DEVELOPMENT OF A QUESTIONNAIRE TO MEASURE RE-FLECTIVE ATTITUDE TOWARD CONVERSATION

SEONGSEOG PARK & BYEONGGON MIN

Department of Korean Language Education, Seoul National University, South Korea

Abstract

This study aimed to develop a valid questionnaire to measure upper secondary learners' Reflective Attitude toward Conversation. In the preliminary stage, 27 items were formulated under the three constructs we defined theoretically: Thoughtful Action, Content and Process Reflection, and Premise Reflection. In the exploratory stage, 467 high school students responded to the 27 items. The results of an exploratory factor analysis showed that 12 of the 27 items had good communalities (> 0.30) and good factor loadings (> 0.40). Further, the 3-factor-model showed a non-significant chi-square test result and a suitable TLI value (.979). In the confirmatory stage, 1,055 high school students responded to the 12 items finalized across two sessions. The results of the confirmatory factor analysis with the first confirmatory stage sample (568 students), showed that the NFI (.849) and CFI (.879) did not satisfy the necessary criteria (>.900). After revising two items, the results of the CFA with the second sample (487 students), revealed a CFI of .912, a GFI of .936, an RMR of .037, and an RMSEA of .076, which satisfied the necessary criteria, although the NFI of .897 still did not satisfy the necessary criterion. We recommend that future studies use the developed measure.

Keywords: reflective attitude toward conversation, thoughtful action, content reflection, process reflection, premise reflection, oracy education

Over the past three decades, reflection has been an important topic in comprehensive pedagogy research fields, including teacher education (e.g., Hatton & Smith, 1995; Korthagen, 1985), medical expert education (e.g., Kember, McKay, Sinclair, & Wong, 2008; Wald, Borkan, Taylor, Anthony, & Reis, 2012), adult learning (e.g., Dirkx, 2006), higher education (e.g., Rogers, 2001), service learning (e.g., Ash & Crayton, 2004), language or communication education (e.g., De Grez, Valcke, & Roozen, 2009; Lamy & Goodfellow, 1999; Lau, 2015) and so on. This growing interest in reflection mainly resulted from Schön (1983)'s argument that practice can be enhanced through the practitioner's reflection on his or her own practice. The inclusion of "reflection and self-development competencies" in the revised national curriculum of Korea as one of the core competencies of Korean language education in 2015, has sparked growing interest in learners' reflection and reflective education across the entire spectrum of the discipline of Korean language education. Researchers who have a main interest in listening and speaking education (oracy education) have thought that "turning experience into learning" (Boud, Keogh, & Walker, 1985, p. 7) is a suitable pedagogy for learning how to behave in a conversation, and, hence, we have to focus on learners' reflection on their experience of conversation. The inclusion of "reflection competency" in our national curriculum mentioned above was a powerful event that supported our ongoing ideas about reflective education for learning conversation.

However, if we are to examine and discuss learners' reflection, we must be able to measure the reflection objectively. Even though we can examine and discuss learners' reflection by our insights using some kind of qualitative method, it would then be hard to compare our results with the results from other research. It would also be difficult to generalize our results to other future pedagogies. For this reason, we judged that we have to develop a tool for measuring learners' reflection on (and in) conversation. Some kinds of tools have already been developed for comprehensive academic fields (e.g., Kember et al., 2000; Wald et al., 2012) or for other academic fields (e.g., Larrivee, 2008), but we could not find any tools that exactly correspond with our aim and academic field—a tool suitable for a pretest-posttest design in upper secondary listening and speaking education. Accordingly, we developed a set of self-reported questionnaire-items for measuring learners' reflection on (and in) their conversation. In this article, we report the entire process and results of developing our *Reflective Attitude toward Conversation (RAC)* questionnaire. We will first explain the definition and constructs of RAC.

Definition of RAC

Attempts to define the concept of reflection in education, at a general and comprehensive level, began with John Dewey. Dewey (1910) defines reflection (or reflective thought) as "[an] active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends" (p. 6). He further explains the process of reflective

thinking as one that starts with "a state of perplexity, hesitation, [and] doubt" and develops in the action of searching for facts or proof "to corroborate or to nullify the suggested belief" (p. 9). In pointing out that "[i]f the suggestion that occurs is at once accepted, we have uncritical thinking, the minimum of reflection" (p. 13), he emphasized that reflective thinking is capable of suspending judgment while maintaining a state of doubt in search of facts and proof that is essential for forming good thinking and the best suggestions for problem solving.

The concept of reflection has been elaborated by two notable educationists, Donald A. Schön and Jack Mezirow. Schön (1983) discusses reflection as a way of thinking to solve practical problems. He argues that the educator, who aims to enhance learners' ability to practice, has to focus on learners' reflection because practical contexts are disorderly and dynamic, and, accordingly, we cannot represent the know-how of problem solving acquired through each person's accumulation of experience in the generalized form of knowledge. Furthermore, Schön makes a distinction between *reflection in action* and *reflection on action* depending on the timing of reflection. He notes that the latter, that is, *post hoc* reflection on the action already taken in a perplexing situation encountered for the first time, can be done by both experienced specialists and unexperienced novices. But the former, that is, *in situ* reflection taking place to instantly find the best solution when faced with such a situation, can be carried out much better by specialists than by novices.

