This study explored the relationships of self-reflection and insight with individuals' performances on various language tasks. The Self-Reflection and Insight Scale (SRIS; Grant, Franklin, & Langford, 2002) assessed individual differences in three factors: engagement in reflection, need for reflection, and insight. A high need for reflection was associated with a low vocabulary level and a small number of intrusive errors on a recognition task. A high level of insight was associated with a high accuracy rate and high response speed on a sentence-sensibility judgment task. In addition, groups who differed in need for reflection and insight did not differ on a working memory test, suggesting that need for reflection and insight account for the variation in the language tasks independent of working memory capacity. The findings suggest connections between self-reflection, insight, and proficiency in language processing.

Key words: self-reflection, insight, individual differences, lexical access, sentence reading

People vary in the way that they process information. This study investigated how individual differences in language processing might relate to people's tendencies to self-reflect and their levels of understanding about their own mental processes. Some stable differences at various levels of cognitive functioning have been reported in the literature. For example, at the perceptual level, when asked to find familiar figures from among a group of distracters, some people preferred to respond as quickly as possible, whereas others preferred to respond as accurately as possible (Kagan, Rosman, Kay, Albert, & Phillips, 1964). Kagan (1966) called this contrast impulsivity versus reflection. At the level of higher order cognitive functioning, Agor (1984, 1989) found that some people tended to use an analytical approach, breaking a problem down into manageable parts, whereas others tended to use an intuitive approach. Individual differences have also been
found through research on learning styles in educational settings (e.g., Entwistle, Hanley, & Hounsell, 1979; Gregorc, 1982; Li & O'Boyle, 2008). In a recent review of the literature, Kozhevnikov (2007) introduced Nosal's model about individual differences at different levels of information processing. As Kozhevnikov pointed out, Nosal’s model highlighted the connection of the different tendencies in self-monitoring and control to the different characteristics of performances on cognitive tasks. Nosal’s position converges with other researchers' theoretical arguments about the important role of cognitive control and regulation in cognition and intelligence (e.g., Sternberg, 1985; Toates, 2006). Consistent with this view, empirical studies on individual differences have reported that some individuals seem to exert more control over their mental states and processes than others (e.g., Keller & Ripoll, 2001).

In the literature of personality, researchers consider individuals to be rather consistently different in terms of whether they regularly monitor and reflect upon their thoughts, feelings, and behaviors. Frequent self-reflection is thought to be associated with a high level of self-consciousness (e.g., Fenigstein, Scheier, & Buss, 1975; Grant, Franklin, & Langford, 2002; Ingram, 1990; Watson, Morris, Ramsey, Hickman, & Waddell, 1996). However, in the past, the construct of self-reflection was mainly studied by researchers and practitioners in areas such as assessment, counseling, and coaching for the purpose of promoting behavior change and mental health. There are few empirical studies aimed directly at the relationship between different dispositions in self-reflection and individual differences in cognitive functioning, even though one's tendency to engage in self-reflection seems logically related to the level of cognitive control and regulation, which in turn should be closely related to how a person processes information, as suggested by research and theories on individual differences (for a detailed review, see Kozhevnikov, 2007).

The current study attempted to explore the relation between self-reflection and one critical aspect of human cognition—language. Language-processing tasks, such as reading and verbal comprehension, are an important aspect in everyday functioning and in academic achievements. Yet, among the models and theories about individual differences in information processing (e.g., Kozhevnikov, 2007; Miller, 1987), it remains to be discussed how individuals may consistently differ in various aspects of language ability, such as vocabulary level, the process of lexical access, the speed and accuracy of sentence reading, and the retrieval of text information. More importantly, given the critical role of cognitive control and regulation to information processing, how might these potential differences in language processing relate to the different dispositions in self-reflection? On one hand, it seems that constant self-reflection could lead to better understanding about the mechanism and process of language comprehension, and thus may help in making cognitive adjustment and exerting efficient control over the process. On the other hand, however, self-reflection could be interruptive to the comprehension process due to the consumption of extra mental resources. In addition, excessive self-reflection and a feeling of uncertainty might go hand in hand, therefore resulting in a negative impact on the process of reading or verbal comprehension. This study aimed to examine individual differences in some basic elements involved in language comprehension and to investigate the relation between self-reflection and language processing.
SELF-REFLECTION, INSIGHT, AND LANGUAGE

