



OPEN ACCESS

EDITED BY

Gavin T. L. Brown,
The University of Auckland, New Zealand

REVIEWED BY

Ioannis Dimakos,
University of Patras, Greece
Damian J. Rivers,
Future University Hakodate, Japan

*CORRESPONDENCE

Larry J. Xethakis
✉ xethakis.larry.john.k@tokai.ac.jp

RECEIVED 30 March 2025

ACCEPTED 04 June 2025

PUBLISHED 20 June 2025

CITATION

Xethakis LJ, Rupp M, Plummer BRB and
Kawagoe T (2025) Development and initial
validation of the Interaction Anxiety in Group
Work (IAGWS) Scale. *Front. Educ.* 10:1602748.
doi: 10.3389/feduc.2025.1602748

COPYRIGHT

© 2025 Xethakis, Rupp, Plummer and
Kawagoe. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in
this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Development and initial validation of the Interaction Anxiety in Group Work (IAGWS) Scale

Larry J. Xethakis^{1*}, Michael Rupp¹, Brendan R. B. Plummer² and
Toshikazu Kawagoe¹

¹Department of Community and Social Studies, Tokai University, Kumamoto, Japan, ²Sojo
International Learning Center, Sojo University, Kumamoto, Japan

Introduction: Social interaction, especially in small groups, has become a widely used teaching methodology in the language classroom and can provide learners with a wide range of benefits. It can also place cognitive and affective demands on learners, provoking feelings of social anxiety, which can become an obstacle to learning. However, there is a lack of scales to measure social anxiety as it operates in small group work. The aim of this study was to develop and evaluate the psychometric properties of a short measure of interactional anxiety, the Interaction Anxiety in Group Work Scale (IAGWS), grounded in the self-presentation theory of social anxiety.

Methods: The study followed a four-phase development process: (1) item development followed by content validity assessment by experts and English learners; (2) item assessment and exploratory factor analysis to determine dimensionality; (3) the structure was tested using confirmatory analysis, and construct validity and reliability were determined; (4) concurrent validity was assessed through correlation analysis with related scales, and temporal reliability was measured.

Results: The final scale comprised 11 items with three dimensions: Becoming the Center of Attention (6 items), Working with New People (2 items), and Coping with Ambiguous Situations (3 items). The goodness-of-fit indices ($\chi^2 = 221.379$, $p < 0.001$; TLI = 0.966; CFI = 0.975; RMSEA = 0.077; SRMR = 0.028). AVE values ranged from 0.682 to 0.745. CR ranged from 0.865 to 0.946, omega from 0.871 to 0.953, and alpha from 0.871 to 0.954. Scores on the IAGWS exhibited significant, positive correlations with scores on each criterion instrument. ICC values for temporal reliability were all above 0.8.

Discussion: The IAGWS has been shown to be a reliable and validate measure of interaction anxiety. It can be used in research on the impact of social anxiety on group work, as well as the effectiveness of interventions aimed at reducing the detrimental influence of anxiety.

KEYWORDS

social anxiety, group work, interaction, language learning, psychometric, validity, reliability, emotion

Introduction

Group work and anxiety

Social interaction is at the heart of L2 use and the L2 learning process (Zhou, 2016), and learners working in pairs or groups has become the context for a major part of the learning that takes place in the L2 classroom (Fushino, 2010). Pair-work and group-work form the backbone of two of the most common teaching methodologies, Communicative Language Teaching and Task-based Language Teaching (Leeming, 2011), and the communicative activities employed in these approaches necessitate the use of pair- and group-work (Dörnyei and Murphey, 2004). When engaged in pair- and group-work activities, learners focus on language as it is used to communicate ideas, thoughts, and opinions, with an emphasis on practical skills rather than on the structure of the language. Using the L2 in such a manner offers a range of advantages for learners, such as exposing them to a variety of input while providing them with multiple opportunities for output (Zhou, 2016), encouraging the development of their language skills (Swain et al., 2002), as well as aiding the development of their communicative competence (Fushino, 2010). Furthermore, interacting with others in pairs or groups helps to develop trust between learners (Dörnyei and Murphey, 2004), allows them to notice common interests, and enhances active engagement in activities (Ito et al., 2022). Interacting with other learners, working together, and sharing information in pairs or groups can thus provide learners with a variety of language-related and social benefits.

Together with these benefits however, group-based learning approaches also place a number of demands upon learners (Cantwell and Andrews, 2002; Linnenbrink-Garcia et al., 2011). On an individual level, learners need communication skills to share ideas, express opinions and negotiate meaning (Fushino, 2010), which can be a challenge for many learners when trying to express themselves in the L2 (Kebłowska, 2012). The social dynamics for learners working in pairs or small groups are also more complex than in a teacher-centered environment. When working in small groups, learners are often asked to take part in new activities, with unfamiliar procedures and where the best course of action can be ambiguous, they often have to work with classmates they do not know, and in some activities, they may have to express their own opinions or disagree others' opinions, possibly opening them up to criticism or evaluation from others (Topham et al., 2016). These demands can create a feeling of unease or even situation-specific anxiety in some learners (Cantwell and Andrews, 2002), which can then become an obstacle to their learning (Zhou, 2016). Anxiety has also been shown to reduce learners' willingness to communicate (Liu and Jackson, 2008), as well as their degree of motivation (Yashima et al., 2009), and group work itself has been noted as a prominent factor behind learners' anxiety in some studies (Maher and King, 2022; Miura, 2019). Learners who experience feelings of anxiety are less willing to work in groups in the first place (Fushino, 2010; Miura, 2019), and when put in groups, are less likely to actively participate (MacIntyre and Gregersen, 2012). Realizing the benefits of pair- and group-work activities depends significantly on learners' active engagement and the quality of their interactions with the group (Molenaar et al., 2014; Webb, 2009),

and the presence of a single non-participatory group member can disrupt the efforts of the group and lower its effectiveness, resulting in outcomes that are worse than when learners work individually (Rhee et al., 2013). Anxiety can thus present a significant obstacle to the effectiveness of group-based language learning.

The negative impact of anxiety on language learning has been extensively studied (see Horwitz, 2017 for a review of important studies in this area). The primary focus of this research, however, has been on forms of situation-specific anxiety related to using a foreign language, or foreign language anxiety (FLA; e.g., Bekleyen, 2009; Cheng et al., 1999; Horwitz et al., 1986; Saito et al., 1999; Young, 1990), with an emphasis on anxiety related to speaking (e.g., Aida, 1994; Gregersen and Horwitz, 2002; King, 2013; Liu and Jackson, 2008; Yashima et al., 2016a). Recent work by King (e.g., King, 2013, 2014; King et al., 2020; Maher and King, 2020) among others (e.g., Miura, 2019; Yashima et al., 2016b; Zhou, 2016) has begun to examine the role and importance of interpersonal and social factors underlying learners' feelings of anxiety. Intriguingly, some of this research (King et al., 2020; Miura, 2019; Yashima et al., 2016b) has suggested that in addition to FLA, learners' concerns over interacting with others, or in other words, social and interactional anxiety, also underlie learners' experience of anxiety in the language classroom and when working in groups.

King et al. (2020) found that while improving the social atmosphere in the classroom produced increased interaction between students, this interaction took place in the L1, and there was no significant increase in L2 spoken production. This finding suggests that the improved classroom atmosphere eased learners' feelings of social anxiety, but not necessarily their unease over using the L2. Using student-led class discussions, Yashima et al. (2016b) were able to reduce learner silence and increase their oral L2 production. However, some learners also reported that feelings of anxiety were a greater factor inhibiting their willingness to take part in the discussion than were worries over language knowledge or competency. Yashima et al. (2016b) surmised that "the issue was more of affect than of English knowledge" (p. 123) for these learners, who even when able to express their opinions in L2, had difficulty in overcoming their feelings of unease. This result provides more evidence that social anxiety and concerns over interacting with other learners may be a prominent source of negative emotion in the language classroom.

Following up on these studies, Xethakis et al. (2024) explored learners' experience of anxiety when working in groups in the language classroom. Their aims were to uncover possible sources of this negative emotion by asking learners to describe an anxiety-provoking situation they encountered when working with other learners, and also to examine the relationship between learners' levels of small-group anxiety and the category of situations they reported. Adopting FLA (Horwitz et al., 1986) and the self-presentation theory of social anxiety (Schenkler and Leary, 1982) as conceptual frameworks, the study employed qualitative content analysis (Mayring, 2014; Schilling, 2006) to categorize learner responses based on the situation described. The analysis identified a range of situations, which were then grouped into two primary categories based on the underlying source of the anxiety. Unsurprisingly, situations related to communicating in the L2, in particular, conveying one's intended meaning and understanding others, formed one prominent category, with its source being

learners' feelings of FLA. The second, and even more prominent category (mentioned in 60% of responses) comprised situations related to interacting with other learners, with the most frequently mentioned situations being interacting with new people, expressing opinions, and uncomfortable silences, and with interaction anxiety (Leary, 1983a), a form of social anxiety, as their source. In addition, learners' level of small-group anxiety was significantly related to their reported source of anxiety, with learners who reported a high degree of anxiety twice as likely to describe an interaction-related situation than one related to L2 communication.

When taken in conjunction with the findings of King et al. (2020), Miura (2019) and Yashima et al. (2016b), these results suggest that interaction anxiety is an important factor underlying learners' experience of anxiety when working in small groups, and that therefore, this form of social anxiety likely constitutes a significant obstacle to the effectiveness of group-based language learning. The important role that interaction and small group work play in the language learning process, and the prominence of group-based learning activities in the classroom, would seem to imply the need for more research into the impact of social anxiety as it manifests in small group work and the need for a greater understanding of the role of interaction anxiety more specifically. There have been a number of studies on the impact of social anxiety on learners in their L1 (e.g., Russell and Shaw, 2009; Russell and Topham, 2012; Topham et al., 2016), and also the impact of anxiety on learners' attitudes toward group work (e.g., Cantwell and Andrews, 2002), but there has been little research focused on learners working in small groups in the language learning classroom (see Zhou, 2016 and Miura, 2019 for two exceptions, however). In addition to further qualitative and observational studies, there is a need for a valid and reliable instrument to assist both researchers and practitioners in the assessment of learners' self-perceptions of interaction anxiety specific when engaged in small group work for research in this area to proceed on an evidence-based foundation. This study represents an initial step in this direction by developing a context-specific measure of interaction anxiety and testing its validity and reliability.

