Danping Wang* and Martin East

Integrating translanguaging into assessment: students’ responses and perceptions

https://doi.org/10.1515/applirev-2023-0087
Received April 29, 2023; accepted April 30, 2023; published online May 26, 2023

Abstract: This paper explores how beginners in a second language (L2) perform on and perceive an online writing test that is designed based on the notion of translanguaging. The test was administered during emergency remote teaching when many L2 courses navigated creative solutions to online testing. Situated in an ab initio Mandarin Chinese course in New Zealand, 163 students’ first-time digital compositions in Chinese and responses to an immediate follow-up survey on their translanguaging practices were analysed as part of evaluating a new assessment design. Students’ digital compositions demonstrated purposeful translanguaging in assessment conditions, judiciously negotiating their existing linguistic knowledge when completing the task. The writing assessment showed augmented task completion when learners’ trans-semiotic repertoires were recognised as a legitimate resource for identity expression. The survey found that most students supported the creative design that integrated digital multimodal composition and translanguaging, replacing the monolingually-focused handwriting-based test tasks. Some students were sceptical of the translanguaging approach and found it unexpected, unnecessary, and inauthentic. The study suggests that L2 writing test design might incorporate translanguaging as a creative and transformative assessment facet to genuinely engage beginning learners in meaningful writing tasks when their proficiency level is limited.

Keywords: assessment; digital multimodal composition; messy creativity; second language writing; translanguaging

*Corresponding author: Danping Wang, The University of Auckland, Auckland, New Zealand, E-mail: danping.wang@auckland.ac.nz. https://orcid.org/0000-0002-5077-5678
Martin East, The University of Auckland, Auckland, New Zealand, E-mail: m.east@auckland.ac.nz. https://orcid.org/0000-0003-3681-5028

Open Access. © 2023 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.
1 Introduction

Translanguaging is the phenomenon whereby users of a second language (L2) draw on linguistic resources both within and beyond the L2 (which may include aspects of learners’ first language or L1) to maintain communicative effectiveness. A recent trend in research in applied linguistics focuses on challenging the conventional understanding of languages as exclusively made up of bounded linguistic features (Otheguy et al. 2015; Toohey 2019). In particular, the introduction of translanguaging in L2 education has fundamentally challenged the structuralists’ static view of the boundaries of languages as socially constructed for “nation-state interests” (García and Li 2014: 17), as well as the monoglossic hegemony and orthodox practices dominating language teaching for decades (García and Anne Kleifgen 2020). Many studies have shown that integrating translanguaging into language education liberates students from the constraints of monolingual/single language focused policies and practices driven by neo-colonial ideologies and creates a plurilingual space for all learners to make sense of their learning experiences (Cenoz and Gorter 2021; García and Kleifgen 2020). Challenging the monolingual orthodoxy, translanguaging is both a “new linguistic reality” (García and Li 2014: 29) and a creative way of being, acting, and languaging in different socio-cultural contexts in which people can learn to use and live with their full semiotic repertoires (Jones 2020). Following Li and Lin (2019), who called for pushing limits and breaking boundaries in language education, Darvin (2020) asserted that more research is needed to examine how the paradigmatic shift introduced by translanguaging facilitates unlocking creativity in language education.

Despite the fact that translanguaging has been incorporated into classroom teaching as a creative aspect of pedagogical practice (see, e.g., Cenoz and Gorter 2021; Lau and van Viegen 2020), it is rarely explored in the realm of assessment or standardised testing (Schissel et al. 2019). This is not necessarily surprising. When testing is viewed as a means to measure or determine levels of proficiency in a given target language, translanguaging arguably introduces an irrelevant variable into the assessment, thereby confounding measurement. Taking account of translanguaging practices would require a broadening out and redefinition of the constructs of interest in the assessment. Shohamy (2011), for example, acknowledged that L2 assessment remains predominantly monolingual (i.e., target language oriented). From this perspective, students are not ‘allowed’ to use non-target language to answer questions or complete, for example, a writing task. That is, as Shohamy (2011: 421) argued, L2 assessment has been “based on monolingual and homogenous constructs that do not enable other languages to ‘smuggle in’.” Allowing students to use their L1 in their L2-focused tests undermines the purpose of such tests and could
be viewed as highly unprofessional or unethical by many educational institutions and teaching professionals. Otteguy et al. (2015) pointed out that translangauing has been severely restricted in mainstream language assessment practices, where students’ linguistic proficiency is confined to testing in a target-language-only monolingual design. Chalhoub-Deville (2019: 472) noted that “reluctance among test developers to engage in ML [multilingual] assessment is one of the main challenges for adopting a flexible approach to translangauing” in testing situations. García and Li (2014: 133) also lamented the fact that “standardised assessments are usually administered in one language only”. This may be profoundly troubling for L2 learners who are at the initial stage of learning when their linguistic knowledge remains limited, and their responses in the target language are also necessarily constrained.

Singleton and Flynn (2021: 10) argued that “translangauing is unlikely to be effective in the initial stages of language learning” due to learners’ unbalanced language skills. Their argument drew on Cen Williams’s Welsh immersion approach of teaching two separate linguistic groups who used one language at a time (García and Kleifgen 2020), which is a form of parallel monolingualism. L1 and García (2022) reiterated that translangauing is not about shuttling linguistic codes between one’s L1 and L2 or using L1 in L2. It emphasises a holistic trans-semiotic repertoire where one’s language knowledge and skills ‘messily’ co-exist (Swain 1985). Moreover, when students learn an L2, they are not driven to achieve L1 or near-L1 standards for career or social economic status (Ushioda 2017), therefore the curriculum’s expectation informed by the one-language-at-a-time pedagogy may represent a mismatch between what students desire and what they experience in language classrooms (García and Kleifgen 2020).