Assessment of Disposition in Self-Reflection

The Self-Reflection and Insight Scale (SRIS) is a 20-item scale developed by Grant et al. (2002). Compared to other similar scales for self-reflection or self-consciousness (e.g., Fenigstein et al., 1975; McKenzie & Hoyle, 1999; Trappnell & Campbell, 1999), one of the advantages of the SRIS is its delineation of the two different factors: self-reflection (SRIS-SR) and insight (SRIS-IN). The two subscales, SRIS-SR and SRIS-IN, were designed to measure “the inspection and evaluation of one’s thoughts, feelings, and behaviors” and “the clarity of understanding of one’s thoughts, feelings, and behaviors,” respectively (Grant et al., 2002, p. 821). The SRIS-SR includes items such as “I frequently examine my feelings” and “I have a definite need to understand the way my mind works.” The SRIS-IN includes items such as “Thinking about my own thoughts makes me more confused” and “I usually know why I feel the way I do.” Respondents indicate on a 6.0 scale how much each statement describes them. The internal consistency coefficients for the two subscales were 0.91 and 0.87, respectively. The test-retest reliability across a 7-week interval was 0.77 for SRIS-SR and 0.78 for SRIS-IN. This delineation of self-reflection versus insight converges with theoretical arguments on self-consciousness (e.g., Pinku & Tzelgov, 2006) and has found some support from empirical studies. For example, Grant et al. (2002) found that SRIS-IN, but not SRIS-SR, was positively associated with cognitive flexibility. Lyke (2009) found that SRIS-IN, but not SRIS-SR, was positively associated with satisfaction with life and subjective happiness. In addition, some data suggested that reflection did not always lead to insight and sometimes can even be negatively related to the level of insight (Grant et al., 2002). Accordingly, the current study chose this scale to assess self-reflection and, at the same time, to examine the role of insight in the potential connection between self-reflection and the advantages or disadvantages in language capacities.

A second advantage of the SRIS as a measure for self-reflection is the differentiation of execution and motive of self-reflection. Originally, Engagement in Reflection and Need for Reflection were two separate subscales underlying SRIS-SR. The Engagement in Reflection subscale was intended to evaluate the frequency of the actual act of self-reflection, for example, “I frequently examine my feelings,” whereas the Need for Reflection subscale was intended to assess the level of motive to conduct self-reflection, for example, “I have a definite need to understand the way my mind works.” In their studies, Grant et al. (2002) found that these two original subscales often loaded on the same factor and eventually combined them as one factor, SRIS-SR. However, researchers tend to believe that acts and motives are theoretically independent (e.g., Grant et al., 2002; Trappnell & Campbell, 1999). Another reason that engagement in reflection and need for reflection should be examined separately is the inconsistent findings on the relations between SRIS-SR and SRIS-IN. As discussed previously, some data suggested that reflection leads to insight, whereas some other data indicated a negative association between reflection and insight (Grant et al., 2002). Possibly, the nondifferentiation between acts of reflection and intents for reflection caused this obscurity. The current study examined engagement in reflection and need for reflection separately, as well as their relations with language abilities. Therefore, the current study followed Grant et al.’s original design.
and considered three factors: engagement in reflection, need for reflection, and insight.

Overview of the Experiments

This study included four experiments that examined how individuals who differed in the engagement of self-reflection, the need for self-reflection, or the level of insight might differ in their performances on common cognitive tasks involving language processing. More specifically, the participants in each experiment were categorized into a high-score group and a low-score group based on each of the three SRIS subscale scores. Experiment 1 examined the differences in vocabulary level between different groups. Experiments 2 and 3 compared the performances of different groups in various aspects of language processing with materials such as common words and simple sentences. Finally, Experiment 4 investigated whether the groups differed in their working memory capacity, which is considered an important factor underlying individual differences in reading and verbal comprehension. Neither the original Grant et al. (2002) study nor studies conducted in this lab found any gender differences in the SRIS scores. Therefore, for the sake of succinctness, gender information is excluded.

Experiment 1

Vocabulary knowledge is an important aspect of language ability and is considered a single index for academic achievement (Dale & Reichert, 1957). It certainly plays a critical role in reading and verbal comprehension. In addition, research has shown that word learning and conceptual development are intertwined (e.g., Clark, 1993; Lucariello, 1987; Rhys-Jones & Ellis, 2000; Tomasello & Farrar, 1986). Therefore, it would be informative to examine what roles self-reflection and insight might play in the development of vocabulary knowledge. This experiment looked at the vocabulary levels of people who scored on the high end versus the low end of the subscales of the SRIS. It was hypothesized that both high self-reflection and high insight would be associated with high vocabulary level. Studies have shown that people understand abstract concepts at a later stage of conceptual development (Gilhooly & Logie, 1980; Keil, 1979). Subjective experiences play an important part in the representation of abstract concepts (Barsalou, 1999; Wiemer-Hastings & Xu, 2005). Frequent introspection and clear insight should be associated to some extent with the comprehensive understanding of one’s subjective experiences. Comprehensive understanding of these experiences should be reflected as rich knowledge of complex, abstract concepts denoted usually by words learned at a later stage of one’s life. As a result, more reflective and insightful individuals should possess richer vocabulary knowledge overall.