Mezirow (1990) expands the concept of reflection to include not only thinking for problem-solving, but also thinking for evaluating one's premises. First, he distinguishes non-reflective actions, namely *habitual action* and *thoughtful action without reflection*, from reflective ones. Thoughtful action without reflection typically accompanies only *introspection*, which Mezirow does not consider to be reflection —but, if reflection has arisen, introspection must have also arisen. Instead Mezirow argues that thoughtful action with *process reflection*, *content reflection*, or *premise reflection* can be considered reflective actions. Process or content reflection focuses on the how or what the problem solving is, while premise reflection focuses on evaluating whether the beliefs, assumptions, or values that one's reflection are based on are valid or not. Thus, while process or premise reflection is associated with thinking for problem solving, as Dewey and Schön have discussed, premise reflection is associated with thinking for evaluating one's belief system, and Mezirow contributes to the establishment of the concept of reflection especially at this point.

There is no doubt that our participation in conversation is a practice that is performed in disorderly and complex contexts, and we always have to reflect on our participation practices in conversation. This claim can be supported by various discussions in pragmatics and interpersonal communication theory. When a person participates in a conversation with a partner, he/she has to interpret the conversational implicature from the partner's utterance (Grice, 1975), and this kind of interpretation can be done through the logic of abduction (Hobbs, 2004; Hobbs, Stickel, Appelt, & Martin, 1993). While deduction or induction usually leads to only one answer, abduction—or inference to the best explanation—it does not offer only one answer

but instead allows one to choose the best alternative among many alternatives (Lipton, 2004). Thus, abductive interpretation in conversation always needs thoughtfulness to fully examine the evidence in the context and a reflective attitude to deliberate on whether the process, content, or premises of one's thinking are valid or not.

Furthermore, most of our everyday conversations can be considered as interpersonal communication, which is associated with establishing, maintaining, developing, or terminating the relationship between two persons. Because of its nature, we have to consider all of the inner dialectics from each participant's psychology and the outer dialectics from the relationship between two participants when we participate in a conversation. According to the dialectic theory of interpersonal communication (Baxter & Montgomery, 1996; see Guerrero, Anderson, & Afifi, 2018, pp. 130-134 for the review), communicators have inner dialectics, including *connection —autonomy*, *predictability—novelty*, and *openness—closedness*, and are faced with outer dialectics, including *inclusion—exclusion*, *conventionality—uniqueness*, and *revelation—concealment*. Within these dialectics, dialectical extremes act as antagonisms in which one rises and the other declines, and these antagonisms occur in an irregular pattern. Thus, it is hard to predict what dialectical extremes one will pursue even by one's own self, which depends entirely on who the partner is and under what contexts the conversation occurs. Therefore, there cannot be any exemplary answers to participate in an interpersonal communication conversation, and we must always ponder and reflect on how it is best to participate in each conversation.

From this point of view, we argue that the problem of engaging in conversation is a problem of disorderly and complex practice, and, therefore, as Schön claims, reflection must be involved. Furthermore, a conversation as the interpersonal communication involves a variety of psychological factors such as beliefs, assumptions, or values, so, in line with Mezirow's claims, we should not only reflect on the validity of problem solving but also reflect on the validity of our belief system. Following the approach of Allport (1935), which defined attitude as "[the] mental and neural state of readiness," (p. 810) we termed *Reflective Attitude toward Conversation (RAC)* to indicate the mental state of readiness to conduct the above reflections. Based on the general concepts of reflection suggested by Dewey, Schön, and Mezirow, we specifically define RAC as *an attitude to carefully and thoughtfully examine the contents and processes of one's own conversations, as well as one's beliefs, assumptions, or values about conversations that affects the examination of contents and processes, in the ongoing or terminated conversation situation*.

Constructs of RAC

Many scholars who have dealt with reflection have suggested a variety of classifications of reflection. For example, Hatton and Smith (1995), who conducted their research based on learners' reflective writings, suggested that reflection can be classified into four categories: *descriptive writing*, *descriptive reflection*, *dialogic reflection*, and *critical reflection*. Similarly, Jay and Johnson (2002) suggested *descriptive*

reflection, *comparative reflection*, and *critical reflection*. Their classifications may be associated with aspects of reflective writing, at least in the case of descriptive writing, but it does not seem to indicate reflection as the mental state itself. Recently, more scholars seem to follow the classification suggested by Mezirow (1990). For example, some scholars have argued that reflection can be divided into three categories, including *content reflection*, *process reflection*, and *premise reflection* (Kreber & Castleden, 2009; McAlpine, Weston, Berthiaume, Fairbank-Roch, & Owen, 2004). Furthermore, Kember et al. (2000, 2008) have suggested four categories, *habitual action*, *understanding*, *reflection*, and *critical reflection*, which include a non-reflective one (i.e., habitual action).

With our aim of establishing the psychometric constructs of RAC, we paid particular attention to the psychometric research of Kember et al. (2000). They adopted a psychometrical statistical method—i.e., factor analysis—to develop a self-reported questionnaire, which consists of four constructs and sixteen items for measuring students' degree of reflective thinking after finishing a university program for expert training. Among their four constructs, that is *habitual action*, *understanding*, *reflection*, and *critical reflection*, three constructs, excluding *understanding*, follow Mezirow (1990)'s categories directly. The *habitual action* is the same as Mezirow's *habitual action*. The *reflection* is the integration of *content reflection* and *process reflection* suggested by Mezirow. The *critical reflection* is synonymous for *premise reflection* by Mezirow. The only modification was done in *understanding*. This was originally intended to be Mezirow's *thoughtful action*, but was modified to meet their measurement objective given that reflection in college education begins with an understanding of professional knowledge.