Interaction anxiety

Working in a small group—where a learner's responses and actions often depend on what other learners do or say—is an example of a contingent social interaction. In situations of this type, e.g., a conversation, a chance meeting, or a group discussion, each participant's actions are in response to, or contingent upon, the other's (or others') actions (Leary and Kowalski, 1995). These situations differ from those, such as when giving a speech or performing on stage, where an individual's actions depend primarily on their own plans and intentions and are not dependent, or contingent, upon others' actions. Interaction anxiety, as a conceptually distinct form of social anxiety, was first outlined by Leary (1983a) as a situation-specific form of social anxiety which manifests in contingent social encounters with a theoretical basis in the self-presentation theory of social anxiety (Schenkler and Leary, 1982).

According to Schenkler and Leary (1982), an individual's actions in a social situation are guided by their impressions and evaluations of others, and conversely, others' impressions and evaluations guide their own actions. Because of this, most individuals attempt to present themselves in a way so as to create their desired impression in the situation at hand. Social anxiety is therefore characterized by concerns over interpersonal evaluation in actual or anticipated social interactions (Schenkler and Leary, 1982). More specifically, feelings of social anxiety arise when an individual is motivated to make a certain impression on others, or in other words has a high degree of impressional motivation, but is unsure of their ability to successfully make the desired impression, or has low impressional efficacy (Catalino et al., 2012). Simply put, feelings of anxiety come about in situations where an individual has a strong motivation to make a particular impression, and/or when the individual does not feel confident in their ability to make such an impression. Conversely, in situations where an individual has little motivation or high confidence in their ability, very little anxiety occurs.

The two antecedents of social anxiety, impressional motivation and impressional efficacy, are in turn influenced by both situational and dispositional factors. The first of these concerns the characteristics of the specific social setting and the people in it, for example, being placed in an unfamiliar situation, or in situations including explicit aspects of evaluation, such as an interview. Dispositional factors relate to an individual's psychological and physical characteristics, such as their sense of self-esteem, their intelligence, or their attractiveness (Leary and Kowalski, 1995). While both factors can influence the severity of feelings of unease, there are differences in the degree of influence that each factor has in triggering feelings of social anxiety. Several studies have shown that situational factors play a much larger role in provoking feelings of social anxiety. Nezelek and Leary (2002) found that 72% of the variance in participants' social anxiety was related to anxiety experienced in different daily-life situations, while Catalino et al. (2012) found an even higher proportion of the variance, 85–89%, in experienced social anxiety was due to situational differences, whereas only 11–15% was due to individual, or dispositional, differences. These results suggest that an individual's experience of social anxiety is more highly influenced by the situation at hand rather than their innate disposition. In the context of this study, this suggests that being placed in the social situation of small group work would be more likely to provoke feelings of anxiety than an individual's innate level of social anxiety.

There are a number of situational factors which can heighten an individual's motivation or adversely impact their sense of efficacy and thereby increase their sense of interaction anxiety (Leary and Kowalski, 1995). Becoming the center of attention, which arouses a greater sense of self-awareness and thus a greater concern with the kind of impression being made on others; the value of the hoped for outcome, i.e., social situations which may have important consequences engender higher motivation; and, first impressions, as these have a greater impact on individuals who are unfamiliar to us, all tend to raise an individual's impressional motivation. Impressional efficacy is primarily influenced by a feeling of uncertainty as to how to act to present the desired impression. Situations which involve a degree of uncertainty

include those which are ambiguous (situations where the rules of how to behave are not immediately obvious), those that are novel (situations which one has not experienced before), and those including unfamiliar others.

As noted above, small group work can present learners with a number of challenges: learners are often asked to take part in novel activities, with procedures that are unfamiliar and where the best course of action can be ambiguous; they frequently must work with classmates they do not know or are unfamiliar with; and in some activities, they have to express opinions or challenge others' opinions. Each of these is similar to a situational trigger of anxiety described by Leary and Kowalski (1995) which suggests that interaction anxiety is a particularly relevant form of negative emotion in the context of small group work.

For many learners, working together in pairs or small groups may not provoke feelings of unease greater than that expected when talking with new people or doing new activities, and in fact putting learners into this social situation has long been advocated as a means to reduce FLA (e.g., Young, 1991; Dörnyei and Murphey, 2004; King and Smith, 2017). However, for a sizeable number of learners being placed in a group can provoke situation-specific feelings of anxiety. Russell and Shaw (2009) found that 28% of learners experience anxiety when working in groups in their L1, and Russell and Topham (2012) found a similar level (26%). In the L2 context, working in groups was the second most cited source of social anxiety in a study Maher and King (2022), and an even higher proportion of learners, 34%, reported experiencing anxiety when doing group work in a study by Xethakis et al. (2024). These results strongly suggest that while working in groups can help ease some learners' feelings of unease, for others, it is a significant source of anxiety.

The central hypothesis of this study is that interaction anxiety—a situation-specific form of anxiety which arises in contingent social interactions—is a particularly salient form of anxiety for learners participating in small group work in English language classrooms, which is hereafter understood to mean a group of 3 to 6 learners engaged in tasks or activities involving L2 communication, and is a primary source of learners' unease and apprehension when placed in such situations. Consequently, for research in this area to progress there is a need for a valid and reliable measure of this negative emotion. However, as will be examined below, there is at present no suitable instrument to measure learners' experience of this form of anxiety in the classroom.

Measuring social and interaction anxiety in the context of small group work

The term social anxiety is often used as an overarching term, or a more general concept, which encompasses a number of forms of anxiety such as, shyness, stage fright, social phobia, and communication apprehension, in addition to interaction anxiety. As such, there exist a wide range of scales for measuring these differing forms of social anxiety, and thus there are a range of potential scales that might be employed to measure interaction anxiety as it manifests in the context of small group work in the

language classroom. Among measures that examine generalized feelings of anxiety arising from social interactions, and whose psychometric properties have been examined to some extent, four of the more prominent are the Social Avoidance and Distress Scale (SAD; Watson and Friend, 1969), the Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987), the Social Interaction Anxiety Scale (SIAS; Mattick and Clarke, 1998), and the Interaction Anxiousness Scale (IAS; Leary, 1983a). However, while each of these instruments are suitable for measuring more generalized feelings of social anxiety, they have a common shortcoming which could hinder their use in the context of group work, that is, a number of the items in each measure deal with situations that have very little to do with those that might be encountered in group work in the classroom, such as meeting an acquaintance on the street, attending a party, or speaking with authority figures.

While the LSAS does have items which deal directly with the context of group work, e.g., talking with people you don't know very well (Item 11) or participating in small groups (Item 2), it also contains items concerned with situations such as eating and drinking in front of others, or making returns to a store. As its name suggests, the IAS does focus specifically on interaction anxiety, i.e., the form of social anxiety most relevant to interacting in contingent social situations, however, as with the other measures, it also includes primarily items that concern out-of-class situations. As the content of these instruments are not specific to the classroom, and include items that refer to situations that are not encountered in the classroom, they may not accurately reflect a learner's level of anxiety when put into the context of small group work in the language classroom.

There are several scales measuring more generalized social anxiety intended specifically for use with children or adolescents, such as the Liebowitz Social Anxiety Scale for Children and Adolescents (Masia-Warner et al., 2003), the Social Anxiety Scale for Children-Revised (La Greca and Stone, 1993), and the Multidimensional Anxiety Scale for Children (March et al., 1997). While these instruments are focused on situations children encounter in an educational environment and as such have items concerned with working in groups and talking with unfamiliar children, they also contain items concerned with joining a club, writing on the board, being teased, and talking behind one's back, and thus may not prove to be valid measures of learners' anxiety when working small groups.

Instruments that are focused on the concept of communication apprehension might offer an alternative to these general social anxiety scales, and in fact, communication apprehension has served as the theoretical basis for the only other scale developed to examine social anxiety as it operates in the context of small group work (Fushino, 2006, 2010). The Personal Report of Communication Apprehension (PRCA-24; McCroskey and Richmond, 1992; McCroskey et al., 1985a) is the most commonly used scale in research on this phenomenon. Similar to the more general social anxiety scales discussed above, the PRCA-24 is not specifically focused on classroom situations, but it is aimed at capturing respondents' experience of anxiety in the context of four social situations: dyads, groups, meetings and public speaking. While the group and dyad subscales are relevant to this study, the two other subscales, the meeting and public speaking subscales,

focused on contexts related more closely to public speaking or audience anxiety rather than interaction anxiety. As of yet, there is no research on the validity of the group work and dyad subscales when used as stand-alone measures of communication apprehension, and furthermore, the structural validity of the PRCA has been questioned. In a study using responses from Japanese university students, Pribyl et al. (1998) found that the group work and meeting subscales did not clearly differentiate into separate factors, stating, “factor structures extracted from the data in this study suggest that Japanese students may not recognize the differences between the concepts of a meeting and a group discussion” (p. 51). Finally, the items in these subscales are concerned with rather general situations (e.g., I am tense and nervous while participating in group discussions), rather than more specific situations which might be addressed with teacher interventions.

The PRCA-24 has served as the basis for several other scales which were developed for use in more educational contexts. The first of these is the Class Apprehension about Participation Scale (CAPS; Neer, 1987). This scale was developed to measure CA as it manifests in the classroom. As such, the CAPS does contain items dealing with facets of group work, e.g., participation and non-participation, speaking up, and expressing opinions, however, it focuses primarily on whole-class activities and discussions rather than solely on small group work. The Japanese Communication Fear Scale (JCFS; Sakamoto et al., 1997) was developed specifically for use in the Japanese context in response to the issues with the PRCA-24. However, this scale is more focused on social relationships between interlocutors, e.g., familiar vs. unfamiliar other, speaking with older and younger individuals, as well as on situations not related to the classroom, such as the experience of communication apprehension in club activities or part-time jobs.

