Modelling trait and state willingness to communicate in a second language: An experience sampling approach

Ju Seong Lee
The Education University of Hong Kong, China
https://orcid.org/0000-0002-8756-0910
jslee@eduhk.hk

Ming Ming Chiu
The Education University of Hong Kong, China
https://orcid.org/0000-0002-5721-1971
mingchiu@eduhk.hk

Abstract
This study investigates whether demographics, L2 enjoyment, L2 anxiety, psychological well-being, or informal digital learning of English (IDLE) affect willingness to communicate in L2 (L2 WTC) at trait or state levels. Across two weeks, 16 Hong Kong EFL students completed pre- and post-trait-level surveys and generated 1,120 state-level responses via the experience sampling method (ESM). The survey findings revealed that students who were older, had higher language proficiency, or had higher pre-WTC digitally reported higher L2 WTC in the classroom. Students who had studied English longer, had higher language proficiency, had higher pre-psychological well-being, had a higher standard deviation L2 WTC over 14 days, or had higher teacher appreciation showed higher L2 WTC outside of class. Students with higher language proficiency, higher pre-WTC digitally, or higher post-psychological well-being reported higher L2 WTC in digital settings. The ESM findings revealed that students with higher teacher appreciation, currently engaged in productive IDLE or both receptive and productive...
types of IDLE, currently experienced greater L2 enjoyment, or currently experienced
greater well-being had higher L2 WTC now. Our findings are relevant to: (a) L2 WTC,
informal language learning, and positive psychology theories, and (b) strategies that
can be used to enhance students’ L2 WTC overall and at a particular moment.

*Keywords:* willingness to communicate in a second language; informal language
learning; positive psychology in SLA; experience sampling method

1. Introduction

A person’s willingness to communicate in a second language (L2 WTC), or “read-
iness to enter into discourse at a specific time with a specific person or persons,
using an L2” (MacIntyre et al., 1998, p. 547), is a key precursor of L2 use (Elahi
Shirvan et al., 2019). Previous research suggests several potential antecedents of
trait L2 WTC: demographics (age, gender, and discipline; MacIntyre et al., 2003),
L2 enjoyment (Botes et al., 2022a), L2 anxiety (Dewaele, 2019), and informal
digital learning of English (IDLE; Lee et al., 2024).

Ryff (1989) proposed a six-component framework for understanding a
person’s psychological well-being: (a) purpose in life, (b) personal growth, (c) au-
tonomy, (d) environmental mastery, (e) self-acceptance, and (f) positive relations
with others. While some studies have suggested a conceptual link between psy-
chological well-being and L2 WTC (Dewaele et al., 2019), no past study has tested
whether well-being is directly linked to L2 WTC (see Zhang et al., 2018; Shamsi
& Bozorgian, 2021). As well-being is part of a mechanism that might increase L2
WTC, our study tested a direct link between the two constructs.

Moreover, L2 students are increasingly using *digital* platforms for L2 com-
munication, while also engaging in *in-class* discussions (primarily with peers on
curriculum-related topics) and having *out-of-class* conversations (e.g., authentic
L2 interactions with friends and foreigners on various topics). Lee et al.’s (2022)
structural equation model showed that EFL students’ L2 WTC differed across
these three settings. However, past studies have not determined whether L2
WTC, its antecedents, or its mechanisms vary across in-class, out-of-class, and
digital settings (see Sudina, 2023).

Furthermore, trait-like (e.g., stable enjoyment) and state-like (e.g., mo-
mentary enjoyment) factors can affect L2 WTC at multiple levels (Cao & Philp,
2006; Pawlak & Mystkowska-Wiertelak, 2015; Pawlak et al., 2016; Peng, 2012;
Zhang et al., 2019, 2020). However, existing studies have often relied on self-re-
port measures obtained in classroom or laboratory settings, which are susceptible
to recall bias and limit the external validity of the findings (Gregersen et al., 2023).
As the experience sampling method (ESM) captures one’s real-time self-reports of thoughts, moods, or behaviors in natural settings, it can mitigate the burden of recall bias and improve the external validity of findings (Beyens et al., 2020). While real-time assessment of L2 WTC remains elusive, ESM can: (a) advance the precision of measurements, (b) enhance external validity, and (c) provide data to model both trait-like and state-like factors in L2 WTC across multiple levels.

In this study, we seek to determine whether demographics, L2 enjoyment, L2 anxiety, psychological well-being, or IDLE are linked to L2 WTC at the trait (in-class, out-of-class, and digital) or state (momentary) level. Specifically, we examined pre- and post-trait-level survey responses and state-level experiences of each participant via advanced statistics (systems of equations and multilevel cross-classification ordered logit, respectively). This study is the first to theoretically propose and empirically show L2 WTC mechanisms at the trait and state levels.

2. Theoretical framework

2.1. L2 WTC

Unlike emotions that focus on “short-lived, feeling-arousal-purposive-expressive phenomena” (Reeve, 2005, p. 294), L2 WTC concerns both stable and temporary factors (MacIntyre & Wang, 2021). Stable trait-like factors that remain consistent across time and situations (e.g., gender and extroversion) alone do not determine L2 WTC (MacIntyre et al., 2003). Dynamic interactions between L2 learners and transient state-like factors (e.g., anxiety in speaking situations and group dynamics; Dewaele & Dewaele, 2018) might also affect L2 WTC. Hence, MacIntyre et al. (1998) proposed a versatile, comprehensive L2 WTC pyramid framework of traits and states. The trait-like and stable social and individual context (Layer VI) comprises intergroup atmosphere and personality. For example, students with more international experience, a greater desire to learn about global issues and interact with foreigners, or less communication anxiety are more likely to initiate an L2 conversation (Lee, 2018; Öz et al., 2015; Yashima, 2002). The affective-cognitive context (Layer V) includes intergroup attitudes, social situation, and communicative competence. For instance, students who are more confident in speaking or who enjoy learning language with peers or teachers have higher L2 WTC (Dewaele & Dewaele, 2018; Song et al., 2022). Motivational propensities (Layer IV) comprise interpersonal motivation, intergroup motivation, and L2 self-confidence. For example, students who have higher self-perceived communicative competence or lower L2 anxiety have higher L2 WTC (Elahi Shirvan et al., 2019; Lin, 2019; MacIntyre & Wang, 2021; see also Jin & Lee, 2022).
L2 WTC situational antecedents (Layer III) are more transient, context-specific, and state-like than the previous layers; they include a desire to communicate with a specific person and state communicative self-confidence. When students are asked to speak about unfamiliar topics, or encounter technical difficulties during synchronous online class, or receive less affective peer support, they report lower L2 WTC (Kruk & Pawlak, 2022; Lee & Liu, 2022). L2 WTC is a behavioral intention; hence it is the closest precursor to communication behavior (Layer I). Notably, students with greater L2 WTC show greater L2 communication frequency and fluency (Sato & Lam, 2021; Zabihi et al., 2021).

Lee et al.’s (2022) structural equation model showed that EFL students’ L2 WTC varied across three contexts: one academic (in-class) and two non-academic (out-of-class and digital). L2 WTC in the classroom refers to students’ willingness to speak in L2 with familiar interlocutors (e.g., peers) in class, often about curriculum topics in structured and predictable communication scenarios. L2 WTC outside the classroom refers to students’ willingness to communicate in L2 in more authentic, unanticipated communication scenarios with familiar (e.g., friends) or unfamiliar (e.g., foreigners) interlocutors on a variety of topics. Lastly, L2 WTC in digital settings refers to students’ willingness to talk or write in L2 about a wide range of topics in natural, volatile communicative situations with both familiar (e.g., offline friends) and unfamiliar interlocutors (e.g., online gamers). Considering the established link between higher psychological well-being and increased social media use among university students (Bano et al., 2019), we propose that psychological well-being may positively influence L2 WTC at both trait and state levels in non-academic settings, particularly in digital environments. Overall, we posit that the factors influencing L2 WTC may directly or indirectly affect L2 WTC across levels across these three settings.