Based on preceding research, we set our constructs of RAC to three: *Thoughtful Action*, *Content and Process Reflection*, and *Premise Reflection*. Since habitual action is not a reflective action and is in an inverse relationship with the other three constructs, it was considered when formulating reverse discrimination items for the measurement of the other construct (i.e., Thoughtful Action), but not set as a separate construct. The description of each of the three constructs we have set is as follows.

With respect to the first construct *Thoughtful Action*, Mezirow explains it as a pause to "[ask] what am I doing" (Mezirow, 1990, p. 6) and included introspection in it as an act of thinking about one's thoughts or feelings (Mezirow, 1991, p. 107). He further notes that although thoughtful action, when regarded separately, may not appear to be reflection because it only involves paying attention to what is being done based on what has already been learned, reflection and thoughtful action can converge in the process of problem solving like brooding over what to do next in a series of actions or how to understand a new experience (Mezirow, 1991, p. 107), as in playing chess, putting forward an argument, or using one's wit (Mezirow, 1990, p. 6). It is for this reason that Mezirow (1990) makes a distinction between *thoughtful action with reflection* and *thoughtful action without reflection*. This led many re-

searchers to assume that *thoughtful action* can be a sub-construct constituting reflective attitude, considering the character of a construct representing a comprehensive psychological manifestation that bundles various related components together. On the other hand, as can be confirmed in the scope of reflection presented by Mezirow (1990, p. 7), such thoughtful action comes out with reflection at the time of its implementation, that is, during action, not after the action. Considering this process, we defined Thoughtful Action as *a reflective attitude toward conversation, seeking to react carefully and thoughtfully after some introspection instead of reacting instantaneously when faced with a problem in the ongoing conversation, especially with respect to an emotionally charged problem*.

The second construct, *Content and Process Reflection*, was termed by merging the content and process reflections suggested by Mezirow (1990, 1991). While Mezirow mentions these two categories of reflection separately, he presents neither clear-cut definitions nor the differences between them. The only explanation is that the former focuses on the content (i.e., *what*) and the latter on the process (i.e., *how*) when the agent examines whether his/her action taken for problem solving was appropriate or whether the implementation method would be adequate for future use in a similar situation (Mezirow, 1990, p. 6; Mezirow, 1991, pp. 107–108). Reflecting on the content and process of a problem and problem solving are actually close-knit, and their boundaries are often difficult to pinpoint. It may be for this reason that Kember et al. (2000) merged them into a single construct, *reflection*. We found it also inevitable to integrate them into a single construct, making the merge explicit by naming it *Content and Process Reflection*. Considering that problems faced during a conversation is primarily associated with the practice, the construct CPR can be regarded as comprehensively referring to *reflection in action* (= *reflection in conversation* in this study) and *reflection on action* (= *reflection on conversation* in this study) suggested by Schön (1983). In fact, Mezirow notes that both content reflection and process reflection are involved in "thoughtful action with reflection" and "ex post facto reflection" (Mezirow, 1991, p. 108). In line with this, we defined the construct Content and Process Reflection as *a reflective attitude toward conversation, seeking to identify the optimal manner of conducting a conversation in the ongoing or terminated conversation situation and to examine the content of one's own expression especially the content involving emotions*.

Finally, the third construct, *Premise Reflection*, refers to reflection on one's own underlying beliefs, assumptions, and values, that is, the *why* for becoming aware of the presence of a problem prior to reflecting on the content of the problem or the problem-solving method. In this regard, Mezirow states that it is impossible to interpret our own experiences free of prejudice because we are bound to our own perspectives of meaning (Mezirow, 1990, p. 6) and highlights the necessity of premise reflection allowing us to raise questions about the justification of the premise underlying the problem posing or problematizing in the first place during the reflection process (Mezirow, 1990, p. 12). Thus, premise reflection includes posing questions about the patterns of our habitual expectations and meaning perspectives already

formed while interacting with the world, others, and ourselves (Mezirow, 1990, p. 12), which is indispensable for giving reflection-based problem solving the right orientation. Additionally, Mezirow states that it is difficult for premise reflection to become an integral component of a spontaneous action process (Mezirow, 1990, p. 13), presumably due to the need for in-depth deliberation, which cannot unfold instantaneously in situations requiring spontaneous action. Based on these arguments, we defined the construct Process Reflection as *a reflective attitude toward conversation, seeking to critically examine in the terminated conversation situation whether all of one's own beliefs, assumptions, or values that affect one's own reflection are desirable and whether they are not indiscriminatingly assimilated ones*.

METHOD

Procedure

Figure 1 presents the three stages of the RAC questionnaire development conducted in accordance with the general procedures for developing a psychometric instrument. In the preliminary stage, we defined RAC and extracted its constructs as the preliminary step for formulating items for RAC measurement. To this end, we reviewed existing theoretical discussions and arguments (e.g., Dewey, 1910; Mezirow, 1990, 1991; Rogers, 2001; Schön, 1983) and studies on self-report measurement of reflection (Kember et al., 2000; Larrivee, 2008). We then formulated preliminary items matching the operational definitions of RAC and its constructs, and subjected them to content validity examination by a panel of five specialists, holders of PhD degrees in listening and speaking education.