There is one measure which is specifically focused on anxiety experienced in group work situations in the classroom, the Communication Apprehension in Group Work scale (CAGW) by Fushino (2006, 2010). This scale was developed for use as part of a larger instrument aimed at examining relationships between communication competence, beliefs about group work, willingness to communicate, and communication apprehension in group work in the L2 classroom. The communication apprehension subscale was developed in reference to research on the content of the PRCA-24 (McCroskey and Richmond, 1992; McCroskey et al., 1985a,b) and its items are intended to measure learners' experience of CA when working in a group in the language classroom. The measure has shown good psychometric properties as part of the larger instrument (Fushino, 2006, 2010). However, two slightly different versions of the scale have been employed (Fushino, 2006, 2010), and when these versions were combined and tested for use as a stand-alone measure, the results suggested two dimensions underlie the scale, one concerned with communication apprehension and another concerned with a more general dislike of or negative attitude toward group work (Xethakis et al., 2024).

Finally, when considering the influence of anxiety in the context of small group work in the language learning classroom, the question of FLA should be addressed. This form of anxiety is a valid concern for all situations in the language classroom. However, the most commonly used measure of this form of anxiety, the Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz et al.,

1986), is focuses on anxiety as it operates in primarily whole-class situations, while, as noted above, several studies (Maher and King, 2022; Miura, 2019; Xethakis et al., 2024) have reported that small group work itself can be an influential source of learner anxiety in the foreign language classroom. In addition, while anxiety related to social situations forms a component of FLA (See Horwitz, 2017 for a discussion of the components of FLA), previous research has shown that that FLA and social anxiety can be differentiated (e.g., King, 2014; Xethakis et al., 2024).

Considering the importance of small group work in the language learning classroom and the detrimental impact of social anxiety on the effectiveness of group work, there is need for further research into interaction anxiety, as the form of social anxiety most relevant to the small group context, in order to better understand its operation, and also to provide educators with tools to lessen its impact, in order to improve the effectiveness of group-based language learning approaches. The lack of an adequate measure of interactional anxiety as it operates in the group work context hinders the advance of research in this area, and thus, the aim of this study was to develop a theoretically grounded and psychometrically valid measure of interactional anxiety as it operates in small group work. The resulting instrument, the Interaction Anxiety in Group Work Scale (IAGWS) may not only aid researchers in gaining better understanding of this important construct, but could also be a valuable tool to help teachers identify students who might feel uncomfortable working in small groups so that teachers can offer assistance or plan interventions beforehand to help make group work more effective.

Methods

The study adapted the process for scale development and validation outlined by Boateng et al. (2018) into a four-step process: (1) item development; (2) scale development; (3) scale evaluation; (4) assessment of concurrent validity and temporal reliability. These phases and the procedures followed in each phase are outlined in Figure 1.

Conceptual delineation

The initial step in scale development is to outline the conceptual definition underlying the instrument and its hypothesized dimensions (Boateng et al., 2018). In this study the concept of interaction anxiety in small group work is based on the self-presentation theory of social anxiety (Leary and Kowalski, 1995; Schenkler and Leary, 1982), and is conceptualized as the strength of feelings of unease or worry related to social interactions students may encounter when working in a small group of 3–6 students to complete an activity or accomplish a task in the classroom. This construct is also hypothesized to comprise three dimensions, which align with three situational antecedents of interaction anxiety as described above. The first of these is becoming the center of attention, which triggers a greater sense of self-awareness and thus a greater concern with the impressions being made on others. The second of these is working with new people, as impressions matter more with those who are unfamiliar to us, and thus it is

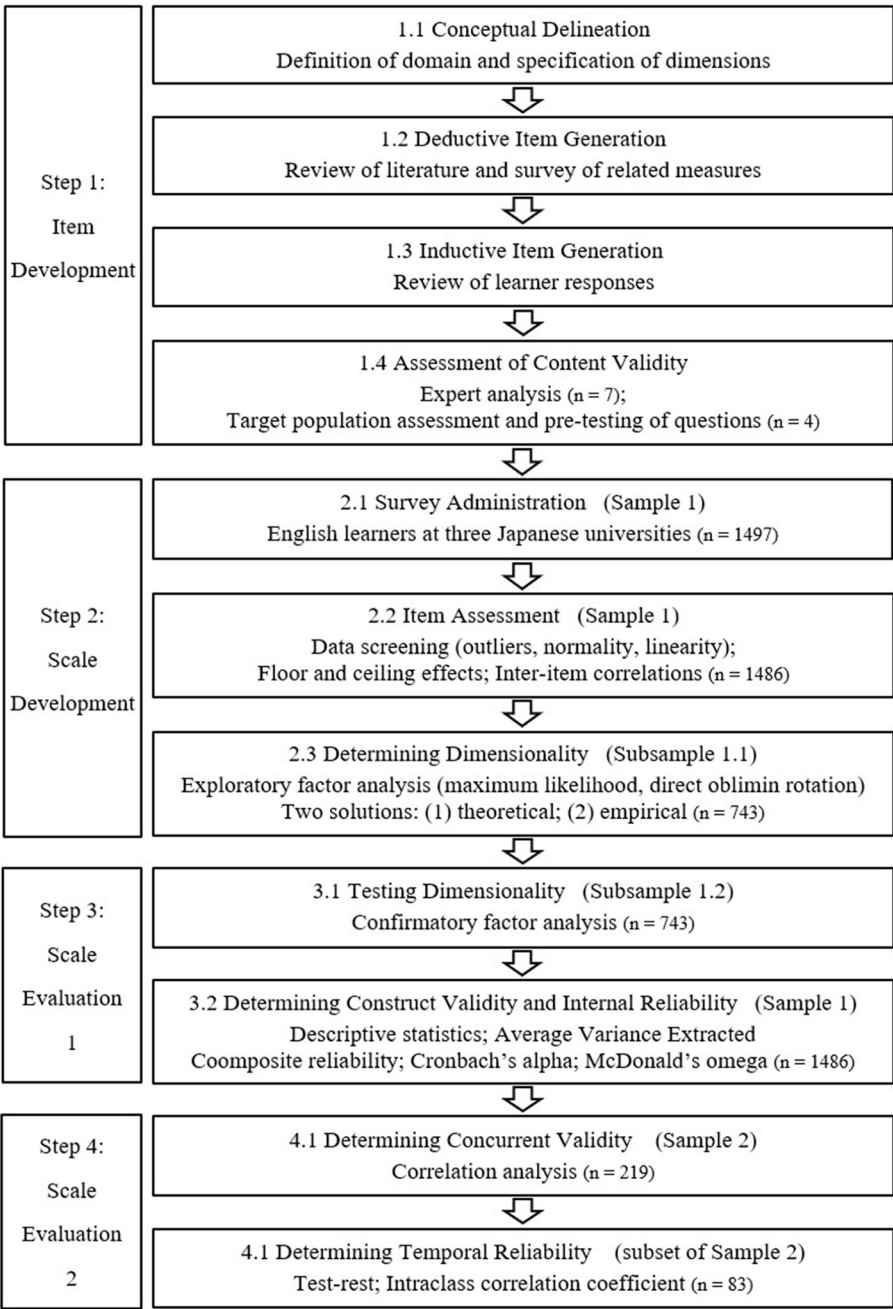


FIGURE 1
Scale development procedures (Boateng et al., 2018).

also more difficult to be sure of how to act to give them a positive impression. Coping with ambiguous situations, situations where the rules of how to behave are not immediately obvious, is the third hypothesized dimension.

Data collection

Data collection for this study comprised three phases. The first phase was related to the item development step in

Figure 1, and comprised four activities: deductive item generation, inductive item generation, assessment of content validity by experts ($n = 7$), and assessment and pre-testing of questions by target population ($n = 4$). The second phase involved the administration of an initial 15-item survey to undergraduate English language learners ($n = 1497$) in order to develop and evaluate the scale. The third phase was concerned with assessing the convergent validity and temporal reliability of the scale with data gathered from a second independent sample ($n = 219$) for this purpose.

Phase 1: item development

Deductive item generation (1.2 in [Figure 1](#)) was the first activity in this phase, and comprised a literature review and a review of existing scales measuring relevant constructs (e.g., social anxiety, communication apprehension, group work attitudes, etc.). A search for relevant studies and scales was undertaken on three databases: PubMed, Scopus and CiNii (a database of literature published primarily in Japanese journals). The search employed a number of keywords and combinations of these: group, group work, classroom, interaction, communication, attitudes, emotion, anxiety, apprehension, social anxiety, scale, and validity.

Inductive item generation (1.3) employed qualitative data from 407 learners who responded to the open-ended question, “*In as much detail as you can, write about an anxious learning experience you had in group-work, and how you felt about it*,” from a previous study on the situational antecedents of learner anxiety when working in small groups ([Xethakis et al., 2024](#)) to develop items based on learner responses.

These two processes provided a pool of prospective items, which were then subjected to content validity assessment (1.4). An initial assessment was carried out by seven experts, who had experience in survey development, survey validation and testing, research experience in the influence of affective factors in group work, or extensive (> 10 years) experience in communicative English teaching at the university level. It was felt that a combination of expert evaluators who were familiar with the technical aspects as well as those with practical classroom experience would provide a broad perspective on the validity of the prospective items.

A second form of content validity assessment was carried out in conjunction with the pre-testing of questions, using a focus group cognitive interview ($n = 4$). When used in the survey development process, this form of cognitive interview allows for dialogue between participants and the sharing of perspectives which can provide greater understanding of the context in question than individual interviews ([Farmer et al., 2022](#)). In this study, this process combined content evaluation with the pre-testing of questions phase. A form including all 47 items was given to four learners enrolled in English communication classes including small group work, who were then asked to describe their reactions and thought processes as they answered each item to provide feedback on item content, relevance, ease of understanding and to help define the construct from the learners’ point of view. This was used to evaluate participants’ understanding of the questions, their thought processes when answering the questions, and whether the questions were appropriate for the context. The focus group discussion was recorded and conducted with two of the authors serving as moderators, with one moderator primarily leading the group, while the other took notes. The two moderators reviewed the recording immediately after the focus group to confirm the content of the notes and uncover any further points of interest.

Phase 2: survey development

The data set employed in this phase of the development process (Sample 1) was gathered from 1497 university students at three universities, two private and one public, in a medium-sized city in south-western Japan. All participants were enrolled in English communication courses that included group work as a regular part of in-class activities. Respondents comprised 943 males (63.0%), 523 females (35.5%), 7 individuals who identified as other (0.5%), and 15 who did not answer (1.0%), with an average age of 18.7 years. The respondents comprised a convenience sample, however, with the combination of private or public universities and the preponderance of non-English majors, it was considered to be somewhat representative of the general population of tertiary English learners. Permission to conduct this study was obtained from administrators at each university after ethical review.

