberg & R. H. Turner (Eds.), *Social psychology: Sociological perspectives* (pp. 525-561). New York, NY: Basic Books.

Howell, W. S. (1982). *The empathic communicator*. Belmont, CA: Wadsworth Publishing Company.

Huddy, L. (2001). From social to political identity: A critical examination of social identity theory. *Political Psychology*, *22*, 127-156.

Huff, K. C., Song, P., & Gresch, E. B. (2014). Cultural intelligence, personality, and cross-cultural adjustment: A study of expatriates in Japan. *International Journal of Intercultural Relations, 38*, 151-157.

Hwang, B., & Hughes, C. (2000). The effects of social interactive training on early social communicative skills of children with autism. *Journal of Autism and Developmental Disorders*, *30*, 331-343.

Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated learning in problem-solving scenarios. *Educational Technology & Society*, *15*, 38-52.

Inquisit 4.01 (2014). Computer software. Seattle, WA: Millisecond Software.

James, C. (1990). *Foundations of social theory*. Cambridge, MA: Belknap.

Jarvis, B. G. (2010). DirectRT (Version 2006) [Computer Software]. New York, NY: Empirisoft Corporation.

Job, J.M. & Klassen, R.M. (2012). Predicting performance on academic and non-academic tasks: A comparison of adolescents with and without learning disabilities. *Contemporary Educational Psychology, 37*, 162–169.

Joseph, D. L., & Newman, D. A. (2010). Emotional intelligence: an integrative meta-analysis and cascading model. *Journal of Applied Psychology, 95,* 54-78.

Kassing, J. W. (1997). Articulating, antagonizing, and displacing: A model of employee dissent. *Communication Studies*, *48*, 311-332.

Kealey, D. J., & Protheroe, D. R. (1996). Effectiveness of cross-cultural training for expatriates: An assessment of the literature on the issue. *International Journalof Intercultural Relations*, *20*,141-165.

Kelemen, W. L., Frost, P. J., & Weaver, C. A. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory & Cognition, 28*, 92-107.

Kinnunen, R., & Vauras, M. (1995). Comprehension monitoring and the level of comprehension in high-and low-achieving primary school children's reading. *Learning and Instruction, 5*, 143-165.

Kitayama, S., Duffy, S., Kawamura, T., & Larsen, J. T. (2003). Perceiving an object and its context in different cultures: A cultural look at new look. *Psychological Science, 14*, 201-206.

Kitayama, S., Mesquita, B., & Karasawa, M. (2006). Cultural affordances and emotional experience: socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology*, *91*, 890-903.

Kitayama, S., Markus, H. R., Matsumoto, H., & Norasakkunkit, V. (1997). Individual and collective processes in the construction of the self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology, 72,* 1245-1267.

Klafehn, J., Li, C., & Chiu, C. Y. (2013). To know or not to know, is that the question? Exploring the role and assessment of metacognition in cross-cultural contexts. *Journal of Cross-Cultural Psychology, 44,* 963-991.

Kohn, M. L., & Schooler, C. (1969). Class, occupation, and orientation. *American Sociological Review*, *34*, 659-678.

Kornell, N., & Terrace, H.S. (2007). The generation effect in monkeys. *Psychological Science, 18*, 682–685.

Krigolson, O. E., & Holroyd, C. B. (2006). Evidence for hierarchical error processing in the human brain. *Neuroscience*, *137*, 13-17.

Lane, H. C. (2007). Metacognition and the development of intercultural competence. In B. P. Woolf, E. A¨imeur, R. Nkambo & S. Lajoie (Eds.), *Workshop on Metacognition and SRL. Artificial Intelligence in Education (AIED)* (pp. 23–32). Marina del Rey, CA.

Legault, L., & Inzlicht, M. (2013). Self-determination, self-regulation, and the brain: autonomy improves performance by enhancing neuroaffective responsiveness to self-regulation failure. *Journal of Personality and Social Psychology*, *105*, 123-138.

Leung, A. K. Y., & Chiu, C. Y. (2010). Multicultural experience, idea receptiveness, and creativity. *Journal of Cross-Cultural Psychology*, *41*, 723-741.

Leung, K., Ang, S., & Tan, M. L. (2014). Intercultural competence. *Annual Review of Organizational Psychology and Organizational Behavior*, *1*, 489-519.

Lin, X., & Lehman, J. D. (1999). Supporting learning of variable control in a computer-based biology environment: Effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching, 36*, 837-858.

Lu, J. G., Quoidbach, J., Gino, F., Chakroff, A., Maddux, W. W., & Galinsky, A. D. (2017). The dark side of going abroad: How broad foreign experiences increase immoral behavior. *Journal of Personality and Social Psychology, 112*, 1-16.