Method

Participants. One hundred and sixty-five (46% female) undergraduate students were recruited from introductory psychology classes at a large university. All considered English as their first language. They received extra credit points for their participation in this experiment.
Measures and procedure. The SRIS measured the tendency to engage in self-reflection, the level of need for self-reflection, and the level of insight. Most participants completed the questionnaire in 5 min. The Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981) assessed the level of vocabulary knowledge. The PPVT-R was designed to measure a person’s vocabulary knowledge in American English. Its reported reliability ranged from 0.61 to 0.88. In each trial of the current experiment, the participants were presented with four pictures simultaneously on a computer screen, at a rate of 12 s/slide, and the pronunciation of a word over a headphones. The pronunciation of each word was prerecorded with a male voice in American English as a .wav file and could be played repeatedly to the participants as needed. The participants indicated which one of the four pictures best depicted the meaning of the word and wrote down their answers for scoring purposes. Most participants finished this task within 30 min. About half of the participants received the SRIS first; the other half of the participants received the PPVT-R first.

Results and Discussion

Two participants did not complete the PPVT-R and thus were excluded from the analysis. The standard PPVT-R scores ranged from 68 to 127, and the percentile ranks ranged from 2% to 96%. No outliers were detected. A tertiary split was performed based on each of the three SRIS subscale scores. The participants who scored above the 67th percentile or below the 34th percentile on each subscale were compared in terms of their standard PPVT-R scores. Table 1 presents the group means and standard deviations of the standard PPVT-R scores.

T tests did not show any difference in standard PPVT-R scores between the high- and low-engagement-of-reflection groups, $t < 1$, but indicated that the group with a high level of need for reflection had significantly lower standard PPVT-R scores than the group with a low level of need for reflection, $t(107) = 2.251, p < .03, \eta^2 = .05$. That is, as measured by the Need for Reflection subscale, a lower level of need for self-reflection is associated with a higher level of vocabulary proficiency. This is contrary to the prediction that self-reflection might be positively related to vocabulary level, which certainly suggests that the relationship between self-reflection and vocabulary knowledge is more complex than expected. One possible explanation is that the need for self-reflection or introspection is only beneficial to the individual when it is at a moderate level. Excessive need for reflection can be detrimental due to its constant occupation of mental resources. As a matter of fact, some researchers have pointed out that it is not always easy to differentiate the positive ways of reflection from the negative ways of rumination. For example, Anderson, Bohon, and Berrigan (1996) argued that questions designed to assess self-reflection sometimes tap on a dysfunctional self-absorption instead of a constructive self-reflection. If, consistent with these researchers’ argument, very high scores on the Need for Reflection subscale to some extent reflect a negative approach of introspection, the result suggests that the need for excessive attention to be directed to the inner world could be a factor against positive progress in language proficiency. This is obviously a speculation that calls for more extensive research.
T tests did not reveal any difference in standard PPVT-R scores between the high-insight and the low-insight groups, $t < 1$. The lack of difference in PPVT-R scores between the high-insight and the low-insight groups is not in line with the prediction that a higher insight level would be associated with a higher level of vocabulary knowledge. This may suggest that a clear understanding of subjective experiences does not necessarily contribute or transform to the accumulation of vocabulary knowledge. Alternatively, this