The measure employed in this phase of the development process comprised 15 items judged to be the most appropriate as a result of expert evaluation and target population feedback. The 15 items reflected the three hypothesized dimensions, with seven items concerning situations hypothesized to relate to becoming the center of attention (e.g., *I feel nervous if I am asked a question by other members.*), two items hypothesized to relate to working with new people (e.g., *I get nervous when I talk with classmates I don’t know.*), and six to ambiguous situations (e.g., *I feel nervous when everyone becomes silent.*). Participants were asked to respond to each item on a six-point Likert scale (from 1 = *strongly disagree*, to 6 = *strongly agree*). The survey form also included two demographic questions asking respondents their age and gender. Items which were not originally in Japanese, or for which there was no Japanese version, were translated following the International Test Commission’s guidelines for translating and adapting tests ([International Test Commission, 2017](#)). Translation was conducted by a specialist with experience in scale development, then back-translated into English by two bilingual English professors, with discrepancies resolved through consultation.

The survey was administered to learners using Google forms during the spring and fall semesters of the 2023 academic year, and participants took the survey during class. The informed consent of the participants was obtained by means of a statement at the beginning of the survey form informing participants that they need not take part in the survey, and that by answering the questions on the form they were giving their consent for their responses to be used in the study.

Phase 3: assessment of concurrent validity and temporal reliability

The data set employed in this phase (Sample 2) was gathered during the spring semester of the 2024 academic year from 219 learners at two universities, one private and one public, in a medium-sized city in south-western Japan, and was independent from Sample 1. All participants were enrolled in English communication courses. This sample comprised 84 female respondents (38.4%), 130 male respondents (59.4%), three who identified as other (1.4%), and two who did not provide a response

(0.9%), with an average age of 18.6 years. As with Sample 1, this sample comprised a convenience sample, though it also comprised primarily non-English majors and thus was considered somewhat representative of the general population of tertiary English learners. Permission to conduct this study was obtained from administrators at each university after ethical review.

The survey form administered in this phase comprised 37 items, including the 11 items of the IAGWS and four measures of differing forms of anxiety for validation (outlined below), with each item answered on a six-point Likert scale (from 1 = *strongly disagree*, to 6 = *strongly agree*). The survey form also included two demographic questions asking respondents their age and gender. Informed consent was obtained in the same way as with Sample 1, and participants took the survey during class.

Interaction anxiousness scale (IAS)

The IAS was developed by Leary (1983a) to measure interaction anxiety in a range of social situations as noted above. The Japanese version of the instrument (Okabayashi and Seiwa, 1992) used in this study comprises seven items and has shown good psychometric properties in a sample of tertiary EFL learners (Xethakis and Rupp, 2023).

Social phobia inventory, short form (Mini-SPIN)

Developed by Connor et al. (2001), the Mini-SPIN is a three-item measure of generalized social anxiety with good psychometric properties in adult (Weeks et al., 2007) and adolescent samples (Ranta et al., 2012). The Japanese items used in this study were taken from Otowa and Morita (2015).

Short fear of negative evaluation (SFNE)

Based on the original Fear of Negative Evaluation Scale (Watson and Friend, 1969), the SFNE was developed for use in the Japanese population by Sasagawa et al. (2004) and is comparable to the widely used Brief Fear of Negative Evaluation Scale (Leary, 1983b). Originally a 12-item scale, the eight-item version used in this study is based on the results of Nihei et al. (2018), who recommended removing the four negatively worded items from the scale. This version of the SFNE has shown good psychometric properties in the target population (Xethakis and Rupp, 2024).

Short form foreign language classroom anxiety scale (S-FLCAS)

Developed by MacIntyre (1992), this 8-item scale is a measure of the situational specific anxiety related to foreign language learning based on the FLCAS (Horwitz et al., 1986) and has been widely used in studies on positive and negative emotions in language learning (e.g., Dewaele and MacIntyre, 2014). Its psychometric properties were confirmed by Botes et al. (2022). The Japanese items were taken from Yashima et al. (2009).

Data analysis

Data from the expert content validity assessment stage was evaluated using the process laid out by Polit et al. (2007), where each item was rated on a scale of 1 (not relevant) to 4 (highly relevant) in terms of its relevance to the purpose, context and conceptual definition of the measure. Content validity was evaluated using the Item-Content Validity Index (I_CVI), where a value for each item

was calculated as the proportion of experts who rated the item as relevant or highly relevant, and a value for I_CVI > 0.78 as evidence of good content validity as suggested by Polit et al. (2007).

The subsequent stage of data analysis comprised two steps (Figure 1), item analysis and determining the dimensionality of the scale (i.e., the number of factors) using exploratory factor analysis (EFA). The data was initially screened for univariate (Z score > 3.29; Tabachnick and Fidell, 2019) and multivariate (Mahalanobis Distance/degrees of freedom > 4.5; Hair et al., 2019) outliers, and the normality and linearity of the scores were assessed. In the item analysis phase, the adjusted item-total correlation was used to examine the polyserial correlation of the items (Zijlmans et al., 2019), with correlations of > 0.30 considered acceptable (Field, 2018). Each item was then checked for floor and ceiling effects, which were considered significant if > 15% of scores on the item were 1 or 6, respectively (Terwee et al., 2007).

While the item analysis process was carried out using data from all 1486 valid responses in Sample 1, in order to determine the number of factors using EFA and test the dimensionality of the scale using confirmatory factor analysis (CFA), the sample was randomly split into two subsamples of equal size. The first of these, Subsample 1.1 ($n = 743$) was used in this phase to determine the number of factors using EFA, and the second subsample, Subsample 1.2 ($n = 743$), was used in the subsequent scale evaluation phase to test the dimensionality of the scale using CFA.

The suitability of the data for factor analysis was determined using the Kaiser-Meyer-Olkin measure of sampling adequacy (Field, 2018). A criterion of 0.4 was established for significant loading an item on a factor (Hair et al., 2019). In the case of items loading on two or more factors, the ratio of variance was calculated and items with a ratio < 2.0 were considered to have significant cross-loading and considered for removal (Hair et al., 2019). Two solutions were investigated using EFA: (1) a theoretically-based solution, adopting the hypothesized three-factors (outlined above), and employing maximum likelihood extraction with direct oblimin rotation; and (2) a data-driven solution, where the number of factors to extract was empirically derived using parallel analysis (carried out using JASP v0.17), together with examination of the scree plot and the eigenvalue (with > 1 being the criterion to stop further extraction), which also employed maximum likelihood extraction.

The models resulting from the EFA process were tested using CFA with data from subsample 1.2 ($n = 743$). Model fit was assessed using the chi-square statistic (χ^2), as well as χ^2/df (Wheaton et al., 1977), and four goodness of fit indices (Brown, 2015; Kline, 2023): the Tucker-Lewis Index (TLI > 0.94), Comparative Fit Index (CFI > 0.94), Standard Root Mean Square Residual (SRMR < 0.08), and Root Mean Square Error of Approximation (RMSEA < 0.07). Cut-off values are those proposed by Hair et al. (2019) based on model complexity and sample size. The analysis employed AMOS v28 with maximum likelihood estimation. Bootstrapping was applied to account for multivariate non-normality.

Following this, data from subsamples 1.1 and 1.2 were re-merged ($n = 1486$) and employed in determining the construct validity and internal reliability of the resultant scale. Means, standard deviations, medians, and interquartile ranges for scores on the scale as a whole and each sub-scale were calculated. Composite scores (scale score/number of items on

the scale) were used to facilitate comparison between scores on each scale. The internal reliability and construct validity of the IAGWS was then determined. The reliability of the IAGWS and its subscales was evaluated using four measures of reliability: McDonald's omega, Hancock and Mueller's (2001) Coefficient H (these values were determined using an Excel-based calculator developed by Brown, 2025), and composite reliability (CR), as well as Cronbach's alpha, with 95% confidence intervals. Values of >0.7 were considered sufficient in all cases (Hair et al., 2019). Cronbach's alpha is known to underestimate a scale's reliability when its underlying assumptions are violated, and because of this its use has been discouraged (McNeish, 2018). Alpha remains one of the most commonly reported indicators of reliability nonetheless (Kalkbrenner, 2023), and for this reason, we report alpha in addition to the alternate indicators of reliability. Correlations between the three subscales were calculated using Spearman's rho to determine the construct validity of the subscales. First, the convergent validity of the factors was estimated using the average variance extracted (AVE), with a benchmark of $AVE > 0.5$, as suggested by Hair et al. (2019). Next, their divergent validity was appraised by comparing the AVE to the maximum shared variance (MSV) and average shared variance (ASV) between subscales, and also by applying Fornell and Larcker's (1981) criterion of the square root of AVE for each subscale greater than the correlations between two subscales.

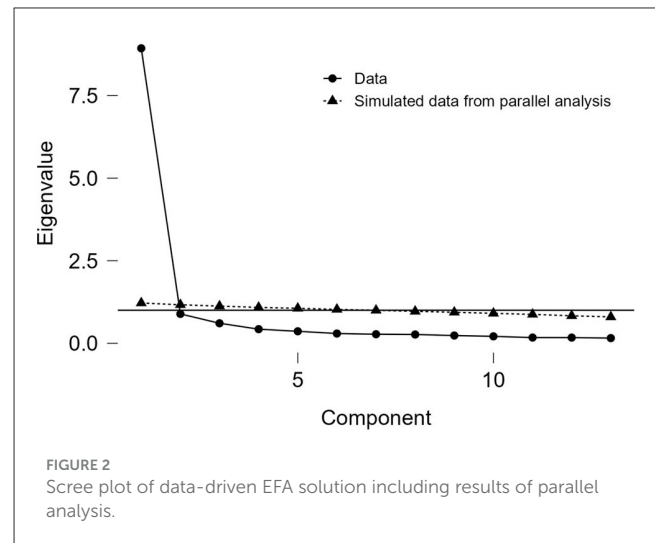
The concurrent validity of the IAGWS was assessed by correlation with scores on the IAS, Mini-SPIN, SFNE, and S-FLCAS using data from Sample 2 ($n = 219$). Spearman's rho was used for correlations due to the degree of non-normality in scores. Cronbach's alpha was calculated as a measure of internal reliability. Means and standard deviations for composite scores (scale score/number of items on the scale) were calculated to facilitate comparison between scores on each scale. To determine temporal reliability, the measure was administered after a two-week interval to a sub-sample of participants from Sample 2 ($n = 83$). The test-retest reliability was determined by intraclass correlation coefficient estimates and 95% confident intervals, based on an absolute-agreement, 2-way mixed-effects model (Koo and Li, 2016), using scores from a subset of Sample 2 ($n = 83$), with two administrations of the survey given over a two-week period.