Trait data help identify enduring factors that influence a person’s overall language communication disposition (MacIntyre et al., 2003). By contrast, state data capture situational factors or specific emotional states that affect communication behaviors in real time (Pawlak & Mystkowska-Wiertelak, 2015). Hence, we combine both trait and state data to begin building a comprehensive theory of L2 communication to inform interventions and instructional strategies (Zhang et al., 2019, 2020).

2.2. Antecedents of L2 WTC

Previous research has suggested several potential antecedents of L2 WTC, including demographics, L2 enjoyment, L2 anxiety, and IDLE. Regarding demographics, studies of students showed that those who were male (Lee & Hsieh, 2019), younger
(Lee & Lee, 2020), English majors (Lee & Lee, 2020), in school for more years (Lee & Lee, 2020; MacIntyre et al., 2003), studying abroad (Kang, 2014; Lee, 2018), more proficient in the language (Zhang et al., 2019; Zhou et al., 2020), or higher achieving (specifically higher course grades; Joe et al., 2017) reported higher L2 WTC.

L2 enjoyment is “a broad positive emotion that language learners experience when their psychological needs are met during challenging language-learning activities” (Botes et al., 2021, p.1; see also Dewaele & MacIntyre, 2014). L2 enjoyment encompasses how much students enjoy learning L2 (personal enjoyment), how effectively their L2 teachers fulfill their psychological needs (teacher appreciation), and how adequately their social and psychological needs are satisfied in the classroom (social enjoyment; Botes et al., 2021). Unlike excitement, which focuses on “a feeling of elation” (Kang, 2005, p. 284), L2 enjoyment encompasses a broader spectrum of positive emotions (e.g., pleasure, satisfaction or fulfillment). Based on Fredrickson’s (2001) broaden-and-build theory, when students have a positive experience learning the L2, feel appreciated by their L2 teacher, or have their social and psychological needs met in the classroom, they are more likely to be motivated to participate in language learning activities and have more opportunities to acquire vocabulary and topical knowledge, which yield greater WTC (Botes et al., 2022a, 2022b; Kruk, 2021, 2022; Lee et al., 2022). Hence, students with more L2 enjoyment than others show greater L2 WTC (Khajavy et al., 2018; Li et al., 2022; see also Botes et al., 2022a).

L2 anxiety, defined as “the worry and negative emotional reaction aroused when learning or using a second language,” is a common problem among language learners (MacIntyre, 1999, p. 27). Unlike fear, which is often an immediate emotional response, L2 anxiety can manifest as immediate, ongoing, or chronic states of unease (Horwitz et al., 1986). L2 anxiety is linked to being judged, fear of making mistakes, and inability to communicate effectively (Lee et al., 2024), so students with higher L2 anxiety than others have lower L2 WTC (Botes et al., 2020; Jin & Lee, 2022). In light of Fredrickson’s (2001) broaden-and-build theory, L2 students’ anxiety can narrow their attention to worry about small details (e.g., grammar) rather than focusing on the broader situation (e.g., context). As a result, they may become less proactive (e.g., participate less in class activities, express fewer ideas in L2), more reticent, and reluctant to speak English (Dewaele, 2019; Kruk, 2022; Lee et al., 2024).

Students who spent more time learning and using English in extramural digital settings (IDLE) reported higher levels of L2 WTC (Lee et al., 2024, 2022; Soyoof, 2022). Unlike extramural English, which encompasses both offline and online English-mediated activities (Sundqvist & Sylvén, 2016), IDLE specifically focuses on online activities (Lee et al., 2024). According to Lee et al. (2022), receptive
IDLE (e.g., consuming English-language content online) helps EFL students expand their lexical and topic knowledge, which can improve their L2 WTC. Also, productive IDLE (e.g., interacting with other English users online) allows students to hone their communication abilities and increase their confidence to communicate with others in English, resulting in higher L2 WTC.

2.3. Psychological well-being as a potential antecedent of L2 WTC

Positive psychology theories in the context of L2 learning emphasize the importance of well-being, which includes both “feeling good” (hedonia) and “functioning well” (eudaimonia; Oxford, 2016; Ryan & Deci, 2000). While hedonia involves subjective positive emotions, which are often tangible and momentary (e.g., ecstasy, warmth, comfort, ease, satisfaction), eudaimonic elements are linked to broader meaning and values. Meaning is an ontological view on the significance of life, which can be achieved through attaining difficult goals (accomplishment) or mutually caring and valued relations with other people (positive relationships, Bershield & Reis, 1998). Meaningful accomplishments and positive relationships often require deep commitment, intense immersion, and involvement in relevant activities (engagement). Together, positive emotions, engagement, relationships, meaning, and accomplishment constitute Seligman’s (2011) five-dimensional PERMA model. Subjective well-being explores how people perceive and feel about their own well-being (Diener, 1984; Diener et al., 2010); it comprises three components: (a) life satisfaction, (b) positive affect, and (c) negative affect. Both external circumstances (e.g., income) and internal factors (e.g., personality traits) can influence subjective well-being (Diener & Chan, 2011).

Ryff (1989) proposed a six-component framework for understanding a person’s overall mental health and well-being (psychological well-being): (a) sense of a meaningful, purposeful life (e.g., making a positive contribution to the world; purpose in life), (b) openness to new experiences for personal development (personal growth), (c) self-regulated ability to make decisions/act independently without being influenced by external pressures or norms (autonomy), (d) ability to effectively manage life events by making constructive choices to reach desired goals (environmental mastery), (e) appreciation of oneself, including acceptance of flaws and weaknesses (self-acceptance), and (f) formation/maintenance of satisfying relationships with other people (positive relations with others). In their school well-being model, Konu et al. (2002) identified four dimensions: (a) school conditions (e.g., group sizes), (b) social relationships (e.g., school climate), (c) means for self-fulfillment (e.g., the value placed on student’s work), and (d) health status (e.g., chronic diseases).
Recent research in L2 learning has shown a growing interest in well-being (Dewaele et al., 2019; MacIntyre et al., 2016, 2019; MacIntyre & Mercer, 2014; Mercer & Gregersen, 2023). For example, Oxford (2016) used Seligman’s (2011) PERMA model to develop the EMPHATICS model that constitutes: (a) emotion/empathy, (b) meaning/motivation, (c) perseverance, (d) agency/autonomy, (e) time, (f) hardness/habits of mind, (g) intelligences, (h) character strengths, and (i) self-factors. Mercer (2021) adopted an ecological perspective and defines well-being in ELT as “the dynamic sense of meaning and life satisfaction emerging from a person’s subjective personal relationships with the affordances within their social ecologies” (p. 3). She underscored the need to integrate hedonic and eudaimonic perspectives and considered both individual and social aspects of well-being in ELT.

Rad and Jafarpour (2023) applied Konu et al.’s (2002) model to enhance students’ well-being by improving school conditions, fostering social relationships through student-teacher interaction, and promoting self-fulfillment. Their intervention enhanced Iranian EFL students’ well-being and their writing skills. However, past studies have mainly focused on the well-being of L2 teachers (Ebadijalal & Moradkhani, 2022; Ergün & Dewaele, 2021; Jin et al., 2021), and no published study has tested a direct link between well-being and L2 communication, such as WTC. Students with greater personal growth, autonomy, environmental mastery, or positive relations with others might have more confidence to interact with others, and possibly greater L2 WTC (Dewaele et al., 2019; Huo, 2022). Moreover, those with greater self-acceptance are more resilient to mistakes or even failures, so the occurrence of such phenomena during L2 interactions are less likely to undermine their L2 WTC (Derakhshan et al., 2022; Dewaele et al., 2019; Rad & Jafarpour, 2023). Therefore, psychological well-being might be linked to L2 WTC.

Methodologically, researchers following the principles of L2 positive psychology have adopted ESM to collect real-time data on moods, thoughts, and behaviors in natural settings, moving away from relying on retrospective surveys. For instance, Gregersen et al. (2023) studied six ESL/EFL teachers with varying levels of PERMA, using ESM to measure their stress levels and well-being. Their findings shed light on the complex relationship between stress, well-being, and health in both work and home environments. While ESM has the potential to reduce recall bias and enhance the external validity of research, its application in L2 research remains limited (Gregersen et al., 2023; Hiver & Al-Hoorie, 2020).