Maddux, W. W., & Galinsky, A. D. (2009). Cultural borders and mental barriers: the relationship between living abroad and creativity. *Journal of Personality and Social Psychology*, *96*, 1047-1061.

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*, 422-430.

Mason, A. & Mason, M. (2005). Understanding College Students with Learning Disabilities. *Pediatric Clinics of North America, 52*, 61–70.

McCurdy, L. Y., Maniscalco, B., Metcalfe, J., Liu, K. Y., de Lange, F. P., & Lau, H. (2013). Anatomical coupling between distinct metacognitive systems for memory and visual perception. *Journal of Neuroscience, 33,* 1897–1906.

Molenberghs, P., Trautwein, F. M., Böckler, A., Singer, T., & Kanske, P. (2016). Neural correlates of metacognitive ability and of feeling confident: a large-scale fMRI study. *Social cognitive and affective neuroscience, 11*, 1942-1951.

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C., Korris, J., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. *Policy Insights from the Behavioral and Brain Sciences, 2,* 129-140.

Morling, B., Kitayama, S., & Miyamoto, Y. (2002). Cultural practices emphasize influence in the United States and adjustment in Japan. *Personality and Social Psychology Bulletin*, *28*, 311-323.

Morris, M. W., Savani, K., & Roberts, R. D. (2014). Intercultural training and assessment implications for organizational and public policies. *Policy Insights from the Behavioral and Brain Sciences*, *1*, 63-71.

Morris, M. W., Savani, K., Mor, S., & Cho, J. (2014). When in Rome: Intercultural learning and implications for training. *Research in Organizational Behavior*, *34*, 189-215.

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.

Neuliep, J. W., & McCroskey, J. C. (1997). The development of a US and generalized ethnocentrism scale. *Communication Research Reports*, *14*, 385-398.

Ng, K. Y., Van Dyne, L., & Ang, S. (2009). From experience to experiential learning: Cultural intelligence as a learning capability for global leader development. *Academy of Management Learning & Education*, *8*, 511-526.

Oyserman, D., & Lee, S. W. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, *134*, 311-342.

Pessiglione, M., Petrovic, P., Daunizeau, J., Palminteri, S., Dolan, R. J., & Frith, C. D. (2008). Subliminal instrumental conditioning demonstrated in the human brain. *Neuron*, *59*, 561-567.

Peters, E. E., & Kitsantas, A. (2010). Self-regulation of student epistemic thinking in science: The role of metacognitive prompts. *Educational Psychology*, *30*, 27-52.

Pierce, K., & Schreibman, L. (1995). Increasing complex social behaviors in children with autism: Effects of Peer-implemented pivotal response training. *Journal of Applied Behavior Analysis*, *28*, 285-295.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33-40.

Portes, A. & Rumbaut, R. G. (2001). *Legacies: The story of the immigrant second generation*. Berkeley: Univ. Calif. Press.

Rabbitt, P. M. (1966). Errors and error correction in choice-response tasks. *Journal of Experimental Psychology*, *71*, 264-272.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Thousand Oaks, CA: Sage.

Reder, L. M., & Schunn, C. D. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In L. Reder (Ed.), *Implicit memory and metacognition* (pp. 45-78). Hillsdale, NJ: Erlbaum.

Richerson, P. J., & Boyd, R. (2005). *Not by genes alone*. Chicago: University of Chicago Press.

Riger, S., & Lavrakas, P. J. (1981). Community ties: Patterns of attachment and social interaction in urban neighborhoods. *American Journal of Community Psychology*, *9*, 55-66.

Rockstuhl, T., & Ng, K. Y. (2008). The effects of cultural intelligence on interpersonal trust in multicultural teams. *Handbook of Cultural Intelligence: Theory Measurment and Applications*, 206-220.

Rockstuhl, T., Ang, S., Ng, K.-Y., Lievens, F., Van Dyne, L. (2013). The incremental value of assessing situational perspective taking in SJTs. Paper presented at the 28th Annual Conference of the Society for Industrial and Organizational Psychology, Houston, TX.

Sam, D. L., & Berry, J. W. (2010). Acculturation when individuals and groups of different cultural backgrounds meet. *Perspectives on Psychological Science*, *5*, 472-481.

Savani, K., Morris, M. W., Naidu, N. V. R., Kumar. S., & Berlia, N. (2011). Cultural conditioning: Understanding interpersonal accommodation in India and the United States in terms of the modal characteristics of interpersonal influence situations. *Journal of Personality and Social Psychology, 100,* 84-102.