At first, preliminary questionnaire items were formulated under five RAC constructs: habitual action, thoughtful action, process reflection, content reflection, and critical reflection. However, the specialists pinpointed two issues that arose in setting constructs and formulating items based on them. First, habitual action and thoughtful action may be conceptually distinguishable, but the preliminary items formulated to measure their respective constructs seemed to be inversely related, and categorizing habitual action under reflection was considered inadequate. Second, while the constructs generally represented the differences in the RAC level, process reflection and content reflection seemed to represent only the difference in focus at the same level rather than the difference in level. Based on the specialists' opinions, we merged habitual action into thoughtful action, thereby moving the items formulated to measure the former to the latter as reverse discrimination items and fused the process reflection and content reflection constructs into one construct. Thus, we set *Thoughtful Action (TA)*, *Content and Process Reflection (CPR)*, and *Premise Reflection (PR)* as the final constructs and prepared 27 preliminary items.

In the exploratory stage, a preliminary survey was administered to 467 students in the 10th and 11th grades (mainly 16 and 17 years old) who were sampled from four high schools located in Seoul, consisting of one gifted education school, one

vocational school, and two general education schools, in October 2016. The main purpose of the preliminary survey was first to examine whether the preliminary items matched the 3-factor construct structure hypothesized in this study at an acceptable level, and second to identify items with proven validity from those that fail to meet the statistical criteria, for the 3-factor model. Exploratory factor analysis (EFA) performed on the results of the preliminary survey confirmed our 3-factor hypothesis, and we selected 12 items as final survey items out of the 27 original items assigned to the hypothesized 3-factor model.

Finally, in the confirmatory stage, we administered the main survey to 568 students in the 11th grade (mainly 17 years old) attending seven high schools located in Seoul (two gifted education schools, one vocational school, and four general education schools) who did not participate in the preliminary survey in December 2016. The confirmatory factor analysis (CFA) performed on the collected survey results was supposed to confirm whether the 12 main survey items have good validity. The original plan was to perform one main survey. However, the validity yielded in the CFA on the first main survey fell slightly short of the cut-off that we set. Therefore, we modified some of the items and re-administered the main survey in March 2017 to 487 students in the 10th grade (mainly 16 years old) attending four high schools (one gifted education school and three general education schools) that did not participate in the preliminary and first main surveys, and we reached the final confirmation.

Figure 1. Procedures of the research
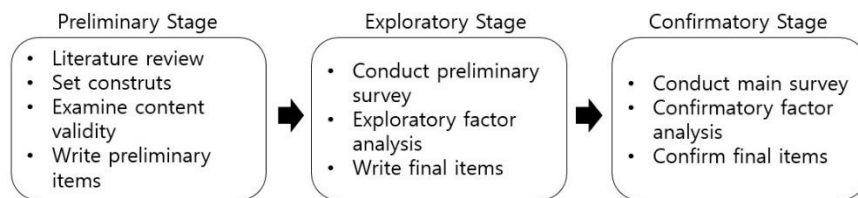
Data collection

Data were collected on three occasions: once in the exploratory stage (preliminary survey) and twice in the confirmatory stage (first and second main surveys). In the exploratory stage, there were 467 participants in the 10th and 11th grades (mainly 16 and 17 years old) attending gifted education, vocational, and general education schools (four schools in total). In the confirmatory stage, the first survey included 568 participants in the 11th grade (mainly 17 years old) attending gifted education, vocational, and general education schools (seven schools in total). In the second survey of the confirmatory stage, there were 487 participants in the 10th grade (mainly 16 years old) attending gifted education and general education schools (four schools

in total). In the three surveys conducted in this study, two schools (one gifted education and one general education) were involved two or more times, but care was taken to ensure that each student participated only once. The number of participants in each survey was determined in consideration of the minimum number (384) to satisfy ±5% margin of error and 95% confidence interval as specified by Backstrom and Hursh (1963). Table 1 presents a detailed overview of the participants in the surveys.

Table 1. Number of participants and schools

| | | School type | | | Total |
		Gifted	Vocational	General	
Preliminary	N of participants	98	157	212	467
	N of schools	1	1	2	4
Confirmatory	1st survey — N of participants	65	92	311	568
	1st survey — N of schools	2	1	4	7
	2nd survey — N of participants	35		452	487
	2nd survey — N of schools	2		3	4

Preliminary items

Table 2 presents the initial 27 preliminary items. These items were subjected to content validity examination, followed by construct adjustment and formulation modification. These are broken down into 9 TA items, 13 CPR items, and 4 PR items. The three TA items marked with "(-)" are the ones left among the original items pertaining to "habitual action" prior to the content validity examination and were moved to the TA construct without changing the formulations in order to be used as reverse discrimination items for the TA measurement items.

Data analysis

We performed factor analyses to verify the validity of the 3-factor RAC model. In accordance with the general process of factor analysis, an EFA was performed on the exploratory stage survey results, and a CFA was performed on the first and second survey results from the confirmatory stage.

The EFA was performed using SAS 9.4. We used the ML method to decompose the reduced correlation matrix. To use the ML method, the number of factors should be set, and the measurement variables should follow a multivariate normal distribution. Since this study hypothesizes three factors in the study design, the first criteria for using the ML method was satisfied. For the second criteria, the skewness and

kurtosis of the 27 items included in the exploratory factor analysis ranged from -1.041 to .853 and from -848 to .936, respectively, demonstrating normal distributions. These results allowed us to use the ML method.