2.4. Aims of this study

Our literature review identified three research gaps. First, while past studies have linked demographics, L2 enjoyment, L2 anxiety, and IDLE to L2 WTC, no
published study to date has shown a link between psychological well-being and L2 WTC (see Shamsi & Bozorgian, 2021; Zhang et al., 2018). Second, although L2 students are increasingly communicating both offline (in-class and out-of-class) and online, past studies have not determined whether L2 WTC, its antecedents, or its mechanisms differ across these settings (Sudina, 2023). Third, while trait-like (e.g., stable enjoyment) and state-like (e.g., momentary enjoyment) factors can influence L2 WTC at both trait and state levels, no published study has modelled their mechanisms at both trait and state levels (Zhang et al., 2018). To address these gaps, this study investigates whether demographics, L2 enjoyment, L2 anxiety, psychological well-being, or IDLE are related to L2 WTC at the trait (in-class, out-of-class, and digital) and state (momentary) levels. We address two research questions (RQ) in this study:

RQ1: How are demographics, L2 enjoyment, L2 anxiety, psychological well-being, or IDLE related to L2 WTC at the trait level (in-class, out-of-class, and digitally)?

RQ2: How are demographics, L2 enjoyment, L2 anxiety, psychological well-being, or IDLE related to L2 WTC at the state (momentary) level?

3. Methods

In this study, we collected university students’ pre- and post-survey responses and experience samples (five times per day for 14 days) regarding L2 WTC and its potential antecedents. We analyzed the survey responses with systems of equations, and the experience samples with multilevel cross-classification ordered logit.

3.1. Context and participants

We recruited 16 Hong Kong EFL undergraduate and postgraduate students (M = 21.8 years old, range: 17-27 years; 81% females, 19% males). 50% of these students majored in English, 19% in non-English language majors, and 31% in other disciplines (e.g., child education, finance, science, and information technology). Among these participants, 31% had studied abroad, typically for less than a year. While Putonghua/Mandarin was their native language, 88% of them had learned English for more than ten years and 12% for seven to ten years. As these participants had high previous English grades (75% A and 25% B), their mean self-perceived English proficiency was 7 (max: 10, range: 3 to 9). Lastly, they spent a mean of 4 hours 34 minutes per day on IDLE during the past month (range: 30 minutes to 10 hours).
3.2. Recruitment procedure

After our university’s ethics committee approved our study, we advertised for study participants at a public university in Hong Kong. We informed them about this study’s goals, benefits, risks, and confidentiality measures (e.g., ensured data anonymization by using arbitrary ID numbers and restricting access to the data). Potential participants were also informed about their right to withdraw from the study at any point. Then, interested students provided their informed consent to participate.

3.3. Statistical power

Statistical power differed across levels of students and their survey responses. For $\alpha = .05$ and a medium effect size of 0.3, statistical power for 16 students is very low but for 1,120 experience samples it exceeded .99 (Konstantopoulos, 2008). As this small sample had low statistical power at the student level, the likelihood that a non-significant student-level result was a false negative was high, but we retained our usual confidence in our significant results.

3.4. Variables

The pre- and post-online survey of traits has five sections: demographics, L2 WTC, L2 enjoyment, L2 anxiety, and psychological well-being. Demographic questions included gender, age, English major, duration of English study, duration of overseas study, language proficiency, course grade, and IDLE frequency.

L2 WTC (Lee et al., 2022) has three subscales: (a) L2 WTC in-class (5 items; Cronbach’s $\alpha = .837$; e.g., “When you are given an opportunity to talk freely in an English class”), (b) L2 WTC out-of-class (5 items; $\alpha = .805$; e.g., “When you and a small group of foreign friends engage in a discussion outside of school”), and (c) L2 WTC digitally (4 items; $\alpha = .910$; e.g., “When you have a chance to talk to other game players in English”). These Likert-type scales ranged from 1 (definitely not willing) to 5 (definitely willing).

L2 enjoyment (Dewaele & MacIntyre, 2014; Lee et al., 2022) has three subscales: (a) personal enjoyment (6 items; $\alpha = .859$; e.g., “I enjoy learning English”), (b) teacher appreciation (5 items; $\alpha = .838$; e.g., “The English teacher is supportive”), and (c) social enjoyment (6 items; $\alpha = .903$; e.g., “The classmates are nice in English class”). The scales of L2 enjoyment, L2 anxiety, and psychological well-being range from 1 (strongly disagree) to 5 (strongly agree).
L2 anxiety (Lee et al., 2024) has two subscales: (a) face-to-face L2 anxiety (6 items; \( \alpha = .938 \); e.g., “I feel nervous when I speak English in front of other students”) and (b) digital L2 anxiety (8 items; \( \alpha = .972 \); e.g., “I feel nervous when I speak English with other fans in an online community”).

Psychological well-being was adopted from Ryff and Keyes (1995, 11 items; \( \alpha = .850 \); e.g., “I like most parts of my personality”).

3.5. Experience sampling method

To measure state-level responses, we collected real-time data from participants using the experience sampling method (Gregersen et al., 2023; Hiver & Al-Hoorie, 2020). We used Google Forms, a free, easy-to-use, survey tool, which piped respondents’ data into an Excel spreadsheet. The ESM survey has six sets of questions: (a) five reporting timeslots (e.g., 9 am-12 pm), (b) IDLE type (receptive IDLE, productive IDLE, and both receptive and productive IDLE) and duration, (c) L2 enjoyment level, (d) L2 anxiety level, (e) L2 WTC level, and (f) psychological well-being level. At each assessment, participants indicated their levels of L2 enjoyment (“How much did you enjoy learning and using English during this time?”), L2 anxiety (“How much anxiety did you have while learning and using English during this time?”), L2 WTC (“How much were you willing to communicate with others in English during this time?”), and psychological well-being (“How happy were you during this time?”). To reduce response burden and fatigue, we used a single-item, Likert-type scale ranging from 1 (not enjoyable/anxious/happy at all, extremely unwilling) to 5 (extremely enjoyable/anxious/happy, extremely willing), as in most ESM studies (Beyens et al., 2020).

A research assistant sent WhatsApp messages to participants between 9:01 am and 9:01 pm daily for 14 days to report their IDLE, L2 enjoyment, L2 anxiety, L2 WTC, and psychological well-being levels. Except for the final timeslot (9:01 pm to 9 am [night until morning]), the time window was divided into four 3-hour intervals for participants (e.g., 3:01 pm-6 pm). Following standard ESM procedures, we required participants to reply to each message within three hours, except for the final timeslot (9 pm to 9 am). For missed responses, the assistant sent reminders or calls before the end of the day. For ESM data collection, participants used Google Forms polls within WhatsApp, which automatically piped these data into an Excel spreadsheet.

3.6. Data analysis

To accurately analyze these data, we addressed the following issues with statistical strategies: (a) similarities versus differences of responses of students across
time with multilevel (ML) cross-classification analysis (Hox et al., 2017), (b) time differences with Q-statistics (Ljung & Box, 1979), (c) ordered outcomes with ordered Logit/Probit (Martinez et al., 2017) and odds ratios (Martinez et al., 2017), (d) multiple outcomes/dependent variables with systems of equations (Kennedy, 2008) (e) sequences across time with vector auto-regression (VAR, Kennedy, 2008), (f) indirect, ML mediation effects with multilevel M-tests (MacKinnon et al., 2004), (g) cross-level interactions (student x time) with random parameters in random effects models (Hox et al., 2017), (h) many hypotheses’ false positives with the two stage linear step-up procedure (Benjamini et al., 2006), (i) comparison of effect sizes with Lagrange multiplier tests (Bertsekas, 2014), and (j) consistency of results across data sets (robustness) with separate multilevel, single outcome models and analyses of data subsets (Kennedy, 2008; see Table 1; see Ahn et al., 2021).