<table>
<thead>
<tr>
<th>Experiment 1</th>
<th>PPVT score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low engagement</td>
<td>99 (13)</td>
</tr>
<tr>
<td>High engagement</td>
<td>99 (11)</td>
</tr>
<tr>
<td>Low need</td>
<td>101 (11)</td>
</tr>
<tr>
<td>High need</td>
<td>95 (14)</td>
</tr>
<tr>
<td>Low insight</td>
<td>98 (13)</td>
</tr>
<tr>
<td>High insight</td>
<td>99 (13)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experiment 2</th>
<th>Accuracy (%)</th>
<th>Reaction time (ms)</th>
<th>Hits</th>
<th>False alarms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low engagement</td>
<td>97 (4)</td>
<td>586 (100)</td>
<td>10.00 (2.66)</td>
<td>2.77 (2.95)</td>
</tr>
<tr>
<td>High engagement</td>
<td>96 (3)</td>
<td>576 (95)</td>
<td>9.25 (3.44)</td>
<td>2.67 (2.27)</td>
</tr>
<tr>
<td>Low need</td>
<td>97 (4)</td>
<td>608 (153)</td>
<td>11.00 (3.16)</td>
<td>3.46 (2.72)</td>
</tr>
<tr>
<td>High need</td>
<td>96 (3)</td>
<td>607 (125)</td>
<td>9.45 (2.77)</td>
<td>1.18 (1.60)</td>
</tr>
<tr>
<td>Low insight</td>
<td>98 (3)</td>
<td>570 (88)</td>
<td>9.71 (2.40)</td>
<td>3.36 (2.59)</td>
</tr>
<tr>
<td>High insight</td>
<td>96 (3)</td>
<td>637 (160)</td>
<td>10.33 (4.07)</td>
<td>2.00 (1.95)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experiment 3</th>
<th>Accuracy (%)</th>
<th>Response time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low engagement</td>
<td>95 (3)</td>
<td>484 (154)</td>
</tr>
<tr>
<td>High engagement</td>
<td>96 (2)</td>
<td>499 (180)</td>
</tr>
<tr>
<td>Low need</td>
<td>95 (2)</td>
<td>504 (132)</td>
</tr>
<tr>
<td>High need</td>
<td>96 (3)</td>
<td>479 (163)</td>
</tr>
<tr>
<td>Low insight</td>
<td>95 (2)</td>
<td>569 (138)</td>
</tr>
<tr>
<td>High insight</td>
<td>97 (2)</td>
<td>455 (190)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Experiment 4</th>
<th>Ospan score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low engagement</td>
<td>41 (12)</td>
</tr>
<tr>
<td>High engagement</td>
<td>41 (12)</td>
</tr>
<tr>
<td>Low need</td>
<td>42 (10)</td>
</tr>
<tr>
<td>High need</td>
<td>40 (12)</td>
</tr>
<tr>
<td>Low insight</td>
<td>40 (9)</td>
</tr>
<tr>
<td>High insight</td>
<td>41 (12)</td>
</tr>
</tbody>
</table>
result may be due to the nature of the PPVT-R. That is, although it includes abstract and complex words, this test mainly contains words with meanings that can be rather easily portrayed by images. People who are more insightful and have clearer understanding about their subjective experiences may have advantages in acquiring and processing concepts lacking imagery referents (Paivio, 1986).

In summary, Experiment 1 revealed that individuals with low levels of need for self-reflection performed significantly better than individuals with high levels of need for self-reflection on the vocabulary test. This pattern of result seems to suggest a negative correlation between need for reflection and vocabulary level. A bivariate correlational analysis showed that the strength of the correlation was rather low, $r(163) = -.16$, $p < .05$. The weak correlation indicates that a significant difference in vocabulary knowledge can only be noted between individuals who differ considerably in their levels of need for reflection. That is, on one hand, the finding of this experiment adds to the literature on the connection between the tendency to self-reflect and the level of vocabulary knowledge. On the other hand, it indicates that future research needs to further explore the relationship between the two, as well as other factors that may be associated with both variables or mediate the relationship between them.

**Experiment 2**

Experiment 2 compared the individuals who scored high or low on the SRIS subscales in terms of their speed and accuracy of lexical access, as well as their levels of processing while reading language materials. A priming paradigm, a common approach in cognitive and language research, was employed for the purpose of this experiment. In a typical primed lexical decision paradigm, researchers present participants a prime (e.g., a sentence) and then ask them to make a quick judgment on whether the target, a letter string following the prime, is a real word. For the purpose of this study, the reaction time and accuracy rate of participants’ lexical decisions on target words were evaluated. In addition, participants’ processing of the prime sentences was assessed by a memory recognition task. The hypothesis of this experiment was that an individual’s disposition in self-reflection and level of insight would be reflected in the process of comprehending prime sentences and making lexical decisions on target words. More specifically, the more reflective an individual is, the deeper this person may tend to process the priming sentences. Based on the level-of-processing theory, the deeper the processing level, the better the performance often is on recall or recognition tasks (Craik & Brown, 2000; Craik & Lockhart, 1972). Therefore, the more reflective an individual is, the better the person's performance should be on the recognition task with the primes. In addition, insightful understanding of one's cognitive processes might be reflected as swift responses to the lexical decision task, especially when under time pressure.

To control complexity, this experiment used a pseudo primed lexical decision design. That is, a set of sentences and a set of letter strings were presented alternatively, but there were no priming effects intended by design. By limiting the interaction of the two tasks, the relations of the SRIS with the performances on the two tasks, lexical decision and memory recognition, could be evaluated independently.
Method

Participants. Thirty-nine (57% female) undergraduate students were recruited from a number of psychology classes at a large university. All considered English as their first language. They received extra credit points for their participation in this experiment.