Results

Step 1: item development

The literature search outlined in Section 2.2.1 resulted in a total of 1,263 hits, from which 102 relevant articles were selected based on a review of each article's abstracts. In addition, a review of existing surveys and reviews of psychological measures revealed an additional 44 papers or chapters which did not appear in the database search, for a total of 146 articles or chapters reviewed to find possible items for the new instrument. This process resulted in 135 items that were considered for inclusion in the initial version of the measure.

The review of learner responses on anxiety provoking experiences revealed a number of situations or contexts that were not included in other measures, and so, example or prototypical responses were selected and adapted for use as potential survey



items. In particular, these items were felt important as they allowed the measure to “capture lived experiences of phenomena by target population,” (Boateng et al., 2018, p. 6), and furthermore, while slightly adapted where necessary, the majority of the inductively generated items employed the respondents own words, and thus were straightforwardly worded, easy to understand, and were more conversational than some of the items uncovered in the review of existing instruments. This process yielded an additional 43 items.

As a result of the deductive and inductive item generation phases, a total of 178 items were included in the initial item pool. These items were then reviewed and rated by authors for relevance and congruence with the conceptual definition and context, which resulted in a reduction in the number of items to a final pool of 47 items which were evaluated for content validity by experts ($n = 7$) and members of the target population ($n = 4$).

Expert comments pointed out the similar content or phrasing of a number of items, as well as several items whose content was ambiguous or not specifically relevant for a typical language classroom and suggested alternatives. As a result of these comments and suggestions, several items were revised, and a number of the repetitive or overly similar items were removed from the survey. The items included in the 15-item measure employed in Phase 2 achieved an average I_CVI of 0.80, which was considered good (Polit et al., 2007).

Participant feedback from the focus group was used to rephrase several items and to remove several that were deemed not relevant or appropriate for the small group work context, as well as to select the most appropriate among similarly phrased items. As a result of the expert assessment and target population feedback the 47 items in the final item pool were reduced to the 15-item measure employed in Step 2.

Step 2: scale development

The data screening process revealed no univariate outliers, however, 11 multivariate outliers were found and removed from the analysis, leaving a sample size of 1,486. The normality of the data set was checked using one-sample Kolmogorov-Smirnov tests together with inspection of the Q-Q plots for each item, and the distribution

TABLE 1 Goodness-of-fit indicators for models evaluated with confirmatory factor analysis.

Model	χ^2	df	χ^2/df	TLI	CFI	RMSEA			SRMR
							95% CI		
Single-factor (empirically derived)	1,034.24*	65	15.9	0.866	0.888	0.142	0.134	0.149	0.0534
Three-factor (theoretically derived)	221.379*	41	5.4	0.966	0.975	0.077	0.067	0.087	0.0286

χ^2 = chi-square statistic; TLI, Tucker-Lewis Index; CFI, Comparative Fit Index; RMSEA, Root Mean Square Error of Approximation; SRMR, Standardized Root Mean Square Residual; CI, Confidence Interval.

* $p < 0.001$.

of all items was found to be non-normal. However, as the effects of non-normality in factor analysis can be reduced when employing sample sizes >200 , as in this study (Hair et al., 2019), and the kurtosis of each item was below the level which can be considered a serious departure from normality (>7 ; Byrne, 2016), it was considered appropriate to employ maximum likelihood estimation in the subsequent factor analysis process.

Following data screening, the item analysis process was carried out. The adjusted item-total correlations were found to be >0.3 for all items and all 15 items were retained. After this, each item was examined for floor and ceiling effects. None exhibited significant floor effects, however, two items, Item 2 (19.8%) and Item 6 (26.5%), were found to have significant ceiling effects, and so both were removed from further analysis, leaving a total of 13 items used in the EFAs conducted in the next step.

The dimensionality of the scale, that is, the number of factors, was determined by EFA using data from Subsample 1.1 ($n = 734$). The KMO value (0.958) indicated the appropriateness of the data set for use in factor analysis (Field, 2018). The hypothesized three-factor EFA solution was investigated first. The pattern of item-loading on the three-factors in this solution was predominantly as expected based on the content of each item, however, two items, Items 1 and 4, cross-loaded on two factors. When the ratio of the variances was calculated it was found to be 1.7 for Item 4 and 1.3 for Item 1 (both less than criterion of 2.0), as a result, both were removed, and the analysis re-run. This solution comprised 11 items loading on three factors and explained 82% of the variance. This solution was tested in CFA using Subsample 1.2 ($n = 734$) as described in the following section.

In the empirically derived solution, the results of parallel analysis, examination of the scree plot, and the eigenvalues all indicated a single factor solution (Figure 2). The loading of all 13 items was >0.6 , with no significant cross-loading. This solution explained 68.7% of the variance and was also tested using CFA as described below.

Step 3: scale evaluation 1

The CFA process was carried out with data from subsample 1.2 ($n = 743$). The KMO value (0.955) indicated the appropriateness of this data set for use in factor analysis (Field, 2018). The validity of both the empirically derived one-factor solution and the theoretically based three-factor solution were tested using CFA, with the results shown in Table 1. The values for χ^2 , and all four goodness-of-fit indices for the one-factor model unambiguously indicated unanimous poor-fit. The values of the three-factor model

TABLE 2 Factor loadings of the Interaction Anxiety in Group Work Scale (IAGWS) and subscales.

	IAGWS item	Factor		
		1	2	3
	Factor 1: becoming the center of attention (BCA)			
1	I get nervous when asked about myself.	0.898		
2	I get nervous when I talk about myself.	0.886		
3	I feel nervous if I am asked a question by other members.	0.884		
4	I get nervous when asked my opinion.	0.871		
5	I feel nervous when I express my opinion.	0.844		
6	I can't ask questions to other members because I get nervous.	0.790		
	Factor 2: working with new people (WNP)			
7	I get nervous when I talk with classmates I don't know.		0.948	
8	Working with classmates I don't know makes me feel anxious.		0.840	
	Factor 3: coping with ambiguous situations (CAS)			
9	I feel anxious when it is difficult to start the conversation.			0.887
10	I feel nervous when everyone becomes silent.			0.797
11	I feel anxious when conversation breaks off.			0.789

on the other hand suggested a fair degree of fit. Both the TLI and the CFI were above 0.96, and the SRMR was far below 0.08. The RMSEA did exceed the criterion of 0.07 for models with less than 12 indicators and a sample size of > 250 (Hair et al., 2019), however, the value of 0.077 was less than the value of 0.08 which is commonly regarded as indicating a fair degree of fit (Brown, 2015), and, in addition, the upper bound of the 95% confidence interval was <0.1 , the point at which the model should be rejected (Kline, 2023). It should be noted that the chi-square value was significant, however, this is to be expected with large sample sizes. The value for the χ^2/df for this model was 5.4, a significant improvement over the value for the one-factor model (15.9). Therefore, on the basis of a holistic

TABLE 3 Descriptive statistics and reliability estimates for the IAGWS and subscales ($n = 1,486$).

Scale/subscale	M	SD	Med	IQR	CR	H	Ω	α	95% CI
IAGWS	3.51	1.20	3.54	2.73–4.36	0.969	0.973	0.953	0.954	0.950–0.957
BCA	3.45	1.28	3.50	2.40–4.33	0.946	0.949	0.946	0.945	0.941–0.950
WNP	3.55	1.45	3.50	2.50–4.50	0.890	0.918	—	0.887	0.874–0.898
CAS	3.61	1.27	3.67	2.67–4.42	0.865	0.876	0.871	0.871	0.859–0.882

IAGWS, Interaction Anxiety in Group Work Scale; BCA, Becoming the Center of Attention; WNP, Working with New People; CAS, Coping with Ambiguous Situations; CI, Confidence Interval. Means, standard deviations, medians (Med) and interquartile ranges (IQR) are for composite scores (scale score/number of items on the scale) to facilitate comparison between scores.

TABLE 4 Convergent and discriminant validity of the IAGWS subscales.

Scale	1	2	3	AVE	MSV	ASV
1. BCA	(0.863)			0.745	0.728	0.702
2. WNP	0.853**	(0.896)		0.802	0.723	0.693
3. CAS	0.822**	0.812**	(0.826)	0.682	0.676	0.668

IAGWS, Interaction Anxiety in Group Work Scale; BCA, Becoming the Center of Attention; WNP, Working with New People; CAS, Coping with Ambiguous Situations; AVE, Average Variance Extracted; MSV, Maximum Shared Variance; ASV, Average Shared Variance; Values in brackets on the diagonal are square root of AVE.

Spearman's rho; ** $p < 0.01$.

view of all the indicators of model fit, the model was deemed to have a fair to good degree of fit, and the three-factor model was adopted.

In line with the conceptual delineation outlined above, the instrument was termed the Interaction Anxiety in Group Work Scale (IAGWS), while the three subscales were named Becoming the Center of Attention (BCA) with 6 items, Working with New People (WNP) with 2 items, and Coping with Ambiguous Situations (CAS), with 3 items. The loadings of the items on their respective factors are shown in Table 2. It should be noted that while a two-item factor deviates from the commonly recommended practice of having no fewer than three indicators per factor (e.g., Hair et al., 2019), Bollen (1989) demonstrated that a two-item factor specified in a multi-construct model can be sufficient to satisfy the conditions for model identification.

Means and standard deviations for the scale as a whole and each sub-scale are shown in Table 3 together with the values for the four estimates of reliability, which all indicated a sufficient degree of reliability. The CR value for the scale as a whole was, and values for the three subscales ranged from 0.865 (CAS) to 0.946 (BCA). The value of Coefficient H was 0.973 for the scale, and ranged from 0.876 (CAS) to 0.949 (BCA). Notably, the value for the two-item DNP subscale was 0.918, suggesting that, although a two-item factor may in some cases engender reliability issues, this is not the case for the DNP in this sample. Values for McDonald's omega were 0.953 for the IAGWS, 0.946 for the BCA subscale and 0.871 for the CAS. McDonald's omega for the WNP could not be calculated due to there being two items on this subscale. Finally, Cronbach's alpha for the whole scale was 0.954, while the values for the subscales ranged from 0.871 (CAS) to 0.945 (BCA).