Table 1 Statistical strategies addressing each analytic difficulty

<table>
<thead>
<tr>
<th>Analytic difficulty</th>
<th>Statistics strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dataset</strong></td>
<td></td>
</tr>
<tr>
<td>• Cross-classified data (surveys across time and across students)</td>
<td>• Multilevel cross-classification (Hox et al., 2017)</td>
</tr>
<tr>
<td><strong>Outcome variables</strong></td>
<td></td>
</tr>
<tr>
<td>• Similar adjacent talk ( t_3 \sim t_4 )</td>
<td>• Q-statistics (Ljung &amp; Box, 1979)</td>
</tr>
<tr>
<td>• Differences across time (Non-stationarity)</td>
<td>• Multilevel cross-classification (Hox et al., 2017)</td>
</tr>
<tr>
<td>• Ordered variable (yes/no)</td>
<td>• Ordered logit/probit; odds ratios (Martinez et al., 2017)</td>
</tr>
<tr>
<td>• Multiple outcomes ( Y_1, Y_2, \ldots )</td>
<td>• System of equations (Kennedy, 2008)</td>
</tr>
<tr>
<td><strong>Explanatory variables</strong></td>
<td></td>
</tr>
<tr>
<td>• Sequences across time ( X_{t2} ) or ( X_{t1} \rightarrow Y_0 )</td>
<td>• Vector auto-regression (VAR, Kennedy, 2008)</td>
</tr>
<tr>
<td>• Indirect, multi-level mediation effects ( X \rightarrow M \rightarrow Y )</td>
<td>• Multilevel M-test (MacKinnon et al., 2004)</td>
</tr>
<tr>
<td>• Cross-level interactions (Student x Time)</td>
<td>• Random effects model (Hox et al., 2017)</td>
</tr>
<tr>
<td>• Many hypotheses’ false positives</td>
<td>• Two-stage linear step-up procedure (Benjamini et al., 2006)</td>
</tr>
<tr>
<td>• Compare effect sizes ( \beta_1 &gt; \beta_2 ? )</td>
<td>• Lagrange multiplier tests (Bertsekas, 2014)</td>
</tr>
<tr>
<td>• Consistency of results across data sets (Robustness)</td>
<td>• Separate multilevel, single outcome models</td>
</tr>
<tr>
<td></td>
<td>• Analyses of subsets of the data (Kennedy, 2008)</td>
</tr>
</tbody>
</table>

3.7. Explanatory model

First, we modeled a vector of Bartlett factor scores (Joreskog & Sorbom, 2018) of student Post-WTC with a system of equations (Kennedy, 2008).

\[
(1) \text{Post-WTC}_yi = \beta_y + \epsilon_{yi}
\]
In Post-WTC, outcome $y$ (WTC in-class, WTC out-of-class, WTC digitally) of student $i$ had a grand mean intercept $\beta_y$, with unexplained components (residuals) at the student-level ($e_{yi}$).

We entered explanatory variables in temporally ordered, sequential sets to estimate the variance explained by each set and to test for mediation effects (Kennedy, 2008). After entering Demographics (gender, age, English major, duration of studying English, duration of overseas study, language proficiency, course grade), we added measures before the intervention Pre (pre-teacher appreciate, pre-personal enjoyment, pre-social enjoyment, pre-anxiety, pre-psychological well-being, pre-WTC), followed by Experience (hours of IDLE receptive only [14-day mean, 14-day standard deviation [SD]], hours of IDLE productive only [mean, SD], hours of IDLE both [mean, SD], enjoy [mean, SD], anxiety [mean, SD], well-being [mean, SD], WTC [mean, SD]), and then Post (post-teacher appreciate, post-personal enjoyment, post-social enjoyment, post-anxiety, post-psychological well-being).

$\text{(2)} \quad WTC_{yi} = \beta_y + e_{yi} + \beta_{yu} \text{Demographics}_{yi} + \beta_{yw} \text{Experience}_{yi} + \beta_{yx} \text{Post}_{yi} + \beta_{yz} \text{Interactions}_{yi}$

A nested hypothesis test ($\Delta \chi^2$LL) determines the significance of each set of explanatory variables (Kennedy, 2008). For greater precision and less multicollinearity, we omitted non-significant variables (which do not cause omitted variable bias, Cinelli & Hazlett, 2020).

Second, we modeled $WTC_{now}$ with a multilevel cross-classification ordered logit (Hox et al., 2017).

$\text{(3)} \quad WTC_{now(1)} = \beta + e_{it} + f_i + g_t$

$WTC_{now}$ of student $i$ at time $t$ had grand mean intercept $\beta$, with residuals at the student-time moment ($e_{yi}$) for each student ($f_i$) at each time ($g_t$).

$\text{(4)} \quad WTC_{now(1)} = \beta + e_{it} + f_i + g_t + \beta_{r} \text{Demographics}_i + \beta_{s} \text{Pre}_i + \beta_{u} \text{Now(1)} + \beta_{v} \text{Previous(1)} + \beta_{w} \text{Earlier(1)} + \beta_{z} \text{Cumulative(1)} + \beta_{iz} \text{Interactions}_{ij}$

Demographics and then Pre were entered as above. Next, we entered Now (IDLE receptive only now, IDLE productive only now, IDLE both now, IDLE neither now, enjoy now, anxiety now, well-being now), followed by Previous (IDLE receptive only in the previous time period [-1], IDLE productive only [-1], IDLE both [-1], IDLE neither [-1], enjoy [-1], anxiety [-1], well-being [-1], WTC [-1]). Then, we entered Earlier (IDLE receptive only [-2], IDLE productive only [-2], IDLE both [-2], IDLE neither [-2], enjoy
Modelling trait and state willingness to communicate in a second language: An experience...  

[-2], anxiety [-2], well-being [-2], WTC [-2]), followed by Cumulative (IDLE receptive cumulative mean, IDLE productive cumulative mean, IDLE both cumulative mean, IDLE neither cumulative mean, enjoy cumulative mean, anxiety cumulative mean, well-being cumulative mean, IDLE receptive cumulative SD, IDLE productive cumulative SD, IDLE both cumulative SD, IDLE neither cumulative SD, enjoy cumulative SD, anxiety cumulative SD, well-being cumulative SD), and then Interactions.

We used an alpha level of .05. Also, we analyzed residuals for influential outliers. To test whether the results differ across link functions (logit vs. probit), we also used multilevel, cross-classification probit.

4. Results

4.1. Summary statistics

Summary statistics are included in Tables 2 and 3 (see Appendix Tables A1 and A2 for correlation-variance-covariance matrices). Notably, the mean daily IDLE via ESM (4.86 hours; range: .94 to 8.88) resembled the mean daily IDLE via survey (4.57 hours, range: .50 hours to 10.00 hours). This similarity implies that a survey asking about IDLE frequency in the previous month is nearly as accurate as IDLE frequency captured by ESM.

Table 2 Summary statistics of students (N = 16)

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-WTC in-class</td>
<td>4.098</td>
<td>.598</td>
<td>3.035</td>
<td>4.160</td>
<td>5</td>
</tr>
<tr>
<td>Post-WTC out-of-class</td>
<td>4.021</td>
<td>.854</td>
<td>1.739</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Post-WTC digitally</td>
<td>4.094</td>
<td>.917</td>
<td>2</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td>8.500</td>
<td>4.761</td>
<td>1</td>
<td>8.5</td>
<td>16</td>
</tr>
<tr>
<td>Age</td>
<td>21.750</td>
<td>2.543</td>
<td>17</td>
<td>21.5</td>
<td>27</td>
</tr>
<tr>
<td>Female</td>
<td>.813</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Studying English duration (years)</td>
<td>11.125</td>
<td>1.088</td>
<td>7.5</td>
<td>11.5</td>
<td>11.5</td>
</tr>
<tr>
<td>Overseas study</td>
<td>.313</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Major English</td>
<td>.250</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Course grade A</td>
<td>.750</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Language proficiency</td>
<td>7.000</td>
<td>1.592</td>
<td>3</td>
<td>7.5</td>
<td>9</td>
</tr>
<tr>
<td><strong>Pre-experience</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-WTC in-class</td>
<td>4.380</td>
<td>.742</td>
<td>2.625</td>
<td>4.780</td>
<td>5</td>
</tr>
<tr>
<td>Pre-WTC out-of-class</td>
<td>4.283</td>
<td>.663</td>
<td>2.931</td>
<td>4.362</td>
<td>5</td>
</tr>
<tr>
<td>Pre-WTC digitally</td>
<td>4.531</td>
<td>.670</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Pre-face-to-face anxiety</td>
<td>7.608</td>
<td>1.216</td>
<td>6</td>
<td>7.936</td>
<td>9.763</td>
</tr>
<tr>
<td>Pre-digital anxiety</td>
<td>7.496</td>
<td>1.480</td>
<td>6</td>
<td>7.261</td>
<td>9.979</td>
</tr>
</tbody>
</table>
### Experience