Ushioda (2017: 469) called for “shifting away from SLA [second language acquisition] frames of reference (concerned with progression toward proficiency in a particular language) in favour of a ‘linguistic multi-competence’ framework” that allows us to better understand the emerging multilingual identity of L2 learners and their communication skills as complex social beings. In particular, students at the initial stage of learning a new language are at a critical period of meaning-making and negotiation as they develop a multilingual self. Failure to make a meaningful connection with their existing life experience, personal aspirations, and diverse motivations may result in dissatisfaction with or even discontinuation of language learning (García 2009). Therefore, it is vital to explore creative approaches that might unleash the transformative potential of translangauing in L2 teaching and assessment designs. More importantly, the project to integrate translangauing into language assessment must take a student-oriented approach to ensure that creative designs are well received by students and effectively used to assess students’ communication skills.
This paper sets out to present how a group of absolute beginners in a Chinese as L2 programme at a New Zealand university performed on and perceived an online writing test that was designed based on the notion of translanguaging. The paper begins with a critical revisit of literature on theories and concepts around monolingualism in language testing and assessment design, and then explains how translanguaging and creativity are interpreted for the purposes of this study. Some key contextual factors are presented to explain the backgrounds in which the creative assessment design was developed. Students’ written outputs and written feedback on the test design are presented and discussed.

2 Literature review

2.1 Multilingualism in language assessment

One of the most influential SLA theories that informs a communicative approach to language teaching is the Output Hypothesis (Swain 1985). The Output Hypothesis asserts that language learners must be pushed to produce output in the L2 so that they can develop the desired lexical and grammatical accuracy. Students’ L2 output represents the outcome of their L2 learning. A historically entrenched emphasis on monolingual (i.e., target language) output is arguably based on the assumption that learners have the same goal for learning a new language as for learning and using their L1 – to use it on par or close to the L1 standard. From an assessment perspective, this means that students’ output proficiency is judged against benchmarks (such as the Common European Framework of Reference [CEFR] or the ACTFL standards) that include L1 or near L1 proficiency as a top-tier goal. That is, the construct of communicative competence (see, e.g., Canale 1983; Canale and Swain 1980) informs the facets of proficiency that we might wish to see and examine in a given communicative test, and that construct traditionally focuses on aspects of the target language without the influence of L1. Leung and Scarino (2016), however, highlighted the importance of reconceptualising the goals and outcomes of learning an L2. They emphasised recognising the diversity of goals that people from different backgrounds learning languages might have in diverse settings and who may be interested in developing different capabilities and achieving different outcomes.

This paper follows Shohamy (2011), who advocated for a recognition of multilingualism as a construct in language testing. This perspective challenges the traditional view of language as a “monolingual, homogenous, and often still native-like construct” (Shohamy 2011: 419). By integrating a heteroglossic language ideology within language assessments, we are able to view dynamic forms of languaging as an asset for elicited performances in assessments while also being
attuned to engagement with languages (Schissel et al. 2019). Another purpose for incorporating multilingual constructs into language testing is to recognise and reflect a fuller linguistic and semiotic repertoire for every individual learner, including those at the beginning stage of learning a new language. It is with this in mind that translanguaging is defined as “the deployment of a speaker’s full linguistic repertoire without regard for watchful adherence to the socially and politically defined boundaries of named languages” (Otheguy et al. 2015: 283). Furthermore, Toohey (2019: 937) pointed out that “knowledge/knowing and language/languaging are relational, processual, and entangled”. Language educators and researchers need to reorient their attention to recognise that the boundaries between named languages are not rigid and abandon their efforts in maintaining a static, binary, and instrumentalist view of language learning. Toohey (2019: 950) also argued that L2 teaching needed to take account of new insights and practices that would allow L2 learners to “participate (perform) before they were competent”. This paper builds on the premise that it is time for us to integrate a broader model of language use into language testing to unlock the learning potential of all learners, especially at the initial stage of learning.

Taking assessment into account will ensure that translanguaging becomes and remains a sustainable framework that teachers can confidently implement throughout the entire learning process (Schissel et al. 2021). Translanguaged multimodal assessment opens spaces for meaningful communication by stimulating students’ interest and participation in completing meaningful tasks. If we were to embrace a translingual framework as the foundation of the entire language learning process, we need to extend our attention to its roles in designing and implementing assessments and push our efforts beyond the current theoretical and pedagogical ‘only target language’ basis. We suggest that designing writing assessment tasks through a translanguaging lens will enable students to demonstrate not only what the new language knowledge means to them individually, but also to connect the new linguistic forms with their existing linguistic repertoires. It also enables test developers to assess more broadly what students know and can do both conceptually and linguistically without limits imposed by external societal definitions of the standard. However, although an increasing number of studies have pointed out the potential of using translanguaging in assessment, García and Li (2014: 133) found that “these tests have not been developed”, at least not as fully as we might like, for L2 assessment.

Some recent studies have shown that the translanguaging approach can be integrated successfully into assessment design. For example, in the linguistically diverse state of Oaxaca in Mexico, Schissel et al. (2021) explored a translanguaging assessment design through an action research case study with 40 pre-service language teachers. The study took into consideration the weightings of the
translanguaged test and drew on a rubric to ensure multilingual students could be assessed equitably and fairly. In another study in U.S. public schools, Lopez et al. (2017) presented a digital classroom assessment system that allowed late-arriving emergent bilingual students to demonstrate their content knowledge and skills by drawing on their entire linguistic repertoires. The latter study also discussed how test developers and teachers could help mediate the translanguaging mode in classroom assessment. Lopez et al. used a computer-based testing platform to integrate translanguaging into their assessment design, and this seems to have offered a practical and promising way to further advance efforts in promoting multilingualism in a world where pen-and-paper testing may be replaced by computer or cloud-based testing in the foreseeable future. Shohamy and Pennycook (2019: 36) also suggested that language tests should incorporate not only multilingualism but also “multimodal and multisensorial assemblances” in order to capture the full scope of resources that students use to engage in real-life communication, which is frequently digital and highly mobile due to the rapid development of technology. The need for creative solutions to the traditional notions of testing language proficiency has become more pressing lately, especially since many educators and institutions are more open to digital transformations after two years’ remote emergency teaching as a consequence of the pandemic.