The AVE values calculated for each subscale were >0.5 , and so each was considered to have sufficient convergent validity (Table 4). The AVE values for each subscale were greater than the MSV and ASV values between subscales, and the square root of AVE for each factor (values in brackets on the diagonal) was greater than

the correlations between factors, which confirmed the discriminant validity of the subscales.

Step 4: assessment of concurrent validity and temporal reliability

Data from Sample 2 ($n = 219$) was employed in this step. The results of the correlation analysis conducted to assess concurrent validity are shown in Table 5. All five scales displayed a high degree of reliability, and the IAGWS was positively and significantly related to the four validation measures, as expected. The strongest association ($r = 0.777$) was between scores on the IAGWS and those on the IAS, a measure of the interaction anxiety construct. A lesser degree of association ($r = 0.645$) was found between the IAGWS and the Mini-SPIN, and even more diminished correlations with a measure of interpersonal evaluative anxiety, the SFNE ($r = 0.528$) and foreign language anxiety, the S-FLCAS ($r = 0.572$). The concurrent validity of the scale is indicated by the direction and relative strength of the correlations between the IAGWS and the other scales. The values for the intraclass correlation coefficient for test-retest reliability (Table 6) ranged from 0.851 (95% CI:0.770–0.904) for the DNP subscale to 0.899 (95% CI:0.844–0.935) for the IAGWS as whole, indicating that the measure possessed good to excellent temporal reliability (Koo and Li, 2016).

General discussion

Social interaction in small groups has become one of the most widely used teaching methodologies in the language learning classroom. Interacting in groups can provide learners with comprehensible input, more opportunities for output and the chance to build relationships with other learners by sharing personal information in a less threatening context than whole-class activities. However, for some learners interacting with others can be a source of anxiety which causes these learners to withdraw from the group, and this impedes both the learning of the individual and the other group members. Concerns over interacting with others, or interaction anxiety, can thus be seen as an obstacle to the effectiveness of a powerful language learning tool. However, there has been no valid and reliable context-specific measure of anxiety which would allow for further research in this area, and which could also serve as a valuable tool for educators to plan interventions to limit the impact of anxiety. The primary aim of this study was

TABLE 5 Correlation analysis of the IAGWS, IAS, mini-SPIN, SFNE, and S-FLCAS.

Scale	M	SD	H	Ω	α	1	2	3	4	5
1. IAGWS	3.22	1.26	0.959	0.952	0.953	—				
2. IAS	3.31	1.23	0.903	0.897	0.898	0.777**	—			
3. Mini-SPIN	3.70	1.32	0.825	0.815	0.813	0.645**	0.725**	—		
4. SFNE	3.54	1.38	0.952	0.951	0.951	0.528**	0.592**	0.656**	—	
5. S-FLCAS	3.89	0.971	0.864	0.831	0.826	0.572**	0.568**	0.526**	0.526**	—

IAGWS, Interaction Anxiety in Group Work Scale; IAS, Interaction Anxiousness Scale; Mini-SPIN, Social Phobia Inventory, Short Form; SFNE, Short Fear of Negative Evaluation; S-FLCAS, Short Form Foreign Language Classroom Anxiety Scale. Means and standard deviations are for composite scores (scale score/number of items on the scale) to facilitate comparison between scores.

Spearman's rho; ** $p < 0.01$.

TABLE 6 Descriptive statistics, reliability, and ICCs for the IAGWS and its subscales.

Scale	Time 1			Time 2			ICC	95% CI
	M	SD	α	M	SD	α		
IAGWS	3.13	1.27	0.968	3.08	1.36	0.959	0.899**	0.844–0.935
BCA	3.04	1.30	0.961	2.99	1.42	0.943	0.866**	0.823–0.926
WNP	3.12	1.46	0.935	3.06	1.50	0.899	0.851**	0.770–0.904
CAS	3.31	1.43	0.902	3.26	1.46	0.885	0.880**	0.814–0.922

IAGWS, Interaction Anxiety in Group Work Scale; BCA, Becoming the Center of Attention; WNP, Working with New People; CAS, Coping with Ambiguous Situations; CI, Confidence Interval. Means and standard deviations are for composite scores (scale score/number of items on the scale) to facilitate comparison between scores.

** $p < 0.001$.

to develop and validate an instrument to measure interaction anxiety as it manifests in small group work in the language classroom, grounded in the self-presentation theory of social anxiety (Leary and Kowalski, 1995; Schenkler and Leary, 1982), and conceptualized as feelings of unease related to social interactions encountered when working in a small group in the classroom.

The scale development process reported in this paper resulted in an 11-item scale, the IAGWS, comprising three dimensions related to situational factors underlying interaction anxiety: concerns over becoming the center of attention, unease at working with unfamiliar or new people, and worries over coping with ambiguous situations. The three-factor model of the IAGWS exhibited a satisfactory degree of construct validity, with values for all four of the goodness-of-fit indicators in the good to acceptable range. The subscales displayed sufficient convergent and divergent validity, with AVE values for each above the criterion of 0.5, and also greater than MSV and ASV values (Hair et al., 2019). The subscale's discriminant validity was also indicated by their performance on the Fornell and Larcker (1981) AVE test, where the square root of each subscale's AVE value exceeded the correlations between factors. The IAGWS also performed well in terms of its reliability. The measure exhibited good internal reliability, with CR values for the scale as a whole and each subscale above 0.7. Similarly, values for Coefficient H, McDonald's omega, and Cronbach's alpha were all above 0.8, with the exception of the WNP subscale (where omega could not be calculated due to the subscale comprising two items). As for temporal reliability, the ICC values for the IAGWS and its three subscales were all above 0.8, indicating good test-retest reliability (Koo and Li, 2016).

The concurrent validity of the IAGWS was shown by its correlations with the four anxiety instruments used as validation measures. Scores on the IAGWS exhibited significant, positive correlations with scores on each instrument. The trends in correlations are in line with expectations. As the IAS is a measure of the same construct, it is to be expected that correlations would be stronger than those for the Mini-SPIN and other measures. Concerns over interpersonal evaluation underlie most forms of social anxiety (Leary and Kowalski, 1995), and thus the degree of correlation between the SFNE and IAGWS is as to be expected. In addition, considering the context of the study, and the fact that, as Horwitz (2017) points out, one of the three analogs of FLA is communication apprehension, the degree of association between the IAGWS and the S-FLCAS should not be surprising. Furthermore, the values in this study were similar to those found by Leary and Kowalski (1993) in testing the construct validity of the IAS, which again suggests that the IAGWS is a valid measure of interaction anxiety in the context it was designed for, small group work in the language classroom.

Analysis of the IAGWS's psychometric properties has provided strong evidence of the validity and reliability of scores produced by the instrument among university English learners in Japan. It can thereby serve as a useful, evidence-based tool in acquiring a greater understanding of this construct and expanding research into the impact of interaction anxiety on small-group language learning. For example, the IAGWS would be valuable in aiding the investigation of the effect of different teaching methodologies on learners' experience of interaction anxiety. In addition, the IAGWS could serve as a tool for teachers' pedagogical interventions. Firstly, the IAGWS can help identify students who might feel uncomfortable working in small groups, information which can inform lesson

planning and task design. Teachers might, for example, add in activities or entire lessons that focus on conversation strategies to help learners deal with the ambiguity that results from communication breakdowns (Barrington, 2021), or use simple icebreaker activities which focus on building rapport between learners to reduce unease over interacting with unknown others (Xethakis, 2024). Secondly, the IAGWS could be used by teachers engaged in action research on the effectiveness of group work in their own classrooms. In its simplest form, this could, for example, involve measuring learners' feelings of anxiety before and after a planned intervention to understand the effects of the intervention. More elaborately, a mixed-methods study could use scores on the IAGWS in combination with qualitative data collection methods (e.g., open-ended questions, interviews, etc.) to gain greater understanding of the issues their learners face when working in groups and the severity thereof, as well as suggest means to alleviate them.

Limitations

While the development process followed in this study resulted in a valid and reliable instrument, there are a number of limitations as well. First, it must be noted that the validation process for any instrument is an ongoing process, and this study represents only the initial stage in the development of the IAGWS. More research is needed to provide further evidence for the validity and reliability of the measure using independent samples. The next limitation concerns the data sets used in the steps of the development process. While the data sets were gathered from both public and private universities and respondents came from a wide range of departments and faculties, which allows for a degree of generalization, there is a need to further validate the scale with more samples in this population. In addition, while the measure displayed validity and reliability in a sample of Japanese university English learners, more research with samples from other countries in Asia, as well as other areas, is needed to confirm the cross-cultural validity and reliability of the scale as well as measurement invariance across different populations. Third, while the structural validity of the three-factor model was confirmed in this study, one of the three factors, WNP, comprises only two indicators, which goes against the commonly suggested guidelines of at least three indicators per factor (Hair et al., 2019). Future research should further examine the situational antecedents of learner anxiety when working or speaking with unfamiliar others in order to add range to this operational construct as well as to uncover other possible situational triggers of interaction anxiety in the small group context. Such research could incorporate in-depth approaches, such as interviews or focus group techniques to explore learners' experience of anxiety and further validate the content of the IAGWS. Additionally, studies investigating learners' physiological responses during group work could offer valuable insights, as well as possibly alternative perspectives. Employing a number of diverse techniques could help to address the limitations of self-report methods, such as, the impact of cultural norms and social desirability and on responses (Heppner et al., 2008), and an underlying assumption that respondents are able to accurately recognize, as well as recall, emotional experiences (Rivers, 2022).

Finally, while the IAGWS exhibited good concurrent validity, to gain a better grasp of the nomological validity of the IAGWS, further studies employing more diverse measures are called for to ascertain the divergent validity of the instrument, as well as studies including criterion variables to investigate the predictive validity of the scale.