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-enjoy teacher appreciates me</td>
<td>9.152</td>
<td>.696</td>
<td>7.983</td>
<td>9.064</td>
<td>10</td>
</tr>
<tr>
<td>Pre-personal enjoyment</td>
<td>8.850</td>
<td>.762</td>
<td>6.962</td>
<td>8.909</td>
<td>9.872</td>
</tr>
<tr>
<td>Pre-social enjoyment</td>
<td>9.376</td>
<td>.532</td>
<td>8</td>
<td>9.500</td>
<td>10</td>
</tr>
<tr>
<td>Pre-well-being</td>
<td>8.797</td>
<td>.705</td>
<td>7.557</td>
<td>8.870</td>
<td>10</td>
</tr>
</tbody>
</table>

#### Post-experience

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-face-to-face anxiety</td>
<td>7.923</td>
<td>1.059</td>
<td>6</td>
<td>8.019</td>
<td>9.618</td>
</tr>
<tr>
<td>Post-digital anxiety</td>
<td>7.177</td>
<td>1.083</td>
<td>6</td>
<td>7.170</td>
<td>9.353</td>
</tr>
<tr>
<td>Post-teacher appreciates me</td>
<td>8.970</td>
<td>.611</td>
<td>7.876</td>
<td>8.983</td>
<td>9.983</td>
</tr>
<tr>
<td>Post-personal enjoyment</td>
<td>8.654</td>
<td>.785</td>
<td>7.120</td>
<td>8.842</td>
<td>9.816</td>
</tr>
<tr>
<td>Post-social enjoyment</td>
<td>9.084</td>
<td>.675</td>
<td>7.832</td>
<td>9.033</td>
<td>10</td>
</tr>
<tr>
<td>Post-well-being</td>
<td>8.598</td>
<td>.646</td>
<td>7.315</td>
<td>8.595</td>
<td>9.771</td>
</tr>
</tbody>
</table>

*Note.* For dichotomous variables, the mean indicates the proportion of participants with the attribute, and $SD$ was not meaningful.

### Table 3  Summary statistics of students in time ($N = 1,120$)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>WTC now</td>
<td>3.688</td>
<td>1.204</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

#### Demographics

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student</td>
<td>8.500</td>
<td>4.612</td>
<td>1</td>
<td>8.5</td>
<td>16</td>
</tr>
<tr>
<td>Day</td>
<td>7.500</td>
<td>4.033</td>
<td>1</td>
<td>7.5</td>
<td>14</td>
</tr>
<tr>
<td>Time</td>
<td>3.000</td>
<td>1.415</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Age</td>
<td>21.750</td>
<td>2.463</td>
<td>17</td>
<td>21.5</td>
<td>27</td>
</tr>
<tr>
<td>Female</td>
<td>.813</td>
<td>.390</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Studying English duration (years)</td>
<td>11.125</td>
<td>1.054</td>
<td>7.5</td>
<td>11.5</td>
<td>11.5</td>
</tr>
<tr>
<td>Overseas study</td>
<td>.313</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>English major</td>
<td>.250</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Course grade A</td>
<td>.750</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Language proficiency</td>
<td>7.000</td>
<td>1.542</td>
<td>3</td>
<td>7.5</td>
<td>9</td>
</tr>
</tbody>
</table>

#### Pre-experience

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$</th>
<th>$SD$</th>
<th>Min</th>
<th>Median</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-WTC in-class</td>
<td>4.380</td>
<td>.719</td>
<td>2.625</td>
<td>4.780</td>
<td>5</td>
</tr>
<tr>
<td>Pre-WTC out-of-class</td>
<td>4.283</td>
<td>.642</td>
<td>2.931</td>
<td>4.362</td>
<td>5</td>
</tr>
</tbody>
</table>
### 4.2. Explanatory models

#### 4.2.1. Post-survey WTCs

Demographics, pre-experience, experience, and post-experience variables were linked to WTC in-class, WTC out-of-class, and WTC digitally (see Table 4).

Demographics (age, language proficiency) and pre-experience (pre-WTC digitally) were linked to WTC in class. Among these students, those who were a year older or with one level higher language proficiency were more likely to report one level higher WTC in class (+1% or +6%, respectively). Also, those with one level higher pre-WTC digitally were 8% more likely to indicate one level higher WTC in class. This model accounted for nearly 78% of the differences in these students’ WTC in class.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>n1</th>
<th>n2</th>
<th>n3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-WTC digitally</td>
<td>4.531</td>
<td>.649</td>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Pre-face-to-face anxiety</td>
<td>7.608</td>
<td>1.178</td>
<td>6</td>
<td>7.936</td>
<td>9.763</td>
</tr>
<tr>
<td>Pre-digital anxiety</td>
<td>7.496</td>
<td>1.434</td>
<td>6</td>
<td>7.261</td>
<td>9.797</td>
</tr>
<tr>
<td>Pre-personal enjoyment</td>
<td>8.850</td>
<td>.738</td>
<td>6.962</td>
<td>8.909</td>
<td>9.872</td>
</tr>
<tr>
<td>Pre-social enjoyment</td>
<td>9.376</td>
<td>.515</td>
<td>8</td>
<td>9.500</td>
<td>10</td>
</tr>
<tr>
<td>Pre-enjoy teacher appreciates me</td>
<td>9.152</td>
<td>.674</td>
<td>7.983</td>
<td>9.064</td>
<td>10</td>
</tr>
<tr>
<td>Pre-psychological well-being</td>
<td>8.797</td>
<td>.683</td>
<td>7.557</td>
<td>8.870</td>
<td>10</td>
</tr>
</tbody>
</table>

#### Now

- **IDLE receptive only now**
  - Mean: .458
  - SD: .721
  - n: 0
  - p: 0
  - %: 3.5

- **IDLE productive only now**
  - Mean: .153
  - SD: .514
  - n: 0
  - p: 0
  - %: 3

- **IDLE both now**
  - Mean: .361
  - SD: .796
  - n: 0
  - p: 0
  - %: 4

- **Enjoy now**
  - Mean: 3.919
  - SD: 1.131
  - n: 1
  - p: 4
  - %: 5

- **Anxiety now**
  - Mean: 1.755
  - SD: .946
  - n: 1
  - p: 1
  - %: 5

- **Well-being now**
  - Mean: 3.719
  - SD: 1.099
  - n: 1
  - p: 4
  - %: 5

#### Previous time period

- **IDLE receptive only (-1)**
  - Mean: .458
  - SD: .722
  - n: 0
  - p: 0
  - %: 3.5

- **IDLE productive only (-1)**
  - Mean: .155
  - SD: .517
  - n: 0
  - p: 0
  - %: 3

- **IDLE both (-1)**
  - Mean: .356
  - SD: .791
  - n: 0
  - p: 0
  - %: 4

- **Enjoy (-1)**
  - Mean: 3.916
  - SD: 1.134
  - n: 1
  - p: 4
  - %: 5

- **Anxiety (-1)**
  - Mean: 1.761
  - SD: .950
  - n: 1
  - p: 1
  - %: 5

- **WTC (-1)**
  - Mean: 3.687
  - SD: 1.205
  - n: 1
  - p: 4
  - %: 5

- **Well-being (-1)**
  - Mean: 3.714
  - SD: 1.101
  - n: 1
  - p: 4
  - %: 5

#### Earlier time period

- **IDLE receptive only (-2)**
  - Mean: .460
  - SD: .724
  - n: 0
  - p: 0
  - %: 3.5

- **IDLE productive only (-2)**
  - Mean: .155
  - SD: .520
  - n: 0
  - p: 0
  - %: 3