Materials. A set of 48 sentences was written with the basic syntactical structure, subject–verb–object, describing common, everyday situations. Half of the sentences were written in question format and the other half in statement format. Sample sentences include “You planned the party,” “You took the pill,” “Did you check the spelling?” and “Did you accept the gift?” The sentences were kept very simple; few modifiers were included. The number of words in the sentences ranged from four to six. A set of 24 high-frequency words and a set of 24 nonword letter strings were constructed. Example words include drop, report, and sense. The numbers of syllables ranged from one to four and were largely matched between words and nonwords. The SRIS measured the tendency to engage in self-reflection, the need for self-reflection, and the level of insight.

Procedure. The 48 sentences and 48 letter strings (words and nonwords) were presented alternatively on a computer screen. Each sentence was on display for 1.9 s, followed by a letter string. Each letter string was on display until the participant responded to it. The participants were given eight practice trials at the beginning to familiarize themselves with the task. They were instructed to read each sentence carefully, but not informed of the need for recall, and to respond to each letter string as quickly and accurately as possible. The computer recorded reaction time and accuracy rate and gave feedback to the participants after each trial. After participants completed this task, they filled out the SRIS. Finally, they were asked to indicate on a sheet of paper with 48 sentences which ones they had seen previously. Half of the sentences on the paper were randomly selected as target sentences from the previous task; half of them were distracters that had not been seen by the participants. They were constructed for this recognition task by using the same syntactical structure but different verbs and different object terms. Most participants finished the experiment in 30 min.

Results and Discussion

Compared with the average lexical decision time and accuracy rate reported in the literature (e.g., Blasko & Connine, 1993; Swinney, Love, Walenski, & Smith, 2007), the participants’ performances on the lexical decision task, including speed (M = 609, SD = 124) and accuracy (M = 97%, SD = 3%), were comparable to reports by studies using the priming approach, which indicates their attentiveness to the task. Participants’ responses to the recognition task were categorized as hits or false alarms. Hits were the correctly identified target sentences, and false alarms were incorrectly selected distracter sentences. The numbers of hits ranged from four to 16, and the number of false alarms ranged from zero to eight. The participants on average only recognized about 10 of the 24 target sentences previously shown on the computer screen. This was probably due to the fact that the participants were not informed beforehand of the recognition task. No outliers were detected for any measures. A tertiary split was performed based on each of the three SRIS subscale scores. The participants who scored above the 67th
percentile or below the 34th percentile on each subscale were compared in terms of their hit scores and false alarm scores in the memory task, as well as their speed and accuracy in the lexical decision task. Table 1 presents the group means and standard deviations.

* T* tests did not reveal any differences in performances on the recognition task or the lexical decision task between the high- and the low-engagement-in-reflection groups, all *t’s* < 1. Neither were there any differences on the lexical decision task between the high- and the low-need-for-reflection groups. However, the *t* test showed that the high-need-for-reflection group made significantly less intrusive errors (i.e., false alarms) than the low-need-for-reflection group, *t*(22) = 2.425, *p* < .03, *η*² = .21, even though the two groups did not differ significantly in the number of correctly identified sentences (i.e., hits), *p* > .20. That is, a higher need for checking on oneself seemed to be associated with a lower level of false identification, but not with a higher level of correct recognition. This may suggest that a strong need for self-inspection mainly renders control over intrusive errors during memory retrieval by staying clear of interference from the distracters. In addition, it could also be the case that individuals with a strong need for reflection are simply more cautious and more reluctant to risk selecting uncertain sentences in fear of making intrusive errors. More research is certainly necessary to elucidate how much the constant need for reflection functions as a factor that resists interference during the retrieval stage, and how much it serves as a factor that promotes a conservative task-taking approach. Furthermore, the finding that the high-need-for-reflection group made significantly fewer intrusive errors than the low-need-for-reflection group suggests that the Need for Reflection subscale of the SRIS is a simple, straightforward tool to assess individuals’ different tendencies to self-reflect. That is, participants’ actual performances on the memory task as measured by the number of intrusive errors converged with their self-evaluation about their levels of intent to conduct self-monitoring.

The high-insight group and the low-insight group did not differ in the memory recognition task or the lexical decision task, all *p’s* > .15. This may suggest that insight is not as relevant to either sentence processing or lexical access. Alternatively, the lack of differences between the high-insight and the low-insight groups might be due to the particular designs of the tasks. That is, different dispositions in insightfulness might manifest themselves only when the tasks require more conscious effort and control. Yet, the tasks in this experiment might not meet this condition. First, lexical access is considered an automatic process (Posner & Snyder, 1975). In this experiment, the participants on average needed only about 600 ms to make the lexical decisions. Second, the recognition task was a surprise task. The participants might not have been as motivated as expected when processing the sentence primes, which may explain the relatively low average recognition rate. With these considerations, Experiment 3 tested individual differences in sentence processing but employed a different type of task that required the participants to pay attention to each presented sentence.