2.2 Creativity in assessment design

In a special issue on creativity in language learning and teaching, guest editor Rodney Jones (2020) incorporated the notion of translanguaging in moving our creative efforts forward. He pointed out that previous endeavours and considerations regarding creativity in language education are confined within the framework of dominant monolingual ideologies. In an earlier publication, Jones (2018) argued that creativity resides in an honest engagement with the ‘messiness’ of most situations in which people are trying to learn new knowledge or make meanings. According to Jones (2018: 82), messy creativity is an acknowledgment of the ultimate nature of language as a messy system, in which “contradictions, contingency, and indeterminacy” are the norm. Specifically, in applied linguistics, Jones (2018: 82) described messy creativity as

a kind of creativity that shakes us out of our comfortable assumptions about the way applied linguistics ought to be carried out, and challenges us to develop ‘messy’ methods to confront this messy creativity, methods that go beyond trying to ‘make sense’ of it through traditional conceptual categories and attempt to approach it from the less traditional perspectives of embodiment and entanglement, affect and action.
Our current research on ‘messiness’ and creativity focuses on developing tools and concepts that make messiness seem neat and orderly, rather than confronting messiness as a natural state or even enabling messiness as legitimate language practice.

The translingual and trans-semiotic messiness may be described as ‘embodied creativity’. Such a paradigm-shifting view of creativity may be viewed as a theoretical response to the translangaging reality in L2 education (Wang 2020), often involving enactment of one’s full repertoire to make meaning in the entire circle of language teaching, learning, and assessment. However, existing research has found that L2 assessment is largely missing out on theoretical advancement illuminated by a translangaging lens.

Integrating translangaging into language assessment requires a creative design that can engage learners in participating in meaningful tasks and can allow them to freely express their minds and identities in the process of demonstrating their evidence of language learning. So-called task-based language assessment (TBLA) may potentially offer a favourable advantage that enables learners to draw on their own resources as they engage in a real-life communication task (East 2021: 494; Norris and East 2021). Indeed, this was used as an assessment design principle in a study conducted by Schissel et al. (2018) on creative multilingual testing design for multilingual speakers. Closely linked to a TBLA approach, formative assessment data can help teachers identify students’ areas of strength or weakness (Ellis 2003; Ke 2006) and understand what the new languaging experience can embody through individualised learning evidence. Overall, being creative involves using existing knowledge or skills in a particular subject or context to experiment with new possibilities in the pursuit of valued outcomes.

A further weakness in the innovation in assessment arena has been the limited attention that has been given to assessment design and innovation in languages other than English. Despite some scholars having explored creative teaching and assessment designs (e.g., García et al. 2017), almost all research has been carried out with students learning English as L2. Most students in existing translangaging research are newly arrived migrants (Lopez et al. 2017), refugees (van Viegen 2020), and minoritised or marginalised youth (Yilmaz 2021), as well as endangered-language speakers (Hickey et al. 2014). The predominant focus on minority speakers in the Global North has been noted by Turner and Lin (2017), who remarked that translangaging should go beyond being a subaltern concept for Eurocentric issues, and that more studies are needed to move translangaging theory forward. The relative lack of research attention on L2 learning in the overall translangaging advocacy reflects the hegemony of English in applied linguistics research (Macedo et al. 2015) and results in a noticeable gap between the theory advancement led by English teaching and the ground level practices of multilingual education. There is
also a growing concern that globalisation, and global English, have resulted in an increasingly low motivation for learning an L2 other than English (Lanvers et al. 2021), and the alarmingly low enrolments in other languages have reinforced English monolingualism in mainstream education. Because of the language learning crisis, what is even more concerning is that Anglophone countries will have more monolingual English speakers in a world where plurilingualism is common in many other countries. East (2015) has warned us that “monolingualism implies inflexibility, insensitivity and arrogance.” This may aggravate the social problems that we have, and the ongoing humanitarian crises caused by populist decision-makers and that work against L2 learning.

The remote emergency teaching precipitated by COVID-19 has served as a catalyst for making profound and creative changes that have thus far been considered impossible in language teaching and assessment (Bond et al. 2021; East et al. 2023; Wang and East 2020). While language teaching methods have evolved drastically over the past few decades, language assessment design remains largely unchanged, especially in the higher education context. University language courses are normally informed by standard proficiency frameworks with aligned outcomes that determine students’ eligibility to apply for scholarships, internships, and study abroad experiences, and that are also linked to prerequisites for higher level courses that require a certain entry-level proficiency. These outcomes have been frequently measured in traditional ways. Because of the disruptions brought about by prolonged emergency remote teaching, many universities have decentralised their assessment policies for L2 programmes, allowing course directors to make more flexible decisions for assessing students equitably and inclusively. Course developers and teachers now have more freedom to explore creative assessment strategies that provide students with a more interactive and authentic learning experience, an experience that is often not well supported in the centralised paper-based examination condition.