Conclusion

The IAGWS has been shown to be a reliable and validate measure of interaction anxiety as it manifests in the context of small group work in the language learning classroom. It is hoped that this instrument will encourage greater research in this area to promote deeper understanding of the impact of anxiety on group work, as well as better knowledge of the construct of interaction anxiety. In addition, the use of this scale should provide educators with a greater awareness of their learners' lived experience and the effectiveness of pedagogical interventions aimed at reducing the detrimental influence of this negative emotion.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Tokai University Ethical Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

LX: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Validation, Writing – original draft, Writing – review & editing. MR: Conceptualization, Data curation, Formal analysis, Writing – original draft, Writing – review & editing, Methodology, Validation. BP: Data curation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing. TK: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – review & editing, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by JSPS KAKENHI Grant-in-Aid for Early-Career Scientists, Grant Number JP22K13172.

Acknowledgments

The authors wish to thank Ian Isemonger for his invaluable advice regarding the technical aspects of the analysis, as well as all of the teachers who assisted in collecting data for this study, and of course, the learners who took the time to respond to the survey.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: the case of students of Japanese. *Modern Lang. J.* 78, 155–168. doi: 10.1111/j.1540-4781.1994.tb02026.x
- Barrington, R. (2021). Including conversation strategies in an English communication course. *SILC J.* 1, 65–69.
- Bekleyen, N. (2009). Helping teachers become better English students: causes, effects, and coping strategies for foreign language listening anxiety. *System* 37, 664–675. doi: 10.1016/j.system.2009.09.010
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., and Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front. Public Health* 6, 149–166. doi: 10.3389/fpubh.2018.00149
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Hoboken, NJ: John Wiley and Sons, Inc.
- Botes, E., van der Westhuizen, L., Dewaele, J.-M., MacIntyre, P., and Greiff, S. (2022). Validating the short-form Foreign Language Classroom Anxiety Scale. *Appl. Linguist.* 43, 1006–1033. doi: 10.1093/applin/amac018
- Brown, G. T. L. (2025). *Coefficient H: Hancock's Solution to Scale Reliability Under Confirmatory Factor Analysis (Version 1)*. Auckland: The University of Auckland.
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. New York, NY: Guilford Press.
- Byrne, B. M. (2016). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. London: Routledge.
- Cantwell, R. H., and Andrews, B. (2002). Cognitive and psychological factors underlying secondary school students' feelings towards group work. *Educ. Psychol.* 22, 75–91. doi: 10.1080/01443410120101260
- Catalano, L. I., Furr, R. M., and Bellis, F. A. (2012). A multilevel analysis of the self-presentation theory of social anxiety: contextualized, dispositional, and interactive perspectives. *J. Res. Pers.* 46, 361–373. doi: 10.1016/j.jrp.2012.02.006
- Cheng, Y. S., Horwitz, E. K., and Schallert, D. L. (1999). Language writing anxiety: Differentiating writing and speaking components. *Lang. Learn.* 49, 417–446. doi: 10.1111/0023-8333.00095
- Connor, K. M., Kobak, K. A., Churchill, L. E., Katzelnick, D., and Davidson, J. R. (2001). Mini-SPIN: a brief screening assessment for generalized social anxiety disorder. *Depress. Anxiety* 14, 137–140. doi: 10.1002/da.1055
- Dewaele, J. M., and MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Stud. Second Lang. Learn. Teach.* 4, 237–274. doi: 10.14746/ssllt.2014.4.2.5
- Dörnyei, Z., and Murphey, T. (2004). *Group Dynamics in the Language Classroom*. Cambridge: Cambridge University Press.
- Farmer, N., Powell-Wiley, T. M., Middleton, K. R., Brooks, A. T., Mitchell, V., Troncoso, M., et al. (2022). Use of a focus group-based cognitive interview methodology to validate a cooking behavior survey among African-American adults. *Front. Nutr.* 9:1000258. doi: 10.3389/fnut.2022.1000258
- Field, A. (2018). *Discovering Statistics Using SPSS* (5th ed.). Thousand Oaks, CA: Sage Publications.
- Fornell, C., and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *J. Market. Res.* 18, 39–50. doi: 10.1177/002224378101800104
- Fushino, K. (2006). The relationship between communication apprehension and beliefs about group work in college English classrooms. *Cooperat. Educ.* 2, 20–37.
- Fushino, K. (2010). Causal relationships between communication confidence, beliefs about group work, and willingness to communicate in foreign language group work. *TESOL Quart.* 44, 700–724. doi: 10.5054/tq.2010.235993
- Gregersen, T. S., and Horwitz, E. K. (2002). Language learning and perfectionism: anxious and non-anxious language learners' reactions to their own oral performance. *Modern Lang. J.* 86, 562–570. doi: 10.1111/1540-4781.00161
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Boston, MA: Cengage Learning.
- Hancock, G. R., and Mueller, R. O. (2001). "Rethinking construct reliability within latent variable systems," in *Structural equation modeling: present and future - A festschrift in honor of Karl Jöreskog*, R. Cudeck, S. Du Toit, and D. Sörbom (Chapel Hill, NC: Scientific Software International Inc), 195–216.
- Heppner, P. P., Wampold, B. E., and Kivlighan, D. M. (2008). *Research Design in Counseling* (3rd ed.). Pacific Grove, CA: Thomson Brooks/Cole.
- Horwitz, E. (2017). "On the misreading of Horwitz, Horwitz and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners," in *New Insights into Language Anxiety: Theory, Research and Educational Implications*, C. Gkonou, M. Daubney, and J. M. Dewaele (Bristol: Multilingual Matters), 31–48.
- Horwitz, E. K., Horwitz, M. B., and Cope, J. (1986). Foreign language classroom anxiety. *Modern Lang. J.* 70, 125–132. doi: 10.1111/j.1540-4781.1986.tb05256.x
- International Test Commission (2017). *International Test Commission Guidelines for Translating and Adapting Tests*. Available online at: https://www.intestcom.org/files/guideline_test_adaptation_2ed.pdf (accessed October 22, 2024).
- Ito, T., Furuyabu, M., and Toews-Shimizu, J. (2022). The effects of individual-level and group-level trust on willingness to communicate in the group language learning. *Front. Educ.* 7:884388. doi: 10.3389/feduc.2022.884388
- Kalkbrenner, M. T. (2023). Alpha, omega, and H internal consistency reliability estimates: Reviewing these options and when to use them. *Counsel. Outcome Res. Eval.* 14, 77–88. doi: 10.1080/21501378.2021.1940118
- Keblowska, M. (2012). "The place of affect in second language acquisition," in *New Perspectives on Individual Differences in Language Learning and Teaching*, ed. M. Pawlak (Cham: Springer-Verlag), 157–167.
- King, J. (2013). Silence in the second language classrooms of Japanese universities. *Appl. Linguist.* 34, 325–343. doi: 10.1093/applin/ams043
- King, J. (2014). "Fear of the true self: Social anxiety and the silent behaviour of Japanese learners of English," in *The Impact of Self-Concept on Language Learning*, ed. K. Ciszér, and M. Magid (Bristol: Multilingual Matters), 232–249.
- King, J., and Smith, L. (2017). "Social anxiety and silence in Japan's tertiary foreign language classrooms," in *New Insights into Language Anxiety: Theory, Research and Educational Implications*, eds. C. Gkonou, M. Daubney, and J. M. Dewaele (Bristol: Multilingual Matters), 91–109.
- King, J., Yashima, T., Humphries, S., Aubrey, S., and Ikeda, M. (2020). "Silence and anxiety in the English-medium classroom of Japanese universities: A longitudinal intervention study," in *East Asian Perspectives on Silence in English Language Education*, eds. J. King, and S. Harumi (Bristol: Multilingual Matters), 60–79.
- Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling* (5th ed.). New York, NY: Guilford Press.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Koo, T. K., and Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* 15, 155–163. doi: 10.1016/j.jcm.2016.02.012
- La Greca, A. M., and Stone, W. L. (1993). Social Anxiety Scale for Children-Revised: factor structure and concurrent validity. *J. Clin. Child Psychol.* 22, 17–27. doi: 10.1207/s15374424jccp2201_2
- Leary, M. R. (1983a). Social anxiousness: the construct and its measurement. *J. Pers. Assess.* 47, 66–75. doi: 10.1207/s15327752jpa4701_8
- Leary, M. R. (1983b). A brief version of the Fear of Negative Evaluation scale. *Pers. Soc. Psychol. Bull.* 9, 371–375. doi: 10.1177/0146167283093007
- Leary, M. R., and Kowalski, R. M. (1993). The Interaction Anxiousness Scale: construct and criterion-related validity. *J. Pers. Assess.* 61, 136–146. doi: 10.1207/s15327752jpa6101_10
- Leary, M. R., and Kowalski, R. M. (1995). *Social Anxiety*. New York, NY: Guilford Press.
- Leeming, P. (2011). Group dynamics and SLA: A review of the literature. *Kwansei Gakuin Univ. Humanit. Rev.* 16, 147–154.
- Liebowitz, M. R. (1987). Social phobia. *Mod. Probl. Pharmacopsychiatry* 22, 141–173. doi: 10.1159/000414022
- Linnenbrink-Garcia, L., Rogat, T. K., and Koskey, K. L. (2011). Affect and engagement during small group instruction. *Contemp. Educ. Psychol.* 36, 13–24. doi: 10.1016/j.cedpsych.2010.09.001
- Liu, M., and Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *Modern Lang. J.* 92, 71–86. doi: 10.1111/j.1540-4781.2008.00687.x
- MacIntyre, P. D. (1992). *Anxiety and Language Learning from a Stages of Processing Perspective* (Unpublished doctoral dissertation). The University of Western Ontario, London, ON, Canada.
- MacIntyre, P. D., and Gregersen, T. (2012). "Affect: The role of language anxiety and other emotions in language learning," in *Psychology for language learning: Insights from research, theory and practice*, eds. S. Mercer, S. Ryan, and M. Williams (London: Palgrave Macmillan), 103–118.
- Maher, K., and King, J. (2020). Observing anxiety in the foreign language classroom: Student silence and nonverbal cues. *J. Psychol. Lang. Learn.* 2, 116–141. doi: 10.52598/jplll/2/1/6
- Maher, K., and King, J. (2022). "The silence kills me": 'Silence' as a trigger of speaking-related anxiety in the English-medium classroom. *English Teach. Learn.* 46, 213–234. doi: 10.1007/s42321-022-00119-4
- March, J. S., Parker, J. D. A., Sullivan, K., Stallings, P., and Conners, C. K. (1997). The Multidimensional Anxiety Scale for Children (MASC): factor structure, reliability, and validity. *J. Am. Acad. Child Adolesc. Psychiat.* 36, 554–565. doi: 10.1097/00004583-199704000-00019
- Masia-Warner, C., Storch, E. A., Pincus, D. B., Klein, R. G., Heimberg, R. G., and Liebowitz, M. R. (2003). The Liebowitz Social Anxiety Scale for Children and Adolescents: An Initial Psychometric Investigation. *J. Am. Acad. Child Adolesc. Psychiat.* 42, 1076–1084. doi: 10.1097/01.CHI.0000070249.24125.89
- Mattick, R. P., and Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behav. Res. Ther.* 36, 455–470. doi: 10.1016/S0005-7967(97)10031-6
- Mayring, P. (2014). *Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Software Solution*. Weinheim: Beltz.
- McCroskey, J. C., Beatty, M. J., Kearney, P., and Plax, T. G. (1985a). The content validity of the PRCA-24 as a measure of communication apprehension across communication contexts. *Commun. Q.* 33, 165–173. doi: 10.1080/01463378509369595
- McCroskey, J. C., Gudykunst, W. B., and Nishida, T. (1985b). Communication apprehension among Japanese students in native and second language. *Commun. Res. Rep.* 2, 11–15. doi: 10.1080/17464099.1985.12289912
- McCroskey, J. C., and Richmond, V. P. (1992). "Communication apprehension and small group communication," in *Small Group Communication*, eds. R. S. Cathcart, and L. A. Samovar (Baltimore, MD: William C. Brown), 342–356.
- McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychol. Methods* 23, 412–433. doi: 10.1037/met0000144
- Miura, Y. (2019). University students' social anxiety and their performance in English courses. *J. Faculty Staff Dev. Higher Educ.* 17, 9–18.
- Molenaar, I., Slegers, P., and van Boxtel, C. (2014). Metacognitive scaffolding during collaborative learning: a promising combination. *Metacognit. Learn.* 9, 309–332. doi: 10.1007/s11409-014-9118-y
- Neer, M. R. (1987). The development of an instrument to measure classroom apprehension. *Commun. Educ.* 36, 154–166. doi: 10.1080/03634528709378656
- Nezlek, J. B., and Leary, M. R. (2002). Individual differences in self-presentational motives in daily social interaction. *Pers. Soc. Psychol. Bull.* 28, 211–223. doi: 10.1177/0146167202282007
- Nihei, M., Arai, H., Maeda, K., Aoki, S., Tsuchiyagaito, A., Iwano, S., et al. (2018). Reassessment of factor structure, reliability, and validity of the Short Fear of Negative Evaluation scale for the Japanese. *Anxiety Disorder Res.* 10, 54–63. doi: 10.14389/jsad.10.1_54
- Okabayashi, N., and Seiwa, H. (1992). A study on the reliability and validity of Leary's social anxiety scale. *Memoirs of the Faculty of Integrated Arts and Sciences, Hiroshima University, III, Studies in Information and the Behavioral Sciences.* 15, 1–9. Available online at: <https://ir.lib.hiroshima-u.ac.jp/00030258> (accessed December, 15, 2024).
- Otowa, T., and Morita, M. (2015). Epidemiology of social anxiety disorder: Its history and conceptual changes. *Anxiety Disord. Res.* 7, 18–28. doi: 10.14389/jsad.7.1_18
- Polit, D. F., Beck, C. T., and Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res. Nurs. Health* 30, 459–467. doi: 10.1002/nur.20199
- Pribyl, C. B., Keaten, J. A., Sakamoto, M., and Koshikawa, F. (1998). Assessing the cross-cultural content validity of the Personal Report of Communication Apprehension scale (PRCA-24). *Jap. Psychol. Res.* 40, 47–53. doi: 10.1111/1468-5884.00074
- Ranta, K., Kaltiala-Heino, R., Rantanen, P., and Marttunen, M. (2012). The Mini-Social Phobia Inventory: psychometric properties in an adolescent general population sample. *Compr. Psychiat.* 53, 630–637. doi: 10.1016/j.comppsy.2011.07.007
- Rhee, J., Parent, D., and Basu, A. (2013). The influence of personality and ability on undergraduate teamwork and team performance. *Springerplus* 2, 1–16. doi: 10.1186/2193-1801-2-16
- Rivers, D. J. (2022). The situation specific arousal analyzer: innovation in the physiological assessment of foreign language education anxiety. *Front. Educ.* 7:802639. doi: 10.3389/feduc.2022.802639
- Russell, G., and Shaw, S. (2009). A study to investigate the prevalence of social anxiety in a sample of higher education students in the United Kingdom. *J. Mental Health* 18, 198–206. doi: 10.1080/09638230802522494
- Russell, G., and Topham, P. (2012). The impact of social anxiety on student learning and well-being in higher education. *J. Mental Health* 21, 375–385. doi: 10.3109/09638237.2012.694505
- Saito, Y., Horwitz, E. K., and Garza, T. J. (1999). Foreign language reading anxiety. *Modern Lang. J.* 83, 202–218. doi: 10.1111/0026-7902.00016
- Sakamoto, M., Pribyl, C., and Keaten, J. (1997). Departmental differences in communication apprehension as measured by the Japanese Communication Fear Scale (JCFS). *Jap. J. Res. Emot.* 5, 10–23. doi: 10.4092/jsre.5.10
- Sasagawa, S., Kanai, Y., Muranaka, Y., Suzuki, S., Shimada, H., and Sakano, Y. (2004). Development of a short Fear of Negative Evaluation scale for Japanese using item response theory. *Jap. J. Behav. Therapy* 30, 87–98.
- Schenkler, B. R., and Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization model. *Psychol. Bull.* 92, 641–669. doi: 10.1037/0033-2909.92.3.641
- Schilling, J. (2006). On the pragmatics of qualitative assessment: Designing the process for content analysis. *Eur. J. Psychol. Assessm.* 22, 28–37. doi: 10.1027/1015-5759.22.1.28
- Swain, M., Brooks, L., and Tocalli-Beller, A. (2002). Peer-peer dialogue as a means of second language learning. *Annu. Rev. Appl. Linguist.* 22, 171–185. doi: 10.1017/S0267190502000090
- Tabachnick, B. G., and Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed). Utah: Pearson.
- Terwee, C. B., Bot, S. D., de Boer, M. R., Van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* 60, 34–42. doi: 10.1016/j.jclinepi.2006.03.012
- Topham, P., Moller, N., and Davies, H. (2016). Social anxiety in learning: Stages of change in a sample of UK undergraduates. *J. Further Higher Educ.* 40, 125–145. doi: 10.1080/0309877X.2014.895307
- Watson, D., and Friend, R. (1969). Measurement of social-evaluative anxiety. *J. Consult. Clin. Psychol.* 33, 448–457. doi: 10.1037/h0027806
- Webb, N. M. (2009). The teacher's role in promoting collaborative dialogue in the classroom. *Br. J. Educ. Psychol.* 79, 1–28. doi: 10.1348/000709908X380772
- Weeks, J. W., Spokas, M. E., and Heimberg, R. G. (2007). Psychometric evaluation of the mini-social phobia inventory (Mini-SPIN) in a treatment-seeking sample. *Depress. Anxiety* 24, 382–391. doi: 10.1002/da.20250
- Wheaton, B., Muthén, B., Alwin, D. F., and Summers, G. F. (1977). Assessing reliability and stability in panel models. *Sociol. Methodol.* 8, 84–136. doi: 10.2307/270754
- Xethakis, L. (2024). Language learning, interaction and social anxiety [Poster Session]. JALT2024 International Conference, Shizuoka, Japan.
- Xethakis, L. J., Rupp, M., and Plummer, B. R. (2024). Why students feel anxious in group work: an investigation into students' perspectives on the sources of their anxious experiences in group work. *Front. Educ.* 9:1461747. doi: 10.3389/feduc.2024.1461747