Table 2. Preliminary items

Construct	no.	Item
Thoughtful Action (TA)	T1	I observe the interlocutor's behavior more carefully if I'm not sure of his intention.
	T2	I try to notice the topic change in conversation.
	T3	I usually cope with problems posed during conversation without reflecting too much. (-)
	T4	I think it is important to consider carefully how to express my thoughts.
	T5	I tend to be indifferent to others' reactions or reaction modes during conversation. (-)
	T6	I do not show my displeasure straightaway at others' unpleasant comments.
	T7	I don't think it is wrong to say in our daily lives whatever comes to my mind. (-)
	T8	I carefully choose my words considering their impact on others' feelings.
	T9	I try to talk more carefully when the conversation takes an unexpected or problematic turn.
Content and Process Reflection (CPR)	C1	When I have faced a problem while talking, I give thought to how to prevent it next time.
	C2	I try to discern whether my interlocutor has put forward a valid opinion.
	C3	After talking with someone, I check whether I have given precise information.
	C4	After clashing with someone, I try to find its cause by thinking back on the conversation.
	C5	After a conversation, I check whether the interlocutor expressed his or her true feelings.
	C6	After a conversation, I check whether the interlocutor stated their facts accurately.
	C7	To solve the problem I had while talking, I think back on similar situations I experienced.
	C8	When I speak with others, I tune myself to the flow of conversation or people's reactions.
	C9	After a conversation, I think back to check if I expressed my feelings honestly to others.
	C10	After a conversation, I check whether I put forward a valid opinion to my interlocutor.
	C11	I always attempt new methods to improve my conversation skills.
	C12	To solve problems encountered while talking, I ask people or consult books or the Internet.
	C13	Among various conversation methods, I try to find the optimal conversation method for me.

	P1	*I think about whether I have desirable conversation habits.*
	P2	*I ask myself whether I unconsciously imitate someone else's bad conversation habits.*
Premise Reflection (PR)	P3	*I compare others' views about desirable conversation to check my own views against.*
	P4	*I critically examine my assumptions to find out if they affect good conversation.*
	P5	*I guard myself against using bad buzzwords only because others are using them.*

Among various factor rotation types, we used promax rotation, which is typically used for oblique rotation. Hypothesizing a 90º angle between two axes in the orthogonal rotation is associated with a prediction that the inter-factor correlation coefficient corresponding to the cosine of that angle is zero. Although it seemed plausible to use an orthogonal rotation to simplify the results even on the assumption that the factors are correlated, as in many studies, we judged it implausible to apply the unrealistic hypothesis that the correlation between the RAC constructs is zero. It is necessary to consider that the sub-constructs pertaining to the process of reflection cannot be easily separated or distinguished from one another because they form a close-knit whole. For this reason, we decided to use an oblique rotation method, which hypothesizes the existence of inter-factor correlations.

To select items adequate for the hypothesized 3-factor RAC model, we conducted comprehensive monitoring to check whether the items had a minimum communality of .30, whether the factor loadings of the hypothesized constructs exceeded .40, and whether the factor loadings of the un-hypothesized constructs were lower than .30. Additionally, we compared the results of the chi-square test and the Tucker-Lewis index (TLI) to examine whether the 3-factor model had a better model fit than the adjacent 2-factor and 4-factor models.

The CFA was performed using LISREL 9.2. To be able to use structural equation modeling in LISREL, the measurement variables should have a normal multivariate distribution, as was the case with the ML method in the EFA. The skewness and kurtosis of the 12 items used in the first main survey ranged from -1.148 to .837 and from -.556 to 1.928, respectively, all demonstrating normal distributions. The skewness and kurtosis of the 12 items used in the second main survey ranged from -1.026 to .780 and from -.683 to 1.097, respectively, all demonstrating normal distributions. Among the overall model fit indices, we applied the root mean square error of approximation (RMSEA), normed fit index (NFI), goodness of fit index (GFI), comparative fit index (CFI), and root mean residual (RMR), which are known to have good model fit criteria, with the cut-off values indicating a good fit set at RMSEA < .08, NFI > .09, GFI > .09, CFI > .09, and RMR < .05.

RESULTS

Results of the exploratory factor analysis

Three items (C2, C5, and C6; see Table 1), which together form a single un-hypothe-sized construct in all analysis methods used for the process of EFA, were detected and were the first items eliminated. These are items on the judgment or monitoring of the counterpart's utterances instead of asking about the agent's own utterances. Then we eliminated seven items (in the order of T3, T6, T1, C3, T5, C12, and C4; see Table 1), starting with the smallest communality until no item had a communality lower than 0.30. Finally, we eliminated five items (in the order of C10, C1, P3, C7, and T2; see Table 1), starting with the one with the highest second factor loading among those that showed factor loadings in excess of 0.30 on two or more con-structs. After eliminating these five items, the second factor loadings of P4 and C8 appeared to be fairly high at .43 and .35, respectively. Item elimination was discon-tinued with these items because more elimination would have disturbed the factor structure. Table 3 outlines the factor structure of the final 12 selected items. Specif-ically, the following results were obtained: First, all four items hypothesized to per-tain to TA (T4, T8, T7, and T9) showed factor loadings in excess of .40 on the second factor. Second, all four CPR items (C11, C13, C9, and C8) showed factor loadings ex-ceeding .40 on the first factor. Third, all four PR items (P2, P5, P1, and P4) showed factor loadings exceeding .40 on the third factor.

   With these final 12 items, we performed chi-square tests and examined TLI to verify that the 3-factor model hypothesized in the paper had higher validity than its adjacent 2-factor and 4-factor models. The test results are outlined in Table 4. These results can be described as follows: In the case of the 2-factor-model, the chi-square test rejected the null hypothesis that "there are only two factors," and the alterna-tive hypothesis "there are two or more factors" was accepted ($p < .001$). The TLI (.899) also did not reach the cut-off criterion of .95. As for the 3-factor model, the null hypothesis that "there are three factors" could be accepted ($p = .08$), and its TLI (.979) also satisfied the cut-off criterion. In the case of the 4-factor model, the null hypothesis "there are four factors" could also be accepted ($p = .50$). However, the TLI (1.002) surpassed the maximum value 1.0, indicating that the model was exceed-ingly saturated. From these results, it could be concluded that the final 12 items se-lected from the 27 preliminary items can be best explained by the 3-factor model hypothesized in this study.