The pandemic has pushed many educational institutions to take digital and even unconventional solutions to deal with widespread emergency remote teaching (Bond et al. 2021). Many frontline teachers have navigated creative digital solutions to redesign or even place their conventional paper-based proctored examinations alongside creative online tests to ensure that, as far as possible, learning objectives can be achieved in fair and reliable conditions (East et al. 2023; Wang and East 2020). The huge shift to remote teaching has posed unprecedented challenges for beginning language courses that cover only limited words, simple grammar structure, and basic cultural concepts. These problems have also prompted new and bold educational innovations in areas that have been an unquestioned norm for decades or even centuries.
2.3 Assessing Chinese writing

One of the impacts on L2 assessment as a consequence of online teaching is teaching and assessing Chinese writing in the digital environment. The linguistic nature of the Chinese language makes the assessment of its writing more challenging and complex than that of English and other alphabetic languages (Xu et al. 2021; Zhang and Lin 2017). This is due to the lack of a phonology-to-logograph link in the Chinese writing system. For European language speakers, orthographical skill and knowledge in Chinese are disconnected from their prior experience of literacy development and writing practice. Particularly for beginners, learning to write Chinese characters by hand has proven to be a considerable challenge (Wang and Li 2022), requiring a tremendous amount of practice time in order to gain long-term memory and retrieve information from doing basic reading and writing tasks. One of the fundamental learning components of a beginner Chinese language course has been to train students to hand-write Chinese characters stroke by stroke.

In writing assessments, beginner Chinese learners are expected to demonstrate their ability to unify the phonological, semantic, and orthographical aspects around one particular character. In completing these assessments, students need to recall the strokes of a character, alongside its meaning and pronunciation. Due to the complexity of the Chinese writing system that requires a prolonged process for initial learners to gain basic skills, beginner Chinese courses often do not have the capacity to integrate informative task-based assessments that develop students’ literacy and communication skills beyond lexical and sentence levels. The centrality of handwriting in Chinese writing assessment has been criticised by many previous learners and has been regarded as a major reason that holds Chinese learners back from being confident users of the language (Gil 2020). It is suggested that future Chinese writing assessment developers should utilise digital solutions to “save time and efforts” for students learning Chinese (Gil 2020: 39).

It is important to note that writing in Chinese is not assessed until students reach Level 3 in the current version of the standardised Chinese proficiency test, Hanyu Shuiping Kaoshi (HSK). In contrast to the various benchmarks set for English writing assessment, the objective of the Chinese writing test is to assess students’ ability to hand-write Chinese characters. Writing tasks in Levels 3 and 4 (equivalent to CEFR B1 and B2) are limited to the sentence level. Essay and story writing are included in Levels 5 and 6 (equivalent to CEFR C1 and C2). Furthermore, the sentence-construction writing tasks in HSK are largely “context-free, with little interaction with how language is used in a real-life situation” (Peng et al. 2021: 329). It is also apparent that Chinese writing assessment for non-L1 speakers has operated under a monologic ideology, with focus on the accuracy and harmony of the script.
The global pandemic has profoundly challenged print-based handwriting-centred Chinese language teaching and learning. In a recent survey of Chinese instructors in U.S. postsecondary institutions, Xu et al. (2021) found that teachers usually welcomed the digital approach to writing instruction during the emergency remote teaching. Zhang’s (2021) research found that learners using keyboard typing were able to perform significantly better in writing assessment of Chinese. Students not only produced longer and more accurate sentences but also had better Chinese character retention rate. However, Wang and Li (2022) pointed out that innovative approaches based on multimodality remain unknown to many Chinese language teachers.

2.4 Second language education in New Zealand: key contextual factors

The present exploration builds on the premise that translanguaging perspectives and practices are bound by macro-meso ideological, sociocultural, and institutional contexts (Shohamy 2011). Given deeply entrenched monolingualism in mainstream education (East 2009; Major 2018), the creativity of the multilingual approach in language teaching and testing can be easily overshadowed or even inhibited by many contextual factors, such as controversies around the translanguaging approach, strict assessment policies in tertiary institutions, and the decline in students’ motivation in learning an L2.

Recent research has explored the implications of, and creative solutions with, translanguaging pedagogies in L2 teaching in the New Zealand context. Wang (2020) surveyed 237 students from three universities and found a large discrepancy between the prescribed monolingual principles in language teaching and the experienced translanguaging reality in day-to-day classroom interaction. In another study related to L2 learning, Walker (2018) analysed the translanguaging patterns in learners’ online interactions and concluded that translanguaging was used as an affordance strategy to expand learners’ semiotic repertoires for meaning co-construction. Focusing on creative learning materials design, Seals and Olsen-Reeder (2020) presented a case of using translanguaging materials to support children from indigenous and marginalised minority groups as part of revitalising and maintaining vulnerable languages. It seems that the translanguaging space co-created by teachers and students is mostly well-received by language learners and users in New Zealand. However, none of these studies has tapped into assessment.

Decisions on major assessments in formal education are usually informed by centralised policy and subject to the educational institution’s infrastructure.
and environment. Assessment within a university is more or less a series of administrative and pedagogical legacies to ensure that the mainstream knowledge system is safeguarded. It also has the power to shape how alternative knowledge can be adapted to fit into the mainstream system and conform with the assessment regulations that are deemed effective for assessing English and European languages. For example, in the authors’ university, the assessment policy required all stage-one (first year) courses to be partially formally assessed under examination conditions. It has been stipulated that, for these courses, “at least 50% of course assessment must occur in invigilated settings, normally achieved through formal examinations and/or tests sat under examination conditions” (Assessment [Coursework, Tests and Examinations] Policy 2019, Item 20). Moreover, the university does not have the capacity to administer computer-based examinations on campus for large-scale courses and prior students in the focal course in question here had traditionally been completing high-stakes paper-based final examinations.

Given the continuous decline in L2 learning in New Zealand, we contend that the rationale to set up centralised high-stakes assessments based on the L1 speaker standard merits a reconsideration. A recent study found that learners of Chinese as L2 at a leading university in New Zealand were found to be more interested in learning the language for personal interest than for material gain or career enhancement (Wang 2020). This finding is in line with de Burgh-Hirabe’s study (2019) which found low instrumental motivation among Japanese language learners at university. Furthermore, recent studies have attested to the fact that learners of languages other than English are more often motivated by personal development to “expand their meaning-making repertoire for identity expression rather than wanting to gain proficiency in a particular language” (Zheng et al. 2020: 796).