**Experiment 3**

Reading comprehension is an important ability to evaluate for the purposes of research, instruction, and diagnosis. Experiment 3 followed the previous experiment to test whether different tendencies in self-reflection or
different levels of insight were relevant to individual differences in sentence comprehension. This experiment utilized a sensibility-judgment task with sentence targets instead of single words, based on the consideration that the design of the task and the time duration required for sentence comprehension may allow more pronounced manifestations of different dispositions in self-reflection and insight during the task.

Method

Participants. Fifty-nine (52% female) undergraduate students were recruited from a large university. All considered English as their first language. They received $5 for their participation in the experiment.

Materials. The experiment used the same set of 48 sentences from Experiment 2. Another set of 48 nonsensical sentences was written for this experiment. These sentences described illogical or impossible situations. Examples of nonsensical sentences include “You chucked the porch” and “Did you smell the noise?” The syntactical structures and the length of the nonsensical sentences generally matched the meaningful sentences. The SRIS measured the tendency to engage in self-reflection, the need for reflection, and the level of insight.

Procedure. The 48 meaningful sentences and 48 nonsensical sentences were presented, one at a time, in a random order on a computer screen. The participants were given 12 practice trials at the beginning to become familiarized with the task. They were instructed to press and hold down a key to display a sentence on the computer screen and then to release this key and press another key to indicate whether the sentence made sense or not. They needed to respond to each sentence as quickly and accurately as they could. The computer recorded the accuracy rate and the time duration between the offset of the sentence and the response as the reaction time. After participants completed this task, they filled out the SRIS. Most participants finished the experiment in 25 min.

Results and Discussion

One participant had unusually long reading times ($z = 3.72$), and another participant generated an exceptionally low score on one of the SRIS subscales ($z = -3.78$ on the Need for Reflection subscale). Thus, they were excluded as outliers from the following analyses. (Analysis including these two cases yielded virtually the same results.) A tertiary split was performed based on each of the three SRIS subscale scores. The participants who scored above the 67th percentile or below the 34th percentile on each subscale were compared in terms of their response times and accuracy rates. Table 1 presents the group means and standard deviations.

The high- and the low-engagement-in-reflection groups were not different in reaction time or accuracy rate, both $p$’s > .10. Neither were the high- and the low-need-for-reflection groups, both $p$’s > .15. However, the high-insight group was significantly more accurate than the low-insight group, $t(35) = 2.25, p < .02, \eta^2 = .15$. In addition, they were significantly faster than their low-insight counterparts, $t(35) = 2.07, p < .05, \eta^2 = .11$. It appeared that higher clarity of understanding about one’s own mental processes was associated with more decisive and more accurate responses in the
process of understanding and judging the meaningfulness of the sentences. Considering that the classifications of the high-insight group and the low-insight group were based on participants’ self-reports, this finding about the actual performance differences between the two groups as measured by observable variables (i.e., reaction time and accuracy) indicates the validity of the participants’ self-evaluation about how well they understand their own mental processes. Taken together with the finding from Experiment 2 about the performance difference between the high- and low-need-for-reflection groups in the memory task, this is positive evidence for the value of the SRIS as a self-report measure.

This experiment used a sensibility-judgment task with sentence materials. Compared to the surprise memory task and the lexical decision task in the previous experiment, this task required more consciously controlled effort from the participants. The findings from this experiment seemed to be in line with the speculation made at the end of Experiment 2, that is, that different dispositions in insightfulness would manifest in a more controlled process. However, in contrast to the previous experiment, the high- and low-need-for-reflection groups did not perform differently on this task. The contrasts between the findings of the two experiments call for more attention to the distinction between self-reflection and insight and more extensive research on the roles of self-reflection versus insight in various cognitive tasks.

**Experiment 4**

Working memory has been considered an important construct in models about cognitive functioning (e.g., Anderson & Lebiere, 1998; Cowan, 1995). Scores on working memory tasks have been shown to be predictive of performance on various cognitive tasks, including reading comprehension (Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, 2001). One possible interpretation of the findings from Experiments 2 and 3 is that working memory capacity is the underlying variable responsible for all of the variation detected in the recognition task and in the sensibility-judgment task. The speculation is that individuals who possess greater mental resources may also be the ones who are more reflective and more insightful. As a result, this connection was reflected as the group differences in their task performances in Experiments 2 and 3. The purpose of Experiment 4 was to test the relationship of working memory capacity with self-reflection and insight. It was hypothesized that there might be a circular connection between these constructs. That is, monitoring and regulation during a cognitive task require mental resources. On one hand, higher working memory capacity could provide affordance for engagement in self-reflection and possibly also insight. On the other hand, constant engagement in self-inspection and clear understanding of inner processes might help to utilize mental resources more efficiently, and thus achieve better task performance.