Savani, K., Phua, D. Y., Morris, M. W., & Hong, Y. Y. The relationship between genotype, intercultural learning ability, and adaptation in a foreign culture. [Unpublished raw data, Nanyang Technological University.]

Schmidt, A. M., & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology*, *56*, 405-429.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*, 460-475.

Schraw G. (1998). Promoting general metacognition awareness. *Instructional Science, 26*, 113–125.

Schwartz, B. (1986). *The battle for human nature: Science, morality, and modern life.* New York, NY: Norton.

Shu, F., McAbee, S. T., & Ayman, R. (2017). The HEXACO personality traits, cultural intelligence, and international student adjustment. *Personality and Individual Differences*, *106*, 21-25.

Selig, J. P., & Preacher, K. J. (2008). Monte Carlo method for assessing mediation: An interactive tool for creating confidence intervals for indirect effects [Computer software]. Available from http://quantpsy.org/.

Spencer-Rodgers, J., & McGovern, T. (2002). Attitudes toward the culturally different: The role of intercultural communication barriers, affective responses, consensual stereotypes, and perceived threat. *International Journal of Intercultural Relations*, *26*, 609-631.

Sterelny, K. (2012). *The evolved apprentice: how evolution made humans unique*. Cambridge, MA: MIT Press.

Sternberg, D. A., & McClelland, J. L. (2012). Two mechanisms of human contingency learning. *Psychological Science, 23,* 59-68.

Tanner, K. D. (2012). Promoting student metacognition. *CBE-Life Sciences Education*, *11*, 113-120.

The World Bank. (2015). International migration at all-time high [Press release]. Retrieved from http://www.worldbank.org/en/news/press-release/2015/12/18/international-migrants-and-remittances-continue-to-grow-as-people-search-for-better-opportunities-new-report-finds

Thillmann, H., Künsting, J., Wirth, J., & Leutner, D. (2009). Is it merely a question of "What" to prompt or also "When" to prompt? *Zeitschrift für Pädagogische Psychologie, 23*,105-115.

Thomas, W. I. & Znaniecki, F. (1918). *The Polish Peasant in Europe and America.* New York, NY: Knopf.

Thomas, D.C., Elron, E., Stahl, G., Ekelund, B.Z., Ravlin, E.C., Cerdin, J., Poelmans, S., Brislin, R., Pekerti, A., Aycan, Z., Maznevski, M., Au, K., Lazarova, M.B. (2008). Cultural Intelligence: Domain and assessment. *International Journal of Cross Cultural Management, 8,* 123-143.

Tomasello, M. (2009). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tung, R. L. (1993). Managing cross-national and intra-national diversity. *Human Resource Management, 32*, 461-477.

Ward, C., & Kennedy, A. (1999). The measurement of sociocultural adaptation. *International journal of intercultural relations*, *23*, 659-677.

Weisbuch, M., & Ambady, N. (2009). Unspoken cultural influence: Exposure to and influence of nonverbal bias. *Journal of Personality and Social Psychology*, *96*, 1104-1119.

Weisbuch, M., Pauker, K., & Ambady, N. (2010). The subtle transmission of race bias via televised nonverbal behavior. *Science, 326,* 1711-1714.

Weisbuch, M., Lamer, S. A., Treinen, E., & Pauker, K. (2017). Cultural snapshots: Theory and method. *Social and Personality Psychology Compass*, *11*, e12334.

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: Common neuronal architecture for the processing of errors and novelty. *The Journal of Neuroscience*, *32*, 7528-7537.

Wirth, J. (2009). Prompting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*, 91-94.

Wong, J. J., & Cho, S. (2007a). A brain-inspired model for recognizing human emotional states from facial expression. In L. Perlovsky & R. Kozma (Eds.), *Neurodynamics of cognition and consciousness*. Berlin: Springer-Verlag.

Wong, J. J., & Cho, S. Y. (2007b). A local experts organization model with application to face emotion recognition. *Expert Systems with Applications*.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367,* 1310-1321.

Author Note

Michael W. Morris and Katrina Fincher, Management Division, Graduate School of Business, Columbia University. Krishna Savani, Department of Strategy, Management, and Organisation, Cultural Science Institute, and Institute on Asian Consumer Insight, Nanyang Business School, Nanyang Technological University.

This research was supported by US Army Research Institute grant W911NF-13-1-0113 awarded to Michael W. Morris, a Nanyang Assistant Professorship grant awarded to Krishna Savani, and Singapore Ministry of Education Social Science Research Thematic Grant MOE2016-SSRTG-006 awarded to Krishna Savani. We thank Hakwan Lau for helpful feedback, and Sooyun Baik, Amie Blocker, Yu Ding, Elizabeth Goodman, Ramya Ramesh, and Li Shi Tan for invaluable research assistance.