Table 3. Selected 12 items' factor loadings

Construct	no.	Factor loading		
		1st factor	2nd factor	3rd factor
Thoughtful Action (TA)	T4	.16	.60	.18
	T8	.26	.57	.14
	T7	-.11	.53	.15
	T9	.30	.45	.09
Content-and-Process Reflection (CPR)	C11	.71	.08	.09
	C13	.59	.15	.27
	C9	.48	.13	.30
	C8	.43	.35	.18
Premise Reflection (PR)	P2	.10	.18	.66
	P5	.19	.22	.49
	P1	.31	.21	.45
	P4	.43	.04	.44

Table 4. The final 12 items' Chi-square tests and TLIs

	2-factor model	3-factor model	4-factor model
Chi-square	$\chi^2 = 118.69$, $df = 43$, $p < .001$	$\chi^2 = 45.09$, $df = 33$, $p = .08$	$\chi^2 = 23.34$, $df = 24$, $p = .50$
TLI	.899	.979	1.002

Results of the confirmatory factor analysis

The first survey in our confirmatory stage was conducted with the 12 items selected based on the results of the EFA. The CFA of the first survey results yielded the factor structure illustrated in Figure 2. The factor loadings of the 12 items on each construct exceeded .40, except for one item (T7). However, T7's factor loading, with .37, was fairly high and could thus be deemed acceptable. Among the overall model fit indices, however, the NFI (= .849) and CFI (= .879) did not reach the cut-off criterion of .90, and we determined that the 12 items used in the first survey were unfit to be the final items for the RAC questionnaire.

Figure 2. Factor structure from the first confirmatory survey
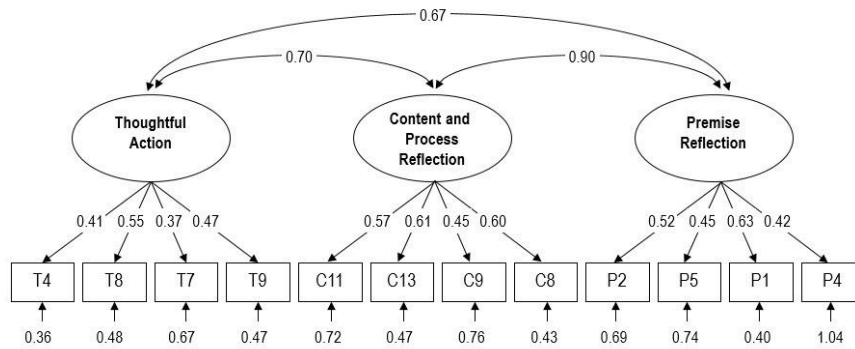
The reason for these rather low model fit indices was partly revealed by the reliability analysis performed on the first survey results. We found that the elimination of C9 improved the reliability of CPR from .67 to .69, and the elimination of P4 improved the reliability of PR from .59 to .62. We discussed this issue with the teachers who had cooperated during the first survey, who informed us that they were asked by some participants about the meanings of these two items. In the case of P4, for example, a student posed a question after misreading "assumptions" as "family" [these two words are homonyms in the Korean language].

Accordingly, we revised the questionnaire with the corresponding modifications: C9 to *"After a conversation, I think back to check if my thoughts and feelings were conveyed to others without distortion"* and P4 to *"I critically examine my beliefs or thoughts about good conversation to find out possible problems."* The revised questionnaire was administered to new students in new schools in the second survey, and the survey results were subjected to CFA. As a result, we obtained the factor structure illustrated in Figure 3. This time, 11 of the 12 items showed factor loadings exceeding .40, with T7 alone showing again .34 as in the first survey. All model fit indices except for the NFI satisfied the cut-off criteria, with NFI very close to the cut-off criterion with .897, as shown in Table 5.

From the above results, we could verify that the final 12 items (see Table 6) had sufficient validity to measure the three RAC constructs hypothesized in this study. Table 7 outlines the measurement results yielded through the final 12 items.

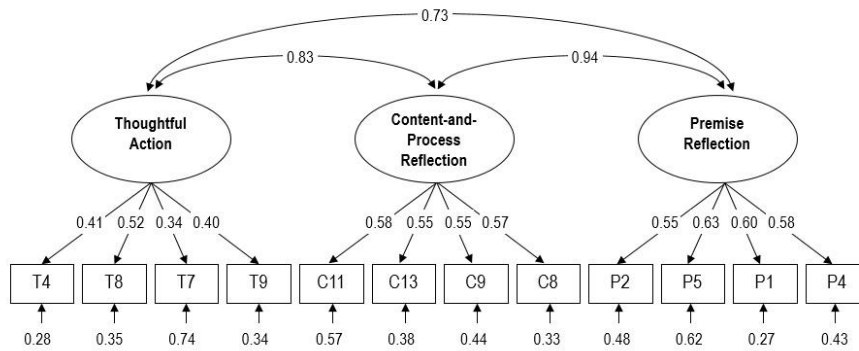*Figure 3. Factor structure from the second confirmatory survey*