- **IDLE both (-2)**
  - Mean: .354
  - SD: .791
  - n: 0
  - p: 0
  - %: 4

- **Enjoy (-2)**
  - Mean: 3.916
  - SD: 1.138
  - n: 1
  - p: 4
  - %: 5

- **Anxiety (-2)**
  - Mean: 1.760
  - SD: .951
  - n: 1
  - p: 1
  - %: 5

- **WTC (-2)**
  - Mean: 3.685
  - SD: 1.207
  - n: 1
  - p: 4
  - %: 5

- **Well-being (-2)**
  - Mean: 3.712
  - SD: 1.104
  - n: 1
  - p: 4
  - %: 5

*Note. For dichotomous variables, the mean indicates the proportion of participants with the attribute, and SD was not meaningful.*
Demographics (studying English duration, language proficiency), pre-experience (pre-well-being), experience (SD WTC over 14 days), and post-experience (the teacher appreciates me) were linked to WTC out-of-class. Among these students, those having studied English for a year longer or with one level higher language proficiency were more likely to report one level higher WTC out-of-class (+9% or +14%, respectively). Students with greater pre-psychological well-being than others were 19% more likely to indicate one level higher WTC out-of-class. Among these students, those with one standard deviation greater WTC over 14 days were 33% more likely to report one level higher WTC out-of-class (largest effect size). Students who perceived that their teacher appreciated them more than others (by one level) were 22% more likely to indicate one level higher WTC out-of-class. This model accounted for over 89% of the variance in these students’ WTC out-of-class.

Table 4 Summary of system of equations’ unstandardized regression coefficients (standard errors in parentheses) and odds ratios modeling WTC in-class, WTC out-of-class, and WTC digitally

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>WTC in class</th>
<th>WTC out-of-class</th>
<th>WTC digitally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.016)</td>
<td>+1%</td>
<td></td>
</tr>
<tr>
<td>Studying English duration (years)</td>
<td>.378</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.047)</td>
<td>+9%</td>
<td></td>
</tr>
<tr>
<td>Language proficiency</td>
<td>.228</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.074)</td>
<td>+6%</td>
<td></td>
</tr>
<tr>
<td>Pre-psychological well-being</td>
<td>.800</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.088)</td>
<td>+19%</td>
<td></td>
</tr>
<tr>
<td>Pre-WTC digitally</td>
<td>.315</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.106)</td>
<td>+8%</td>
<td></td>
</tr>
<tr>
<td>SD WTC over 14 days</td>
<td>1.606</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.224)</td>
<td>+33%</td>
<td></td>
</tr>
<tr>
<td>Post-teacher apprecates me</td>
<td>.965</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.069)</td>
<td>+22%</td>
<td></td>
</tr>
<tr>
<td>Post-psychological well-being</td>
<td>.594</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.172)</td>
<td>+14%</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>.777</td>
<td></td>
<td>.891</td>
</tr>
</tbody>
</table>

Note. Each outcome was modeled with a constant term. *p < .05, **p < .01, ***p < .001

Demographics (language proficiency), pre-experience (pre-WTC digitally), and post-experience (post-well-being) were linked to WTC digitally. Among these students, those with one level greater language proficiency were 9% more likely to report one level higher WTC digitally. Students with one level higher pre-WTC digitally were 16% more likely to indicate one level higher WTC digitally. Among these students, those with one level greater post-psychological well-being were 14% more likely to report one level higher WTC digitally. This model accounted for nearly 78% of the variance in these students’ WTC digitally.
4.2.2. WTC now experience samples

Most of the differences in WTC now were across time (67%) rather than across students (33%; see Table 5). All results discussed below described first entry into the regression, controlling for all previously included variables. Ancillary regressions and statistical tests are available upon request.

Table 5 Summary of multilevel cross-classification logit’s unstandardized regression coefficients (standard errors in parentheses) and odds ratios modeling WTC now

<table>
<thead>
<tr>
<th>Explanatory variable</th>
<th>Regressions predicting WTC now</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td></td>
</tr>
<tr>
<td>Pre-enjoy: teacher appreciates me</td>
<td>.878</td>
<td>.878</td>
<td>.987</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.430) +21%</td>
<td>(.428) +21%</td>
<td>(.328) +23%</td>
<td></td>
</tr>
<tr>
<td>IDLE productive now</td>
<td>.385</td>
<td>**</td>
<td>.631</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td>(.121) +10%</td>
<td>(.147) +15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDLE both now</td>
<td>.498</td>
<td>***</td>
<td>.558</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td>(.083) +12%</td>
<td>(.097) +14%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enjoy now</td>
<td></td>
<td></td>
<td>.631</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.100) +15%</td>
<td></td>
</tr>
<tr>
<td>Well-being now</td>
<td></td>
<td></td>
<td>1.223</td>
<td>***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(.107) +27%</td>
<td></td>
</tr>
<tr>
<td>Variance at each level</td>
<td>Explained variance at each level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student (33%)</td>
<td>.108</td>
<td>.151</td>
<td>.728</td>
<td></td>
</tr>
<tr>
<td>Time (67%)</td>
<td>.065</td>
<td>.103</td>
<td>.492</td>
<td></td>
</tr>
<tr>
<td>Total variance explained</td>
<td>.080</td>
<td>.119</td>
<td>.570</td>
<td></td>
</tr>
</tbody>
</table>

Note. Each model included a constant term. *p < .05, **p < .01, ***p < .001

Pre-experience, IDLE, and current psychological measures were linked to WTC now. Among these students, those perceiving that their teacher appreciated them one level more were 21% more likely to report one level higher WTC now. Among these students, those currently engaged in productive IDLE or both types of IDLE were more likely to report one level higher WTC now (+10% or +12%, respectively). Students perceiving greater enjoyment or well-being now were more likely to indicate one level higher WTC now (+15% or +27%, respectively). The final model accounted for 57% of the variance in WTC now.

All other results were non-significant. Notably, no mediation tests or interactions were significant. Robustness tests with single outcomes, data subsets or multilevel cross-classification probit showed similar results. Analysis of residuals show no influential outliers.

5. Discussion

Unlike previous piecemeal testing studies that examined L2 WTC at the trait or state level, this study proposed and tested a comprehensive theoretical model
of how demographics, L2 enjoyment, L2 anxiety, psychological well-being, or IDLE were related to L2 WTC at both the trait (in-class, out-of-class, and digitally) and state (momentary) levels.

5.1. L2 WTC trait

Specifically, in the classroom, students who were older, had higher language proficiency, or had higher pre-WTC digitally reported higher L2 WTC, aligning with past studies (Lee & Lee, 2020; Zhou et al., 2020). Building on past studies showing that WTC digitally was linked with WTC in-class (Lee & Lee, 2020), this study showed that pre-WTC digitally affects post-WTC in-class to add nuance to our understanding. This is the first evidence showing that students who were more likely to engage in English conversations in digital settings became more willing to subsequently speak English in the classroom.

Outside of class, students who studied English longer, had higher language proficiency, had higher pre-psychological well-being, had a higher standard deviation L2 WTC over 14 days, or perceived that their teacher appreciated them showed higher L2 WTC, aligning with past studies (Lee et al., 2022; Lee & Lee, 2020; MacIntyre et al., 2003; Zhou et al., 2020). Notably, this is the first study to show that students with higher pre-psychological well-being had higher L2 WTC outside the classroom. This result suggests that students with better psychological health were more likely than others to subsequently engage in English communication outside the classroom. This study also shows the first evidence that students with a higher standard deviation L2 WTC over 14 days than others had higher L2 WTC outside the classroom (largest effect size). This result suggests that such students who encounter a wider variety of easy (with friends in familiar settings) and difficult (with strangers in unfamiliar settings) communicative situations than others (higher L2 WTC volatility) are more willing to subsequently communicate in L2. If future studies (especially intervention studies) confirm this finding, then a teacher should expose their students to a greater variety of L2 communicative situations that range in difficulty (not only easy ones and not only hard ones) to improve their L2 WTC outside the classroom.