Address correspondence to Michael W. Morris, Columbia Business School, 3022 Broadway Uris Hall, New York, NY 10027, or to Krishna Savani, Nanyang Business School, 50 Nanyang Ave S3-01C-81, Singapore 639798. E-mail: mwm82@columbia.edu or ksavani@ntu.edu.sg.

Footnotes

[1] It may seem paradoxical that we measured individual differences in metacognition with an explicit self-report scale while claiming that the processes through which they aid learning are implicit. This is not a contradiction, however. The fact that people know their beliefs and habits does not mean that people are aware of how such beliefs or habits affect their thoughts and behaviors. Hence, research on implicit theories measures implicit theories with explicit self-report scales (Dweck, Chiu, & Hong, 1995).

Figure Captions

Figure 1. Relationships tested in the current studies.

Figure 2. Illustration of the task used in Study 1

Figure 3. Simple slopes depicting participants' mean accuracy by trial set, at ±1 SD on cultural metacognition (Study 1).

Figure 4 Simple slopes depicting participants' mean accuracy by trial set, at ±1 SD on cultural metacognition (Study 3).

Figure 5. Simple slopes depicting participants' mean accuracy by trial set, at ±1 SD on surprise difference score (Study 3).

Figure 6. Simple slopes depicting participants' mean accuracy by trial set and experimental condition (Study 4).

Figure 7. Diagram illustrating the task used in the control, explicit metacognition instructions present, and instructions absent conditions (Study 5).

Figure 8. Simple slopes depicting participants' mean accuracy by trial set and experimental condition (Study 5).

Figure 1.

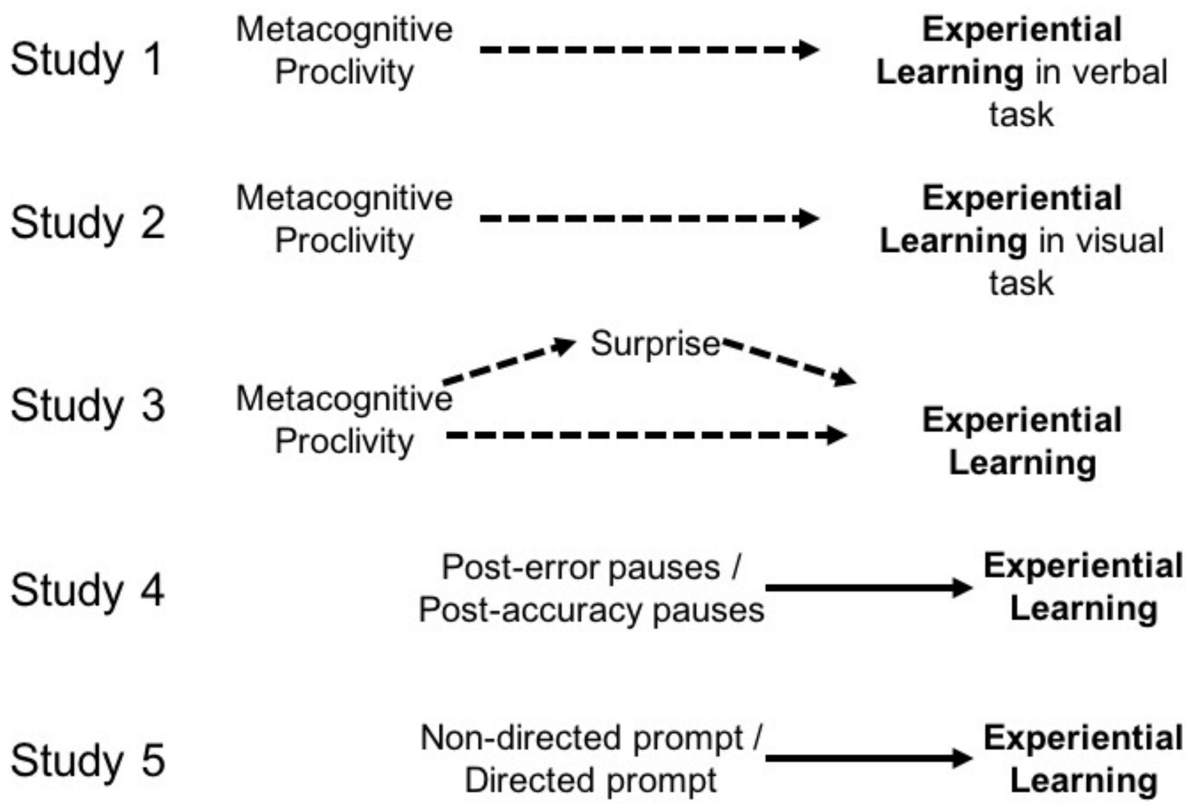
Study 1
Metacognitive Proclivity
Experiential Learning in verbal task
Study 2
Metacognitive Proclivity
Experiential Learning in visual task
Study 3
Surprise
Metacognitive Proclivity
Experiential Learning
Study 4
Post-error pauses / Post-accuracy pauses
Experiential Learning
Study 5
Non-directed prompt / Directed prompt
Experiential Learning

Figure 2 (Study 1).