At the authors’ university, an internal on-site investigation was conducted with 1,188 former students who took Chinese between 2014 and 2017. In total, 141 valid responses were collected. It was found that 47% of students were primarily motivated to study Chinese to fulfil a “desire for self-satisfaction”, such as travel, entertainment, stepping outside the comfort zone, or because they simply enjoyed learning an Asian language. This was followed by 23.1% of students motivated by a “desire for career enhancement” and 11.9% who had a “desire to be integrated with the Chinese culture”. The findings indicated that the majority of students at the university were not learning Chinese for career-related purposes. Having high-stakes assessments irrelevant to real-life communication could therefore do a disservice when it comes to maintaining students’ learning motivation. In the same research site, looking at oral assessment for beginners in Chinese as L2, Wang (2019a) found that students had been assessed in a decontextualised, form-focused, teacher-led, and memorisation-based setting, often arranged in a nerve-wracking one-on-one meeting in a lecturer’s office, twice a semester. This echoes what Shohamy (2007)
described as her unpleasant experiences in taking language tests. She argued (2007: 142), “[i]t was tests that were responsible for turning the enjoyment and fun of learning into pain, tension, and a feeling of unfairness”.

The question therefore arises around how L2 assessment designs can be adapted or transformed when emergency remote teaching realities come up against traditional monolingual ideologies. It is also important to take into consideration students’ responses to new and creative assessment designs (see, e.g., East 2016). Towards this purpose, this study aims to shed light on the two following research questions: (1) How do beginning L2 learners perform on an online writing test designed based on the concept of translanguaging? (2) How do they perceive the new test design?

3 The study

The present study was carried out in the first author’s course for beginners in L2 Chinese during the emergency remote teaching when many L2 courses were compelled to navigate creative solutions to online testing. Data were collected by the first author primarily in the context of teaching and learning in the first semester of the emergency remote teaching and to gather evaluative pilot evidence that might inform future pedagogical practice. The data are drawn on here, in an anonymised and largely aggregated way, essentially as means to discuss the lessons that were learned in the pedagogic context.

In the current site, new and practical assessment policy has emerged to accommodate the new online teaching environment: examinations will be digital, remote, and open book by default, unless exemptions for either in-person examinations or invigilation are approved. New methods of assessment have been explored since the revised assessment policy was introduced. However, although tests and examinations were shifted online, they remain high-stakes, monolingual, and decontextualised. In addition, instructors have been concerned about effective measures to ensure that they can adequately assess students’ learning outcomes while also minimising opportunities for cheating (Daniels et al. 2021). A traditional understanding of assessment seems to permeate despite attempts at innovation.

The massive shift to emergency remote teaching has posed unprecedented challenges to classroom-based conventional L2 courses, where pen-and-paper invigilated assessments have been compulsory. The impact is more profound at the elementary L2 level that only covers limited vocabulary, a small number of simple grammar structures, and very basic cultural concepts. The assessment presented in this evaluative study is one of five short tests designed for the five textbook-based lessons covered in this 12-week course. Each test is worth 2% of the total grade.
The purpose of setting up these five short tests has been to provide low-stakes assessments for students to keep track of their learning progress and to prepare them for the final paper-based test. Students took the first short test under the conventional mode before the course had to be shifted online in March 2020. Figure 1 provides an example of the test paper.

The test paper shows the main activity, which is to recognise and recall the new characters students have learned. Students are tested on their memorisation skills in decoding and encoding new characters. However, “remembering” is at the bottom level of the cognitive domain of learning, which only requires lower order thinking skills. Also, this paper-based test requires completion by hand, and the level of difficulty for this test significantly increased due to the handwriting requirement.

The data reported here were drawn from one question from the second short test (Short Test 2, administered online), and an immediate follow-up survey with all students enrolled in the course (n = 163). Short Test 2 was developed under the revised university assessment policy. The time limit for completing the entire test of five questions (multiple choice, short answers) for Short Test 2 was 1 h, including the time allocated for unexpected technical problems as a consequence of web-based testing. Students were allowed only one attempt to complete the graded assessment within a 24-h window. The last question was ‘Composition’, a short writing task as shown in Figure 2. The stimuli were chosen with two considerations in mind: first, a family photo of the characters from The Incredibles was provided to students to create a

![Figure 1](image_url)

**Chinese 100/100G Beginning Modern Chinese 1**

<table>
<thead>
<tr>
<th>Short Test 2 (2%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last Name:</td>
</tr>
<tr>
<td>AU-ID:</td>
</tr>
<tr>
<td>First Name:</td>
</tr>
<tr>
<td>Stream:</td>
</tr>
</tbody>
</table>

Please write Characters for Pinyin in row one (1-5) and write Pinyin for Characters in row two (6-10).

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>gè</td>
<td>de</td>
<td>dà</td>
<td>méi</td>
<td>yǒu</td>
</tr>
<tr>
<td></td>
<td>(measure word)</td>
<td>(particle)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>这</td>
<td>妈妈</td>
<td>孩子</td>
<td>她</td>
<td>男</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1:** The conventional test paper used for short test 2.
context for them to write about family, which is the main learning content to be tested; second, the Mulan film poster was chosen due to its popularity among students – the majority of the film was shot in New Zealand.