Table 5. Model fit indices from the second confirmatory survey

	Chi-square	NFI	CFI	GFI	RMR	RMSEA
Criterion	None	> .90	> .90	> .90	< .05	< .08
This research	χ^2 = 194.40, df = 51, p < .001	.897	.921	.936	.037	.076

Table 6. Finally confirmed 12 items

Construct	no.	Item
Thoughtful Action (TA)	T4	*I think it is important to consider carefully how to express my thoughts.*
	T8	*I carefully choose my words considering their impact on others' feelings.*
	T7	*I don't think it is wrong to say in our daily lives whatever comes to my mind. (-)*
	T9	*I try to talk more carefully when the conversation takes an unexpected or problematic turn.*
Content-and-Process Reflection (CPR)	C11	*I always attempt new methods to improve my conversation skills.*
	C13	*Among various conversation methods, I try to find the optimal conversation method for me.*
	C9*	*After a conversation, I think back to check if my thoughts and feelings were conveyed to others without distortion.*
	C8	*When I speak with others, I tune myself to the flow of conversation or people's reactions.*
Premise Reflection (PR)	P2	*I ask myself whether I unconsciously imitate someone else's bad conversation habits.*
	P5	*I guard myself against using bad buzzwords only because others are using them.*
	P1	*I think about whether I have desirable conversation habits.*
	P4*	*I critically examine my beliefs or thoughts about good conversation to find out possible problems.*

Note. asterisked(*) items have been revised from those in the preliminary and 1st confirmatory surveys to those in the 2nd confirmatory survey

Table 7. Results of measurement by the final 12 items

| Construct | | Total | General | | | Gifted |
			Total	Male	Female	
Thoughtful Action (TA)	M	4.17	4.17	4.16	4.19	4.17
	SD	.54	.54	.56	.51	.54
Content-and-Process Reflection (CPR)	M	3.77	3.75	3.81	3.67	4.03
	SD	.65	.66	.67	.63	.57
Premise Reflection (PR)	M	3.65	3.63	3.63	3.64	3.86
	SD	.68	.68	.70	.66	.68

CONCLUSION

In this study, we defined *Reflective Attitude toward Conversation (RAC)* based on the transdisciplinary theoretical and empirical discussions we reviewed in the literature, and established three constructs to measure it. Based on the three constructs, we formulated 27 preliminary questionnaire items to measure RAC. From these preliminary items, we intended to select those that showed good validity. To this end, we administered one preliminary survey and two confirmatory surveys to a total of 1,522 students in the 10th and 11th grades. During the process of extracting the constructs, formulating the preliminary questions, and selecting the final 12 items through repeated questionnaire surveys and factor analysis, we could conclude that the final 12 items satisfy all applicable criteria. The following summarizes the concrete evidence for this conclusion.

The final 12 items showed construct validity to corroborate the hypothesized 3-factor model consisting of *Thoughtful Action (TA)*, *Content and Process Reflection (CPR)*, and *Premise Reflection (PR)*. Good construct validity was verified as follows. First, it was verified in an EFA of the adjacent factor models: The 2-factor model's null hypothesis "there are two factors" was rejected, but the 3-factor model's null hypothesis "there are three factors" was accepted. The 4-factor model resulted in an exceedingly saturated model, but the 3-factor model showed appropriate model fit. Second, the CFA also satisfied the cut-off criteria for almost all of the model fit indices.

The measurement values for the three constructs measured with the final 12 items were generally consistent with the views of Mezirow (1990, 1991) that content and process reflections are involved in a higher level of reflection than thoughtful action, and critical reflection is involved in a higher level than content and process reflections. TA yielded the highest mean value with 4.18, followed by CPR with 3.77 and PR with 3.65. As a result of a repeated measures ANOVA, the inter-construct differences were found to be statistically significant ($p < .001$), and contrast tests verified statistically significant differences between TA and CPR, TA and PR, and CPR

and PR ($p < .001$). These results imply that learners have difficulties attaining RAC in descending order from PR, CPR, and TA on average. Moreover, certain degrees of differences were observed between students of gifted education and general education students, who are generally considered to have different levels of academic achievements. While no differences in average score were observed in TA between gifted education and general education students (gifted $M = 4.17$, general $M = 4.17$), the former outscored the later in CPR (gifted $M = 4.03$, general $M = 3.76$) and PR (gifted $M = 3.86$, general $M = 3.63$). These results suggest that although the items developed in this study measure the Reflective Attitude toward Conversation, the level of reflective thinking during or after conversation influences the measurement results to some extent.

The RAC measurement items presented in this study are expected to be used for diverse purposes in qualitative and quantitative studies on reflective conversation education. For example, in order to explore differences in thinking characteristics between high RAC and low RAC learners using a qualitative approach, the questionnaire items developed in this study may serve as a basis for developing criteria for selecting high RAC or low RAC learners in the initial case selection stage. Additionally, if a researcher wishes to develop a reflective education design or syllabus for learning conversation, to experimentally apply it, and to check whether the new intervention is efficient, the researcher may use the questionnaire items developed in this study to find out the effect of the intervention by comparing the pre-intervention and post-intervention RAC scores. Besides these examples, the proposed questionnaire items have a variety of application potentials as a research instrument that can give momentum to diverse studies on reflective education for learning conversation. It is our hope that this study will lead to a large number of follow-up studies.

AUTHORS' NOTE

REFERENCES

Allport, G. W. (1935). Attitudes. In C. Murchison (Ed.), *A handbook of social psychology* (pp. 798–844). London, UK: Clark University Press.