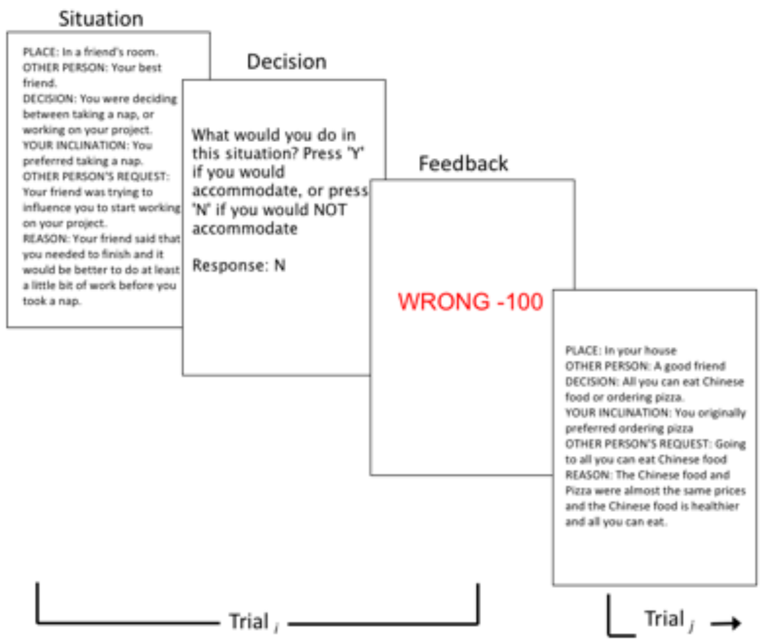

Figure 3 (Study 1).

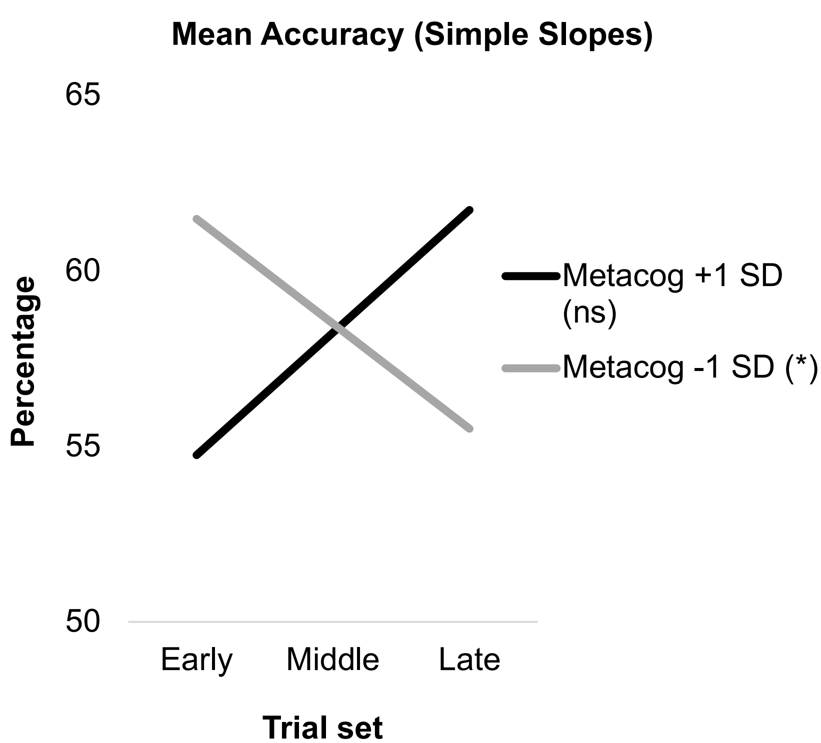

Figure 4 (Study 3).