- Xethakis, L. J., and Rupp, M. J. (2023). Validity and Reliability of the Interaction and Audience Anxiousness Scale in the Japanese SLA Context. *Bulletin of The School of Humanities and Science of Tokai University*, 1, 1–10. Available online at: <https://www.utokai.ac.jp/uploads/2024/05/4d5e43a0f564c97fb104a115026262d6.pdf> (accessed December, 15, 2024).
- Xethakis, L. J., and Rupp, M. J. (2024). The Psychometric Properties of the Short Fear of Negative Evaluation Scale in the Japanese SLA Context. *Bulletin of The School of Humanities and Science of Tokai University*, 2, 25–38. Available online at: http://www2.kuma.u-tokai.ac.jp/bunrikiyou/wp-content/uploads/2024/03/2024_03_LarryJ.pdf (accessed December, 15, 2024).
- Yashima, T., Ikeda, M., and Nakahira, S. (2016a). “Talk and silence in an EFL classroom: Interplay of learners and context,” in *The Dynamic Interplay Between Context and the Language Learner*, ed. J. King (London: Palgrave Macmillan), 104–126.
- Yashima, T., MacIntyre, P. D., and Ikeda, M. (2016b). Situated willingness to communicate in an L2: interplay of individual characteristics and context. *Lang. Teach. Res.* 22, 115–137. doi: 10.1177/1362168816657851
- Yashima, T., Noels, K., Shizuka, T., Takeuchi, O., Yamane, S., and Yoshizawa, K. (2009). The interplay of classroom anxiety, intrinsic motivation, gender in the Japanese EFL context. *J. Foreign Lang. Educ. Res.* 17, 41–64.
- Young, D. J. (1990). An investigation of students’ perspectives on anxiety and speaking. *Foreign Lang. Annals* 23, 539–553. doi: 10.1111/j.1944-9720.1990.tb00424.x
- Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *Modern Lang. J.* 75, 426–439. doi: 10.2307/329492
- Zhou, M. (2016). The roles of social anxiety, autonomy, and learning orientation in second language learning: a structural equation modeling analysis. *System* 63, 89–100. doi: 10.1016/j.system.2016.09.001
- Zijlmans, E. A., Tijmstra, J., Van der Ark, L. A., and Sijtsma, K. (2019). Item-score reliability as a selection tool in test construction. *Front. Psychol.* 9:2298. doi: 10.3389/fpsyg.2018.02298