The students’ responses were used to address the first research question of how students perform on a test that incorporates translanguaging opportunities. The writing task aimed to follow an informative task-based language assessment design (East 2021: 132). It gives students an opportunity to demonstrate their learning by completing a meaningful task, but also by drawing on their existing language repertoire, whether in L2 or L1. That is, a clear instruction was given that words in English could be used when needed to enable communication to occur. The textbook used in this course is Integrated Chinese, Book One. At the time of the assessment, students had only completed the first two lessons, which covered very limited words related to greetings, self-introductions, and family members. Students in the course had received three weeks of face-to-face handwriting training while teaching remained on-campus. One week before the test, students were given a special tutorial to install a virtual Chinese keyboard to their electronic devices as they prepared to take this test. Students’ answers to the writing task question were collected to provide evidence of their performance in their first-time translanguaged-mode test.
Students’ perceptions of the test experience were gleaned from the survey data and were used to answer the second research question. Using the Canvas Quizzes function, it was possible to insert a survey that required all test takers to participate before submitting their test papers. Students were given an additional 10 min to complete one open-ended survey question: “What do you think of our decision to allow students to use English in completing this short test?” A thematic analysis approach was used to analyse students’ responses. Students’ input was collected and coded based on the patterns of meaning observed within the qualitative data. These codes were then placed under two major themes: supportive of or sceptical towards the translanguaging element in the assessment.

4 Findings

4.1 Students’ performances

It should be noted that typing in Chinese is not as straightforward as typing in alphabetical languages. Students are faced with several challenges as part of the complex process of decoding and encoding the script while composing in Chinese digitally. Typing one character or word in Chinese involves six steps (Zhang and Min 2019): transforming ideas into Chinese words; recalling pronunciation; using Pinyin (the Chinese romanisation system) to transliterate the Chinese word; typing the romanised letters on the keyboard; selecting; and finally confirming the correct character based on its orthography from a list of homophones prompted in an input method interface.

Students’ overall performance on their first translanguaging test were nonetheless deemed to be successful. All 163 students completed the writing task and submitted the test and their survey comments within the time limit. Table 1 shows two examples of students’ writing output. Some emerging findings are discussed below.

4.1.1 Purposeful translanguaging

Students’ outputs did not involve a single case of overusing English when completing this task. Although it was their first time undertaking this kind of test, it seems the students were fully capable of undertaking purposeful translanguaging in an assessment context. In this study, purposeful translanguaging refers to a judicious use of one’s linguistic repertoire to complete a writing task in an examination condition. Their compositions demonstrated that they understood that the purpose of the test was to demonstrate their newly acquired Chinese language skills.
As can be seen in the two examples in Table 1, students’ translanguaging practices only occurred at a lexical level and with limited English words for proper nouns. It was clear that students were fully aware of their responsibility to provide an adequate output in Chinese as L2 for teachers to evaluate their learning achievement. The fear of opening the floodgate to non-target language use or having insufficient output evidence for evaluation was not found in the evidence (Wang 2015). In fact, the task was designed with the expectation that students could complete it entirely in Chinese, except for some of the characters’ names that students may find necessary to include. Example 1 in Table 1 is a case where the student only used Mulan to address the audience and kept the rest of the composition in Chinese. Although the question rubric did not prescribe a limitation on use of English in answering this question, it was found that all students only resorted to English substitutes for words that were not covered in the course but were found useful in making a meaningful sentence.

### 4.1.2 Augmented task completion

It seems the translanguaging design provided an expanded opportunity for students to maximise their efforts in task completion. Although the task only asked for 20 characters to demonstrate their learning, all students’ compositions went beyond this task completion requirement. In the initial stage of Chinese language learning, they showed their capability to engage in this kind of writing task and produce relatively long L2 writing output beyond the sentence level. In this short composition, every student had assembled almost all new words (such as kinship terms including father, mother, brother, and sister) and key grammar points (such as to be, yes-or-no
question, numbers, and measure word) covered in the first two lessons. As Example 2 in Table 1 shows, the student strived to use several types of sentences that demonstrated a desire to communicate with Mulan, and genuine interest in knowing more about her. Overall, students’ writing went beyond the expectations of the task. Similar results were found in Lopez et al.’s study (2017: 6) which showed that such computer-based engaging test design could “elicit more meaningful and relevant evidence of the students’ knowledge and skills”.

### 4.1.3 Visualised messy creativity

The digital translanguaging composition task design made it possible for students to enter into the ‘messiness’ of learning and using another language to expand their existing linguistic repertoires. In addition to opening the possibility for language learners to legitimately go through an embodied experience in connecting new knowledge with their available resources, the test unlocked beginner learners’ potential to play with these resources creatively. This kind of creativity has long been discouraged because language learners are expected to produce correct (but not necessarily creative) answers. García and Li (2014) argued that translanguaging is a creative process, emphasising the spontaneous performances of multilingual speakers as they engage in assemblages of the semiotic resources that are made available. Much of Li’s (Li 2016, 2020) work on creative translanguaging practices is exemplified with cases involving English and Chinese. It is hoped that ‘messy creativity’ can also become part of language testing so that creative translanguaging practices can be encouraged at the beginning stage of learning Chinese as L2. By integrating the translanguaging approach into language testing, students’ creativity could be embodied in their individualised languaging processes and products.

### 4.2 Students’ perceptions

Students’ comments on the translanguaging design provided insight into their performance on the writing task. The total number of comments (171 supportive comments and 42 sceptical comments) outnumbered the total student number (163) because some students commented from both sides. Overall, it seems that students were supportive of the digitally-mediated translanguaging assessment design.

#### 4.2.1 Supportive comments

The thematic analysis found 12 categories under the positive group. Table 2 provides an overview and examples for each category. On the whole, students provided highly
informative comments to explain why they enjoyed this assessment task. In general, students found the task engaging and suitable for beginner learners as a means to participate creatively and confidently, even at the initial stage of L2 learning. It enabled students to express their ideas and identities relatively freely without ‘turning off’ their existing repertoire or being overwhelmed by the stress of writing in a new language like an L1 speaker. With the translanguaging design, students could compose a relatively long piece with richer information while maintaining

Table 2: Supportive comments on the translanguaging approach.