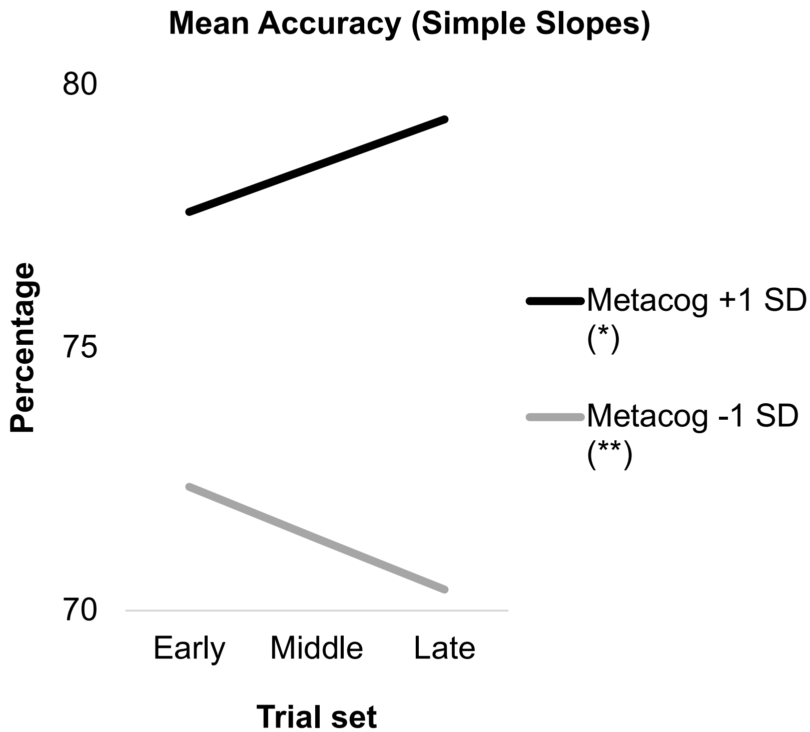

Figure 5 (Study 3).

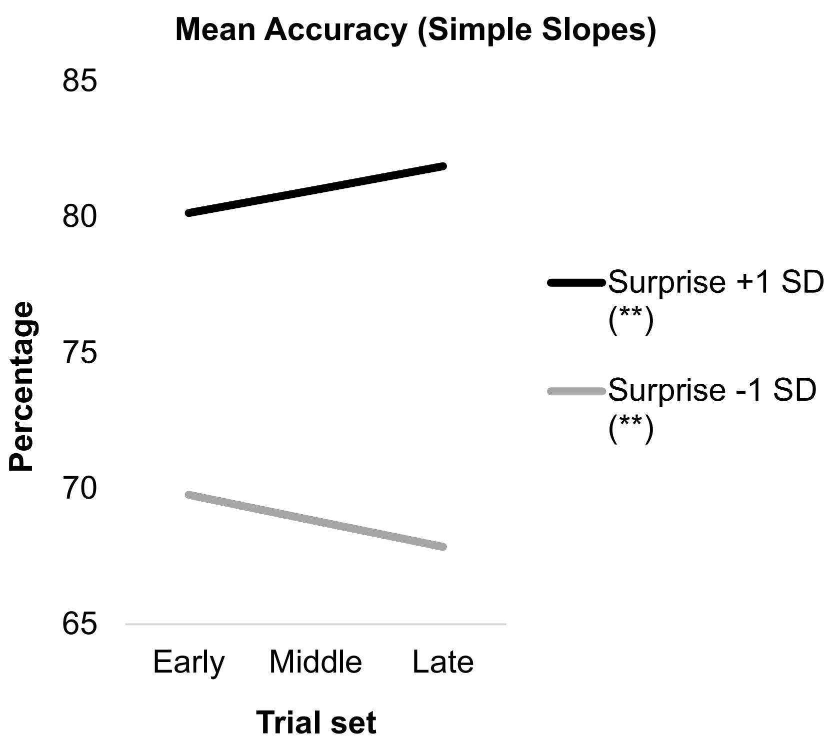

Figure 6 (Study 4).