In digital settings, students with higher language proficiency, higher pre-WTC digitally, or higher post-psychological well-being reported higher L2 WTC. Building on past studies showing that students with higher language proficiency have higher L2 WTC offline (Zhang et al., 2019; Zhou et al., 2020), our findings offer the first evidence that higher language proficiency yields higher L2 WTC online. Consistent with the above WTC in class and WTC out-of-class results, students with higher pre-WTC digitally or higher post-psychological well-being reported higher post-L2 WTC digitally. These findings suggest that students who are psychologically
healthy with a positive attitude toward their lives are more likely to subsequently engage in English communication in digital environments.

5.2. ESM

Our immediate, fine-grained, ESM results revealed that students who perceived that their teacher appreciated them, currently experienced greater L2 enjoyment, currently engaged in productive IDLE or both receptive and productive IDLE, or currently experienced greater psychological well-being had higher L2 WTC now. Building on past trait L2 WTC research (Botes et al., 2022a; Elahi Shirvan et al., 2019; Lee et al., 2024), our findings showed that teacher appreciation, L2 enjoyment and IDLE affected L2 WTC in the moment. These findings imply that when teachers meet L2 students’ psychological needs at the trait level or foster their L2 enjoyment at the state level, students are more likely to speak English in the moment.

Our findings also indicate that students’ current participation in IDLE, particularly production-oriented IDLE, raises the likelihood of immediate L2 WTC. This immediate evidence shows that incorporating IDLE into L2 activities can facilitate more state L2 WTC in class or in everyday life situations. Unlike past studies with large sample sizes (N = 388; Lee & Sylvén, 2021; N = 1,265; Lee et al., 2024), trait-level IDLE was not linked to trait L2 WTC; however, their high correlation (r [mean IDLE, mean WTC] = .618) suggests that the low statistical power of the small sample (N = 16) might have caused a false negative. Another possible interpretation is that watching English-language Netflix or YouTube videos and interacting with other English users on social media may have a transient effect on L2 WTC, but no effect at the dispositional level. Past studies showing a positive relationship between IDLE frequency over six months and L2 WTC suggest a third possibility: Students might need to engage in IDLE over a longer period of time (e.g., six months) to affect their trait L2 WTC (Lee & Sylvén, 2021). Future studies can adjudicate these possibilities.

While demographic or trait-like factors mainly affected trait L2 WTC, state-like factors largely influenced state L2 WTC. These results imply that state-like/momentary affective or pedagogical support, whether in the classroom or in real-life situations (out-of-class and digitally), can immediately facilitate students’ L2 WTC. These findings align with the view that trait L2 WTC is a person’s overall disposition toward L2 communication, so it is more stable and less susceptible to change across contexts. By contrast, as state L2 WTC is a person’s WTC in a specific situation, state L2 WTC is more local, so situation-specific attributes might change it. These results also support MacIntyre et al.’s (1998) framework, in which trait-like factors (e.g., age, language proficiency, and teacher appreciation) align with stable individual, social, and affective-cognitive contexts (Layers VI
and V), whereas motivational tendencies and situational antecedents (Layers IV and III) are linked to state-like factors (e.g., current participation in productive IDLE, current feeling of L2 enjoyment, and psychological well-being).

While past studies theorized a potential link between psychological well-being and L2 WTC (Dewaele et al., 2019), this is the first empirical study to show that students with a higher current state of psychological well-being can yield immediately higher L2 WTC. This result suggests further studies testing whether a teacher who enhances students’ current psychological well-being immediately raises their L2 WTC. Surprisingly, both trait-like and state-like L2 anxiety were unrelated to L2 WTC at both the trait and state levels. Hence, EFL students with lower L2 anxiety may not necessarily become willing to communicate in English overall and in the moment. These results contradict the majority of previous research findings (showing that lower L2 anxiety correlated with higher L2 WTC; Botes et al., 2020; Dewaele, 2019; Jin & Lee, 2022; Lee et al., 2024). This study’s high correlation at the trait-level of mean anxiety and mean L2 WTC ($r = -.540$) suggests that the low statistical power of the small sample ($N = 16$) might have caused a false negative, but the large ESM data ($N = 1,120$) showed high statistical power (> .99).

Interestingly, positive affective factors such as psychological well-being and L2 enjoyment are linked to L2 WTC at both the trait (out-of-class and digitally) and state levels. These findings support the positive psychology perspective in L2 teaching and learning: prioritize developing positive emotions and psychological attributes (e.g., enjoyment, motivation) over reducing negative ones (e.g., anxiety), to promote overall well-being (Dewaele et al., 2019; MacIntyre et al., 2019). Likewise, these results align with those of past studies showing that students with higher psychological well-being are often more autonomous, seek personal growth, are open to new experiences, or maintain positive relationships with others (Dewaele et al., 2019; Ryff, 1989). Future studies can examine the mechanisms by which psychological well-being or enjoyment affect L2 WTC. Notably, daily IDLEs measured via ESM and via the survey were similar (ESM: $M = 4.86$ hours, range = .94 to 8.88 hours; survey: 4.57 hours, range: .50 to 10.00 hours), ameliorating concerns over the accuracy of self-reported survey data (Arndt et al., 2023). Hence, recall bias did not alter these self-reported IDLE frequencies.

6. Implications

6.1. Theory and methodology

Our study has potential theoretical, methodological, and pedagogical implications. Our findings contribute to theory, specifically second language acquisition
theories regarding L2 WTC, informal language learning, and positive psychology. For instance, we found more trait- and state-L2 WTC antecedents, gained a better understanding of how IDLE, trait L2 WTC, and state L2 WTC interact, and revealed that enjoyment and psychological well-being, rather than anxiety, were related to trait- and state-L2 WTC. Students who studied English longer, had higher language proficiency, had higher pre-psychological well-being, had a higher standard deviation L2 WTC over 14 days, or had higher teacher appreciation showed higher L2 WTC outside of class. Students with higher language proficiency, higher pre-WTC digitally, or higher post-psychological well-being reported higher L2 WTC in digital settings. The ESM findings revealed that students with higher teacher appreciation, currently engaged in productive IDLE or both receptive and productive types of IDLE, currently experienced greater L2 enjoyment or currently experienced greater well-being had higher L2 WTC now. Of particular significance, psychological well-being was linked to trait and state L2 WTC only in non-academic settings (out-of-class, digital), which yields two theoretical insights. First, incorporating psychological well-being into the social and individual context (Layer VI) and affective-cognitive context (Layer V) of the pyramid model enhances our comprehension of L2 WTC. Secondly, academic (in-class) versus non-academic (out-of-class and digital) settings require separate L2 WTC pyramids. Hence, our study advances L2 WTC theory.

We employed rigorous data across three contexts and advanced statistical analyses (Al-Hoorie & Vitta, 2019). Unlike past studies with only cross-sectional laboratory data (MacIntyre & Wang, 2021) or classroom studies (e.g., Dewaele & Dewaele, 2018; Lee & Liu, 2022), our real-time ESM data from participants in their natural real-life surroundings reduced respondents’ recall biases, increased ecological validity, and rendered our findings more readily applicable to real-world situations. Furthermore, we applied advanced statistics (system of equations, multilevel cross-classification ordered logit) to address subtle, temporal research questions; avoid unwarranted assumptions; and increase the accuracy of our results.

6.2. Pedagogy

Our findings help teachers develop strategies for supporting students’ out-of-class learning activities and emotional well-being, in order to improve their L2 WTC both at a particular moment and overall. As students with higher psychological well-being had higher L2 WTC out-of-class, digitally, and now, teachers who create and maintain supportive learning environments might help students become more autonomous and open-minded to engage in meaningful conversations with
others (Dewaele et al., 2019; Dewaele & Dewaele, 2019). Doing so might enhance students’ mental health, positive outlook on life, immediate (state) L2 WTC, and trait L2 WTC in authentic situations.