<table>
<thead>
<tr>
<th>Categories of comments</th>
<th>Number of comments</th>
<th>An example of students’ comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interesting, engaging, fun</td>
<td>45</td>
<td>It was engaging because they made me think</td>
</tr>
<tr>
<td>Relevant to the current level</td>
<td>39</td>
<td>It is interesting and relevant to the content that we learnt in class</td>
</tr>
<tr>
<td>Reduce stress level</td>
<td>22</td>
<td>Students can focus on writing when they are not pushed to remember everything at the same time</td>
</tr>
<tr>
<td>Reduces cheating</td>
<td>12</td>
<td>It's great! this way it reduces the temptation to go and check Google translation</td>
</tr>
<tr>
<td>Fair and inclusive</td>
<td>10</td>
<td>I think your decision to allow students to use English to help with a Chinese essay is fair. It is helpful for those students who are not yet used to using a Chinese keyboard</td>
</tr>
<tr>
<td>Increases fluency</td>
<td>9</td>
<td>It is a very good decision to allow students to use English in writing. This is because it helps us to complete the essay more fluently</td>
</tr>
<tr>
<td>Maintains freedom</td>
<td>7</td>
<td>I believe that using English, especially in the early stages, is acceptable as it allows for some freedom for us to express our ideas with limited Chinese words</td>
</tr>
<tr>
<td>Allows complexity</td>
<td>7</td>
<td>I think this was a good decision. It allows us to show the markers what we want to express without disrupting the flow of the sentence/paragraph. By using English words as a substitute for the words we do not know how to say in Chinese, it gives us a chance to make more complex sentences</td>
</tr>
<tr>
<td>Identity expression</td>
<td>6</td>
<td>I don't think it acts as a crutch. It allows us to express ourselves more fully despite our limited Chinese. Otherwise, everyone would end up writing the same story</td>
</tr>
<tr>
<td>Increases confidence</td>
<td>5</td>
<td>I think that when we are allowed to use English in our Chinese essay, it gives us a bit more confidence. I was quite surprised I'm writing in Chinese just after a few weeks of studying</td>
</tr>
<tr>
<td>Encourages long answers</td>
<td>5</td>
<td>As there is an optional condition to use English, it makes me feel to put longer sentences on the answer sheet</td>
</tr>
<tr>
<td>Truthful learning evidence</td>
<td>4</td>
<td>It can help the teachers better understand our real proficiency level not only our memorisation skills</td>
</tr>
</tbody>
</table>
a certain degree of fluency and complexity. Specifically, three aspects are worthy of more extended discussion for potential impact on future language teaching and assessment.

Firstly, it seems that students were mostly impressed by the novel and engaging nature of the new test design. It may be that this is the result of the combined effect of the online testing (reducing the pain or difficulty of handwriting), multimodal stimuli (providing multisensory meaning-making experience), communicative task-oriented composition design (creating the space for individuality and real language use), and, finally, the open attitude towards students’ existing repertoire (empowered by the translanguaging opportunity). It seems the integrated task design enhanced the level of engagement with the composition task for beginner learners. The creative test design fosters students’ creativity.

Secondly, composition as an open-ended writing task provided opportunities for students to share their ideas and identities. It seems that, in this kind of context-rich open-ended task, students have a stronger sense of control over their own creative stories. Students’ desire for identity expression is supported by the diverse stories that emerged, and no identical products were found. In a review study, Smith et al. (2021) found that most existing works on digital multimodal composing support emergent bilingual students’ identity expression. The potential translingual identity enacted by the translanguaging design in testing cannot be fully unleashed in a traditional summative assessment that requires standardised answers as evidence of proficiency. It is important to nurture a plurilingual and translingual identity (Zheng 2017) at an early stage so that L2 users can develop and experience stronger ownership of the language they learn instead of being forever constrained by the boundaries between the named languages and conventional modes for assessment and teaching.

Thirdly, at the operational level, students found that the translanguaging design reduced the temptation to resort to machine translation (which could be perceived as an illegitimate use of a resource or ‘cheating’), even though the test was administered under the ‘open book’ policy. This claim was supported by students’ writing samples. In marking students’ compositions, there appears to be no evidence of students using Google translate or copying from each other. Because students were allowed to use English for words they had not learned or could not remember, they knew their compositions were not expected to include words and expressions more advanced than their actual levels. For words not covered in the course, they could express themselves with English. It seems that students believed that the translanguaging assessment design allowed them to honestly showcase their languaging processes and served as “truthful learning evidence” (quoted from Table 2, the last category of positive comments) rather than a perfect imitation of L1 standards. The reduced propensity to cheat may come from a reduced level of anxiety and stress that
beginners experience in the process of producing meaningful content when drawing on a wider linguistic repertoire.

4.2.2 Sceptical comments

The study found six categories of comments showing sceptical attitudes towards the trans languaging design, as shown in Table 3. Students were not consulted or informed beforehand about this short test which aimed to pilot the trans languaging design. Their comments clearly showed that some students were not comfortable sitting a test that allowed them to do things that had been outlawed in their past educational experiences. It seems their monolingual educational experiences and prior language learning had shaped their beliefs so that they saw trans lingual practices as ‘unnecessary’, ‘unexpected’, and ‘inauthentic’, while the L1 standard was regarded as the norm. Interestingly, most sceptical comments were shorter than the positive comments, usually one sentence long, and lacked further information to explain standpoints.