**Method**

**Participants.** Fifty-five (49% female) undergraduate students were recruited from a large university. All considered English as their first language. They received $5 for their participation in this experiment.
**Measures and procedure.** The Turner and Engle (1989) Ospan task assessed working memory capacity. Participants were presented a series of simple math problems, each with a correct or an incorrect answer (e.g., $8/2 - 1 = ?$), and a series of unrelated words alternatively on a computer screen. Each trial contained two to six pairs of math problems and words. The task was to remember the words in the order they were presented within each trial while judging the correctness of the answers to the math problems. The numbers of correctly judged math problems were recorded. The numbers of correctly recalled words were summed up as the Ospan score. The maximum math accuracy rate is 100%, and the maximum Ospan score is 60. An Ospan score is valid only when the 85% math accuracy criterion is met. The SRIS then measured the tendency to engage in self-reflection, the need for self-reflection, and the level of insight. All participants completed the task within 40 min.

**Results and Discussion**

Two participants did not meet the math accuracy criterion. Their Ospan scores were excluded from the analysis. The Ospan scores of the remaining participants ranged from 19 to 60. No outliers were detected. As in the previous experiments, a tertiary split was performed based on each of the three SRIS subscale scores. The participants who scored above the 67th percentile or below the 34th percentile on each subscale were compared in terms of their Ospan scores. Table 1 presents the group means and standard deviations.

$T$ tests did not reveal any differences between the high- and low-engagement groups, between the high- and low-need-for-reflection groups, or between the high- and low-insight groups, all $t$'s < 1. It seemed that, as measured by Ospan scores, working memory capacity appeared to be independent of the tendency to engage in self-reflection, the need for self-reflection, and the level of insight. Working memory capacity is considered to be an important factor that accounts for a significant part of the variance in intellectual ability (e.g., Conway, Cowan, Bunting, Thrierrault, & Minkoff, 2002; Engle et al., 1999; Kane et al., 2001). As discussed earlier, working memory capacity predicts performance in various cognitive tasks including reading comprehension (e.g., Engle et al., 1999; Daneman & Carpenter, 1980). For example, a meta-analysis by Daneman and Merikle (1996) that included 77 studies showed that the score of working memory span was a good predictor for performance in reading. The results of this experiment showed that the capacity of working memory did not seem to be directly relevant to or restricting of the engagement in self-reflection or the clear understanding about the self. Likewise, self-reflection and evaluation was not necessarily associated with efficient use of limited mental resources. Consequently, the lack of connection between the Ospan and the SRIS subscale scores suggests that the SRIS subscale scores contributed to the group differences in task performances in the previous experiments independent of working memory capacity. Self-reflection and insight therefore should be taken into consideration as individual difference variables in future research on language proficiency and research using tasks involving language materials.
General Discussion

Four experiments were conducted to investigate the roles of self-reflection and insight in some common tasks often used in cognitive and psycholinguistic research. Two of the three SRIS subscales, Need for Reflection and Insight, were found to be particularly relevant to the tasks in this study. A high level of need for reflection was associated with a low vocabulary level and a low level of intrusive errors in memory recognition tasks using sentence materials (Experiments 1 and 2). A high level of insight was associated with a high level of accuracy and fast responses in the sentence-sensibility judgment task (Experiment 3). The individual differences in self-reflection and insight were independent of an important cognitive construct, working memory capacity (Experiment 4), which suggests that self-reflection and insight account for the differences in language ability and processing revealed in the previous experiments. These findings on the connection of language processing with self-reflection and insight are intriguing because the SRIS scores were generated through simple self-evaluation. Despite their differences in reflective tendencies and in levels of insightfulness, individuals seem to have a good understanding about the way they function cognitively and can provide relatively reliable evaluations in these aspects.

The weakness of the self-assessment approach is well acknowledged in psychological research and other fields. Self-assessment about one’s own internal processes can be even more problematic. The findings from this study and an earlier study (Xu, 2009) both indicated that the SRIS appeared to be a simple and effective tool to assess individual differences in this aspect. Xu asked college students to submit written descriptions about various mental processes (e.g., to contemplate, to hypothesize). The clarity of their differentiation of these mental processes was examined in relation to their self-evaluations using the SRIS on their reflective tendencies and their insight levels. The results showed that people with high SRIS subscale scores (i.e., SRIS-SR or SRIS-IN) made clear distinctions among different mental processes consistently across three stages: the initial mental state (i.e., the common condition of a mental process), the mental process, and the final mental state (i.e., the common outcome of a mental process), whereas people with low SRIS subscale scores made much clearer distinctions about the conditions or the outcomes of processes than about the processes per se. The association between one’s actual conceptual understanding about different processes and one’s self-evaluation with the SRIS supports the validity of the SRIS as a self-report measurement scale.