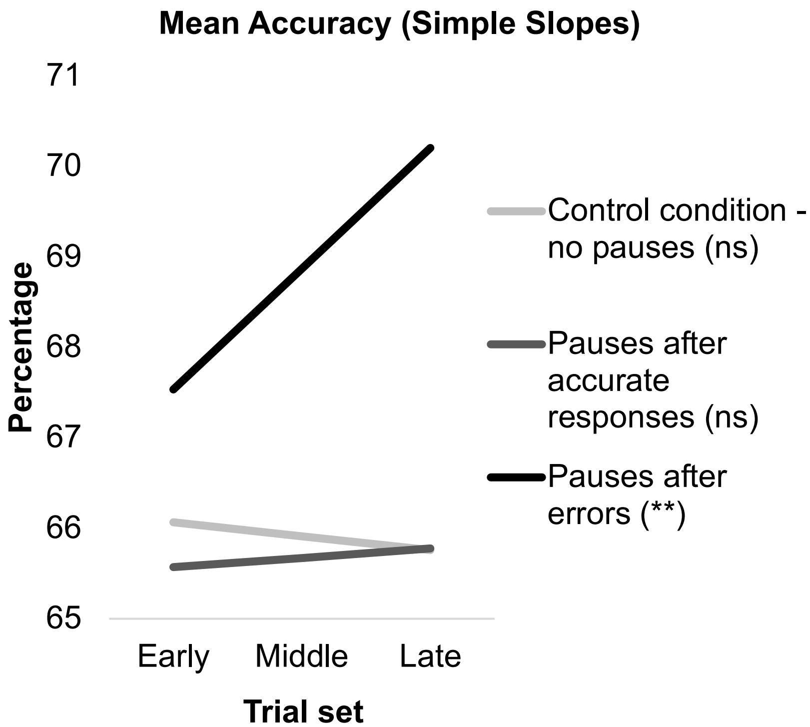

Figure 7 (Study 5).

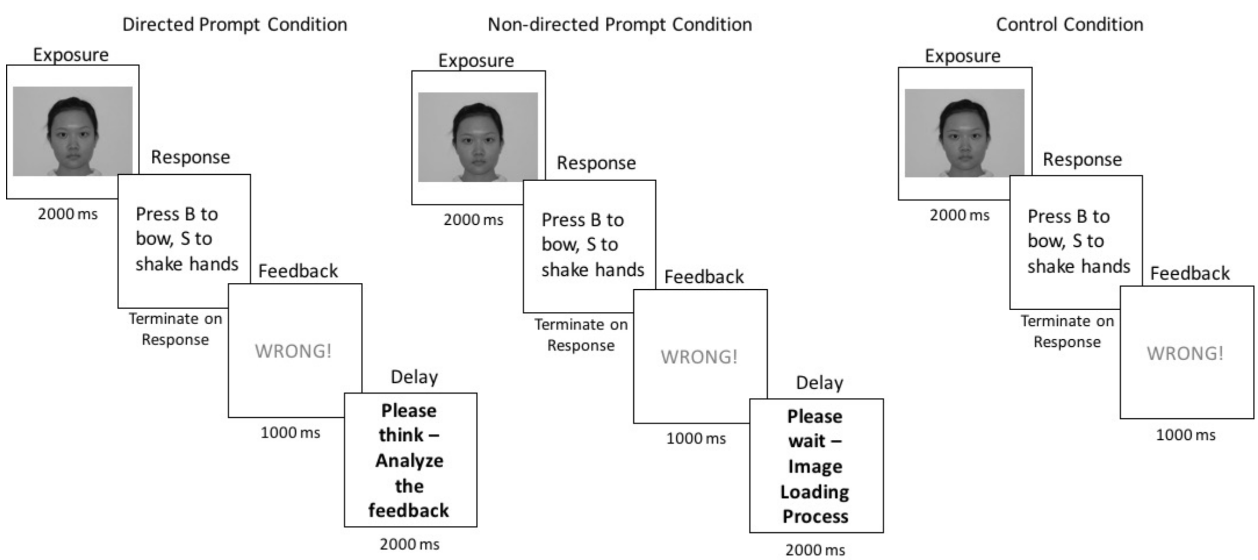

Figure 8 (Study 5).

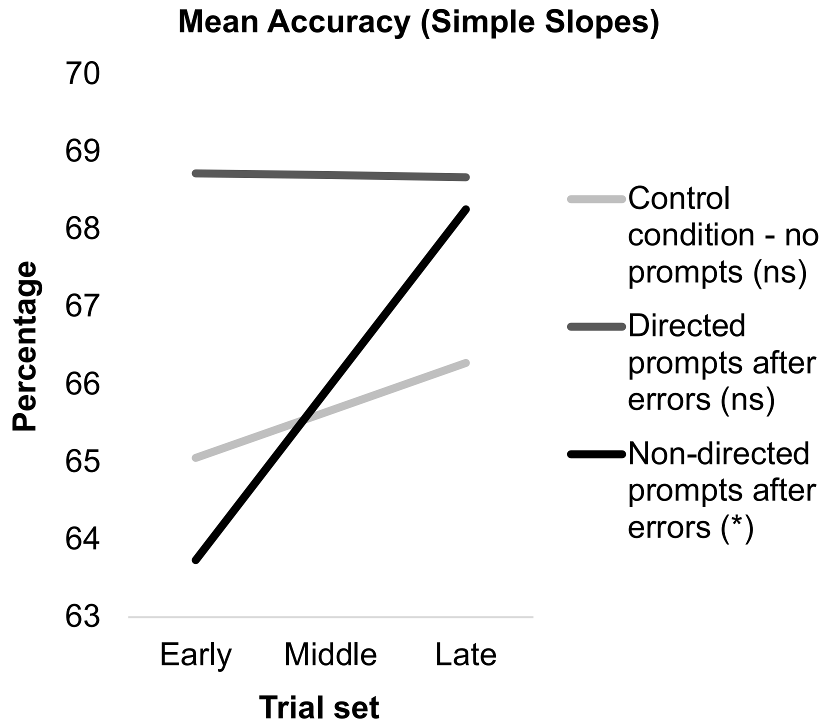
Mean Accuracy (Simple Slopes)
70
69
68
67
66
65
64
63
Percentage
Early
Middle
Late
Trial set
Control condition - no prompts (ns)
Directed prompts after errors (ns)
Non-directed prompts after errors (*)