Students with higher L2 enjoyment, particularly teacher appreciation, showed higher L2 WTC out-of-class and now. Thus, teachers can be caring, friendly, and supportive to build a strong teacher-student rapport (Chiu & Chow, 2011), which might improve L2 WTC outside of the classroom and at a given moment. Both trait-like and state-like L2 anxiety were unrelated to L2 WTC at both the trait and state levels. Hence, teachers can focus on cultivating positive emotions and attributes (e.g., enjoyment and motivation) to improve L2 WTC rather than reducing negative emotions (e.g., anxiety; Dewaele et al., 2019; MacIntyre et al., 2019). As students with higher pre-existing L2 WTC or standard deviation L2 WTC over 14 days showed higher L2 WTC in-class, out-of-class, or digitally, teachers can design and encourage activities that help students experience a variety of easy and difficult communicative situations (e.g., talking to friends in familiar settings vs. talking to strangers in unfamiliar settings). Such activities can help students improve their trait L2 WTC, especially outside the classroom (Lee & Hsieh, 2019). Students’ current participation in IDLE, particularly production-oriented IDLE, increased the likelihood of L2 WTC now. Therefore, teachers who integrate IDLE into L2 activities might immediately boost state L2 WTC (Lee et al., 2022).

7. Limitations and implications for future research

This study’s limitations include self-reported data, single region, limited contexts, and limited explanatory variables. Unlike one-time, retrospective surveys, ESM yields real-time data with greater accuracy about moods, thoughts, and behaviors in natural settings. As ESM is not widely used in L2 research, future researchers can use it to reduce potential recall bias and enhance the external validity of their findings. However, ESM data relies on self-reporting, which still remains subjective. Hence, future studies can capture objective data (e.g., with open-source LangTrack App, Arndt et al., 2023). Moreover, this study only examined EFL learners in one region, so future studies can include students from other regions and countries. Also, as this study only examined state L2 WTC at the global level, future studies can investigate state L2 WTC across three different settings (in-class, out-of-class, digital). Lastly, this study only examined a small set of explanatory variables, so future studies can include other L2 related variables (e.g., motivation and emotion constructs; Al-Hoorie, 2017, 2018; Derakhshan et al., 2021; Kruk et al., 2022).
8. Conclusion

Our research proposed and showed how EFL learners’ demographics, L2 enjoyment, L2 anxiety, psychological well-being, and IDLE were associated with L2 WTC at the trait (in-class, out-of-class, and digital) and state (momentary) levels. Our survey data showed that older students, those with higher language proficiency, or those with higher pre-WTC digitally reported higher *L2 WTC in the classroom*. Among these student participants, those who had studied English for longer durations, had higher language proficiency, felt greater pre-psychological well-being, showed higher variation in L2 WTC, or felt that their teacher appreciated them more reported higher *L2 WTC out of the classroom*. Students with higher language proficiency, higher pre-WTC digitally, and higher post-psychological well-being reported higher *L2 WTC in digital settings*.

Our ESM data indicated that students who felt that their teacher appreciated them were currently engaging in productive IDLE, were currently experiencing more L2 enjoyment, or currently felt greater well-being had higher *L2 WTC at the moment*. Our findings contribute to second language acquisition theories such as L2 WTC, informal language learning, and positive psychology, as well as teachers’ strategies for encouraging out-of-class learning activities and improving emotional well-being. These strategies can help EFL students improve their L2 WTC in general and in the present.

Acknowledgements

The study was funded by the Central Reserve Allocation Committee (CRAC) at EdUHK (grant number 1-55-0A429).
References


Modelling trait and state willingness to communicate in a second language: An experience . . .


APPENDIX A

Ancillary analyses

Table A1 Correlations, variances, and covariances of student variables in the lower left, diagonal, and upper right matrices

<table>
<thead>
<tr>
<th>Student variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Post-WTC in-class</td>
<td>0.33</td>
<td>0.18</td>
<td>0.25</td>
<td>0.74</td>
<td>0.08</td>
<td>0.68</td>
<td>0.05</td>
<td>0.20</td>
<td>-0.07</td>
<td>-0.14</td>
<td>-0.15</td>
</tr>
<tr>
<td>2 Post-WTC out-of-class</td>
<td>0.37</td>
<td>0.68</td>
<td>0.50</td>
<td>0.10</td>
<td>-0.19</td>
<td>0.88</td>
<td>0.18</td>
<td>0.12</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.15</td>
</tr>
<tr>
<td>3 Post-WTC digitally</td>
<td>0.49</td>
<td>0.69</td>
<td>0.79</td>
<td>0.34</td>
<td>-0.21</td>
<td>0.81</td>
<td>0.13</td>
<td>0.36</td>
<td>0.02</td>
<td>0.00</td>
<td>0.13</td>
</tr>
<tr>
<td>4 Age</td>
<td>0.52</td>
<td>0.05</td>
<td>0.15</td>
<td>6.06</td>
<td>1.03</td>
<td>1.25</td>
<td>0.34</td>
<td>0.35</td>
<td>-0.39</td>
<td>-0.59</td>
<td>-0.32</td>
</tr>
<tr>
<td>5 Studying English duration</td>
<td>0.13</td>
<td>-0.22</td>
<td>-0.23</td>
<td>0.40</td>
<td>1.11</td>
<td>0.00</td>
<td>-0.21</td>
<td>-0.05</td>
<td>-0.12</td>
<td>-0.26</td>
<td>-0.34</td>
</tr>
<tr>
<td>6 Language proficiency</td>
<td>0.76</td>
<td>0.69</td>
<td>0.59</td>
<td>0.33</td>
<td>0.00</td>
<td>2.38</td>
<td>-0.07</td>
<td>0.22</td>
<td>-0.18</td>
<td>-0.15</td>
<td>-0.27</td>
</tr>
<tr>
<td>7 Pre-well-being</td>
<td>0.54</td>
<td>0.22</td>
<td>0.62</td>
<td>0.22</td>
<td>-0.07</td>
<td>0.22</td>
<td>0.20</td>
<td>0.42</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>8 Pre-WTC digitally</td>
<td>0.54</td>
<td>0.22</td>
<td>0.62</td>
<td>0.22</td>
<td>-0.07</td>
<td>0.22</td>
<td>0.20</td>
<td>0.42</td>
<td>0.00</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>9 SD WTC</td>
<td>-0.40</td>
<td>0.11</td>
<td>0.08</td>
<td>-0.49</td>
<td>-0.34</td>
<td>-0.36</td>
<td>0.19</td>
<td>-0.01</td>
<td>0.10</td>
<td>-0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>10 Post-teacher appreciates me</td>
<td>-0.40</td>
<td>-0.02</td>
<td>-0.01</td>
<td>-0.41</td>
<td>-0.41</td>
<td>-0.16</td>
<td>-0.34</td>
<td>-0.01</td>
<td>-0.17</td>
<td>0.35</td>
<td>0.08</td>
</tr>
<tr>
<td>11 Post-well-being</td>
<td>-0.42</td>
<td>0.28</td>
<td>0.23</td>
<td>-0.21</td>
<td>-0.52</td>
<td>-0.28</td>
<td>0.58</td>
<td>-0.01</td>
<td>0.41</td>
<td>0.21</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Table A2 Correlations, variances, and covariances of student in time variables in the lower left, diagonal, and upper right matrices

<table>
<thead>
<tr>
<th>Student variables</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 WTC now</td>
<td>1.45</td>
<td>0.06</td>
<td>0.08</td>
<td>0.23</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td>2 Pre-teacher appreciates me</td>
<td>0.08</td>
<td>0.45</td>
<td>0.05</td>
<td>0.04</td>
<td>0.17</td>
<td>0.11</td>
</tr>
<tr>
<td>3 IDLE productive now</td>
<td>0.12</td>
<td>0.14</td>
<td>0.26</td>
<td>-0.06</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>4 IDLE both now</td>
<td>0.24</td>
<td>0.07</td>
<td>-0.14</td>
<td>0.63</td>
<td>0.17</td>
<td>0.09</td>
</tr>
<tr>
<td>5 Enjoy now</td>
<td>0.70</td>
<td>0.22</td>
<td>0.03</td>
<td>0.18</td>
<td>1.28</td>
<td>1.01</td>
</tr>
<tr>
<td>6 Well-being now</td>
<td>0.72</td>
<td>0.15</td>
<td>0.06</td>
<td>0.11</td>
<td>0.82</td>
<td>1.21</td>
</tr>
</tbody>
</table>