As discussed in the introduction section, the two SRIS subscales, Engagement in Reflection and Need for Reflection, were combined during the revision of the SRIS. Correlational analyses were conducted with data compiled across the four experiments to examine the relationship between the subscales of the SRIS (see Table 2). The results showed that neither engagement in reflection nor need for reflection was strongly correlated with insight. But they appeared to be moderately correlated with each other. Despite the correlation between these two subscales, the findings of this study suggest that their relationship with other behavioral measures should after all be examined separately. That is, for the tasks used in the current study, introspective acts and introspective intents are two different aspects,
and intents seem to be more directly relevant to participants’ actual performance. Researchers have acknowledged the challenge to design a scale assessing constructive self-reflection while steering away from dysfunctional self-absorption (Anderson et al., 1996; Grant et al., 2002). The findings about self-reflection in this study may have some implications for this issue. First, the engagement in actual acts of reflection and the need to do so should be assessed separately. Lack of discrimination between the two may obscure the role of self-reflection in language and other cognitive functions. Second, the need for reflection seemed to be related to less reckless intrusive errors in memory tasks, but also with lower vocabulary proficiency. This may suggest that a strong need for reflection fares differently in short-term versus long-term processes. On one hand, it may be an indication of an individual’s effort to uphold the level of performance on an immediate task. On the other hand, it may also be a warning sign that obsession with the current mental states and processes could impede the pace of long-term progress. More research is needed to test these speculations, but researchers need to be aware that the need for reflection can be a complex factor in language or cognition in general, and that future assessment of the need for reflection may need to take into consideration different task contexts.

Table 1

<table>
<thead>
<tr>
<th>Group</th>
<th>PPVT score</th>
<th>Accuracy (%)</th>
<th>Reaction time (ms)</th>
<th>Hits</th>
<th>False alarms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low engagement</td>
<td>99 (13)</td>
<td>97 (4)</td>
<td>586 (100)</td>
<td>10.00 (2.66)</td>
</tr>
<tr>
<td></td>
<td>High engagement</td>
<td>99 (11)</td>
<td>96 (3)</td>
<td>576 (95)</td>
<td>9.25 (3.44)</td>
</tr>
<tr>
<td></td>
<td>Low need</td>
<td>101 (11)</td>
<td>97 (4)</td>
<td>608 (153)</td>
<td>11.00 (3.16)</td>
</tr>
<tr>
<td></td>
<td>High need</td>
<td>95 (14)</td>
<td>96 (3)</td>
<td>607 (125)</td>
<td>9.45 (2.77)</td>
</tr>
<tr>
<td></td>
<td>Low insight</td>
<td>98 (13)</td>
<td>98 (3)</td>
<td>570 (88)</td>
<td>9.71 (2.40)</td>
</tr>
<tr>
<td></td>
<td>High insight</td>
<td>99 (13)</td>
<td>96 (3)</td>
<td>637 (160)</td>
<td>10.33 (4.07)</td>
</tr>
</tbody>
</table>

Table 2

Correlation Coefficients of the Subscales of the SRIS (N = 312)

<table>
<thead>
<tr>
<th>SRIS subscales</th>
<th>Need for Reflection</th>
<th>Insight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engagement in Reflection</td>
<td>.61***</td>
<td>.23***</td>
</tr>
<tr>
<td>Need for Reflection</td>
<td>-</td>
<td>.19***</td>
</tr>
<tr>
<td>Insight</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

***p < .001.

The experiments reported here were some initial attempts to assess the relationship of different dispositions in self-reflection and insight with people’s performance on language tasks. One limitation of this study was that, like most research studies on self-reflection and insight, only college students were included. It is not clear whether and how self-reflection and insight may change over an individual's life span. Therefore, it is important for future research to include different age groups and to explore the developmental differences in these constructs. A second limitation was that the study was conducted in laboratory settings and utilized specially designed tasks. Future research will need to further explore the implications of individual differences in self-reflection and insight in the context of reading comprehension and language proficiency, especially in natural settings. For example, self-reflection and insight should be taken into account while exploring potential differences in individuals’ sensitivity to contextual information. That is, how do individuals differ in terms of making use of contextual information in reading? How do they differ in terms of making inferences while reading a text and retrieving information from memory about the previous text? Furthermore, how do they differ in terms of general reading strategies and styles? What are the connections between self-reflection, insight, and these individual differences in reading performance? The findings from this study suggest that the SRIS can be a promising tool, or at least a blueprint of a promising tool, to assess and determine these connections.
References