Ash, S. L., & Clayton, P. H. (2004). The articulated learning: An approach to guided reflection and assessment. *Innovative Higher Education, 29*(2), 137-154.
https://doi.org/10.1023/B:IHIE.0000048795.84634.4a

Backstrom, C. H., & Hursh, G. D. (1963). *Survey research*. New York, NY: Northwestern University Press.

Bain, J. D., Ballantyne, R., Packer, J., & Mills, C. (1999). Using journal writing to enhance student teach-ers' reflectivity during field experience placements. *Teachers and Teaching, 5*(1), 51–73. https://doi.org/10.1080/1354060990050104

Baxter, L. A., & Montgomery, B. M. (1996). *Relating: Dialogues & dialectics*. New York, NY: Guilford.

Boud, D., Keogh, R., & Walker, D. (Eds.). (1985). *Reflection: Turning experience into learning*. New York, NY: Routledge Falmer.

De Grez, L., Valcke, M., & Roozen, I. (2009). The impact of goal orientation, self-reflection and personal characteristics on the acquisition of oral presentation skills. *European Journal of Psychology of Education, 24*(3), 293-306. https://doi.org/10.1007/BF03174762

Dewey, J. (1910). *How we think: A restatement of the relation of reflective thinking to the educative process.* Boston, MA: Heath. https://doi.org/10.1037/10903-000

Dirkx, J. M. (2006). Engaging emotions in adult learning: A Jungian perspective on emotion and trans-formative learning. *New Directions for Adult and Continuing Education,* 109, 15-26. https://doi.org/10.1002/ace.204

Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Eds.). *Syntax and semantics: Vol. 3. Speech acts.* (pp. 41-58). New York, NY: Academic Press.

Guerrero, L. K., Andersen, P. A., & Afifi, W. A. (2018). *Close encounters: Communication in relationships*. Thousand Oaks, CA: SAGE.

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education, 11*(1), 33–49. https://doi.org/10.1016/0742-051X(94)00012-U

Hobbs, J. R. (2004). Abduction in natural language understanding. In L. R. Horn, & G. Ward (Eds.), *The handbook of pragmatics* (pp. 724-741). Oxford, UK: Blackwell Publishing. https://doi.org/10.1002/9780470756959.ch32

Hobbs, J. R., Stickel, M. E., Appelt, D. E., & Martin, P. (1993). Interpretation as abduction. *Artificial Intelligence, 63*(1-2), 69-142. https://doi.org/10.1016/0004-3702(93)90015-4

Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education, 18*(1), 73–85. https://doi.org/10.1016/S0742-051X(01)00051-8

Kember, D., Leung, D. Y. P, Jones, A., Loke, A. Y., McKay, J., Sinclair, K., Tse, H., Webb, C., Wong, F. K. Y., Wong, M., & Yeung, E. (2000). Development of a questionnaire to measure the level of reflective thinking. *Assessment & Evaluation in Higher Education, 25*(4), 381–395. https://doi.org/10.1080/713611442

Kember, D., McKay, J., Sinclair, K., & Wong, F. K. Y. (2008). A four–category scheme for coding and as-sessing the level of reflection in written work. *Assessment & Evaluation in Higher Education, 33*(4), 369–379. https://doi.org/10.1080/02602930701293355

Korthagen, F. A. (1985). Reflective teaching and preservice teacher education in the Netherlands. *Journal of Teacher Education, 36*(5), 11-15. https://doi.org/10.1177/002248718503600502

Kreber, C., & Castleden, H. (2009). Reflection on teaching and epistemological structure: Reflective and critically reflective processes in "pure/soft" and "pure/hard" fields. *Higher Education, 57*(4), 509–531. https://doi.org/10.1007/s10734-008-9158-9

Lamy, M. N., & Goodfellow, R. (1999). 'Reflective Conversation' in the virtual language classroom. *Language Learning & Technology, 2*(2), 43-61.

Larrivee, B. (2008). Development of a tool to assess teachers' level of reflective practice. *Reflective Practice, 9*(3), 341–360. https://doi.org/10.1080/14623940802207451

Lau, K. (2015). Assessing reflection in English enhancement courses: Teachers' views and development of a holistic framework. *Assessment & Evaluation in Higher Education, 41*(6), 854–868. https://doi.org/10.1080/02602938.2015.1048424

Lipton, P. (2004). *Inference to the best explanation*. New York, NY: Loutledge. https://doi.org/10.4324/9780203470855

McAlpine, L., Weston, C., Berthiaume, D., Fairbank-Roch, G., & Owen, M. (2004). Reflection on teaching: Types and goals of reflection. *Educational Research and Evaluation, 10*(4–6), 337–363. https://doi.org/10.1080/13803610512331383489

Mezirow, J. (1990). How critical reflection triggers transformative learning. In J. Mezirow (Ed.), *Fostering critical reflection in adulthood* (pp. 1–20). San Francisco, CA: Jossey-Bass.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco, CA: Jossey-Bass. https://doi.org/10.1177/104515959100300116

Rogers, R. R. (2001). Reflection in higher education: A concept analysis. *Innovative Higher Education, 26*(1), 37–57. https://doi.org/10.1023/A:1010986404527

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. New York, NY: Basic Books.

Wald, H. S., Borkan, J. M., Taylor, J. S., Anthony, D., & Reis, S. P. (2012). Fostering and evaluating reflective capacity in medical education: Developing the REFLECT rubric for assessing reflective writing. *Academic Medicine, 87*(1), 41–50. https://doi.org/10.1097/ACM.0b013e31823b55fa