Table 1. *Descriptive statistics and Correlations between variables measured in Study 3.*

	Mean	SD	1	2	3	4	5	6
1. Average Accuracy	0.75	0.17	-					
2. Surprised Difference	2.32	2.11	0.37***	-				
3. Cultural Metacognition	5.21	1.10	0.16*	0.099	(0.88)			
4. Cultural knowledge	3.47	1.17	0.048	-0.0036	0.48***	(0.86)		
5. Cultural motivation	4.99	1.18	0.013	-0.010	0.50***	0.43***	(0.87)	
6. Cultural behavior	4.78	1.24	0.074	-0.0046	0.58***	0.45***	0.60***	(0.91)

Reliabilities are reported in parentheses on the diagonal.
**p* < 0.05; ***p* < 0.01; and ****p*<0.001 (two-tailed)

Table 2. *Results of the hierarchical linear regression models tested in Study 3.*

Predictor	Model 1 Results	Model 2 Results	Model 3 Results
Trial set	-0.00097 (.0044), [-0.0096, 0.0077], 0.22	-0.00097 (.0044), [-0.0095, 0.0076], 0.22	-0.00097 (.0043), [-0.0093, 0.0074], 0.23
Cultural Metacognition	0.032 (.012), [0.0081, 0.056], 2.62**		0.023 (.012), [-0.00054, 0.046], 1.91
Cultural knowledge	-0.0030 (.011), [-0.025, 0.019], 0.27		-0.00085 (.010), [-0.021, 0.019], 0.08
Cultural motivation	-0.013 (.012), [-0.037, 0.010], 1.10		-0.011 (.011), [-0.034, 0.011], 0.99
Cultural behavior	0.0020 (.013), [-0.023, 0.027], 0.16		0.0050 (.013), [-0.020, 0.030], 0.39
Metacognition X Order	0.017 (.0047), [0.0076, 0.026], 3.57***		0.014 (.0046), [0.0051, 0.023], 3.07**
Cultural knowledge X Order	-0.0026 (.0049), [-0.012, 0.0069], 0.53		-0.0019 (.0046), [-0.011, 0.0071], 0.42
Cultural motivation X Order	0.00053 (.0046), [-0.0085, 0.0096], 0.12		0.0011 (.0045), [-0.0077, 0.010], 0.25
Cultural behavior X Order	-0.014 (.0051), [-0.024, -0.0039], 2.73**		-0.013 (.0050), [-0.023, -0.0032], 2.61**
Surprise		0.029 (.0045), [0.020, 0.038], 6.42***	0.028 (.0047), [0.019, 0.037], 5.92***
Surprise X Order		0.0087 (.0019), [0.0050, 0.012], 4.62***	0.0079 (.0018), [0.0043, 0.011], 4.34***

Results are reported in the format: *B* (*SE*), 95% CI, z-value
**p* < 0.05; ***p* < 0.01; and ****p*<0.001 (two-tailed)

Table 3. *Results of the linear regression models tested in Study 3.*

Independent variable	*B*	95% CI	*SE (B)*	β	*t*
Cultural Metacognition	0.34	[0.0017, 0.68]	.17	.18	1.98*
Cultural knowledge	-0.078	[-0.37, 0.21]	.15	-.043	0.53
Cultural motivation	-0.078	[-0.39, 0.24]	.16	-.043	0.48
Cultural behavior	-0.11	[-0.43, 0.21]	.16	-.064	0.67
df					208

*$p < 0.05$; **$p < 0.01$; and ***$p<0.001$ (two-tailed)