5 Final remarks

The primary purpose of the evaluation reported here was to pilot a writing task that could include trans languaging as a strategic resource in the context of modifications

<table>
<thead>
<tr>
<th>Categories of comments</th>
<th>Number of comments</th>
<th>An example of students’ comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confusing</td>
<td>14</td>
<td>Not sure how to incorporate English with Chinese in an essay so was more confusing</td>
</tr>
<tr>
<td>Native-speakerism</td>
<td>12</td>
<td>Students should be encouraged to find the appropriate Chinese equivalent wherever possible, to stretch themselves in an attempt to sound more native, as opposed to just substituting words they do not know but want to use, with English equivalents</td>
</tr>
<tr>
<td>Challenging</td>
<td>8</td>
<td>It is more challenging as it pushed me to make long sentences that can easily go wrong</td>
</tr>
<tr>
<td>Unnecessary</td>
<td>4</td>
<td>It didn’t seem necessary to me. Apart from the names, it is possible to write the whole thing in 中文。</td>
</tr>
<tr>
<td>Unexpected</td>
<td>2</td>
<td>I Found it to be quite unexpected. I have never been allowed to use English in language tests</td>
</tr>
<tr>
<td>Unauthentic</td>
<td>2</td>
<td>I Am indifferent about using English. It makes it easier but is obviously less authentic</td>
</tr>
</tbody>
</table>
to courses as part of transition to a digital environment, and, in that context, to elicit stakeholder feedback from the students. Overall, students’ sceptical feedback was helpful in making informed decisions concerning future teaching and learning, as well as improvements for curriculum and assessment design. Following the successful implementation of the first translanguaging mode task, the course has developed more digital multimodal composition tasks to further unlock students’ creativity. This study offers an account of how beginner learners of an L2 perform on and perceive a creative assessment integrating translanguaging opportunities.

This pilot represents one of the first attempts to implement a multilingually-oriented assessment approach in a beginning level Asian language course with a relatively large group of students in a higher education context. Students in the course were absolute beginners who had studied a new language for just 20 h (four weeks, five days a week) by the time they undertook this creative test. Drawing on the translanguaging design, the evaluation provides promising evidence that initial assessment can be developed to “allow learners to participate (perform) before they were competent” (Toohey 2019: 950) rather than to delay or deny their ability to express themselves when they start to explore a new medium of communication. Many efforts are being made to design construct valid tests to measure learners’ proficiency within the predominant monolingual domain. It is time for language assessment research to extend its efforts to reconceptualise and integrate multilingualism as a new (or at least expanded) construct in developing many more creative assessment approaches that can genuinely address contemporary educational issues and that can help “bring forth a world distinct from what we already are” (Toohey 2019: 938).

The findings from this evaluation suggest that translanguaging as a creative assessment strategy is on the whole well received by initial learners for a range of reasons. Based on students’ writing samples and comments, it is evident that translanguaging is a practical, zero-cost, and open-minded strategy that can be easily incorporated when redesigning and reinventing traditional assessments. In integrating translanguaging into assessments, students will have expanded opportunities to create meaningful language products in an engaging and lower-stress space and communicate in empowering ways from the initial stages of L2 learning, rather than keeping silent under the gaze of L1 speaker standards until they are competent. Under the translanguaging lens, learners’ creative potential can be better fulfilled (Wang 2019b), and their existing repertoires are recognised as legitimate resources in their overall ‘entangled’ language learning experience. Their messy creativity and embodied languaging can only be materialised in the tireless efforts to push the limit and break the boundaries in language teaching (Li and Lin 2019), and push it further by incorporating the notion of translanguaging into assessment design for initial learners of a language other than English.
We return to the argument we made at the beginning: communicative tests of language proficiency are built on theoretical frameworks of communicative competence (e.g., Canale 1983; Canale and Swain 1980), interpreted from a single language perspective and benchmarked against standards relative to L1 norms. This is what teachers and assessors are used to and the approach reflects an understanding that assessment is there to measure target language proficiency. However, the Canale and Swain model includes, as strategic competence, the strategies L2 users might draw on “to compensate for breakdowns in communication due to insufficient competence or to performance limitations and … to enhance the rhetorical effect of utterances” (Canale 1983: 339). Taking this classic perspective into account, translanguaging may potentially be promoted as a legitimate strategy within a communicative understanding of competence. Furthermore, as McNamara (2001: 333) made clear, we need to look critically at our practices and the assumptions that underpin them as part of “the evolution of the field.” While language teachers are still struggling to understand the supportive roles of students’ existing repertoires in L2 learning, the evaluation presented here provides some evidence that it is possible to move beyond just a creative idea.

Creativity is nonetheless not possible without constraints (Darvin 2020). Jones (2020: 540) reminded us that “the point of translanguaging is not that language users can do anything they want, but rather that they are able to bring to bear a wider range of resources to respond to the conventions and contingencies of whatever situation they find themselves in.” The findings show that students were aware that the primary purpose of the translanguaging task for a graded test was a measurement of writing proficiency in L2 Chinese. It is important to note that this evaluation was carried out with adult students in an elite university who were studying in a formal course that counted towards their degrees. This serves as a constraint to their languaging behaviour in a test condition. Even so, the overall meso-level contextual readiness and policy flexibility gave a green light for such a creative test to appear on the university’s learning management system. Many previous studies have reiterated that the institutional expectations regarding language proficiency can work against the creative power of a translingual framework (e.g., Otheguy et al. 2015; Schissel et al. 2021). Due to the two years of on-and-off lockdown teaching, course directors had opportunities to explore new testing principles and methods, as well as novel theoretical frameworks, as catalysts for reinventing assessment designs and strategies. It is hoped that the assessment policy will maintain a certain degree of flexibility when on-campus teaching is resumed. Some temporary creative solutions might, however, need to be adapted to become the ‘new normal’ even after the pandemic crisis is largely behind us.
Bearing in mind the above constraints, this study (although small-scale in nature, due essentially to the limited writing samples available) highlights that potentially radical changes can take place in Chinese language teaching and assessment. This single case shows that initial learners are fully capable of participating in virtual communication in Chinese when their existing linguistic repertoire is recognised as a legitimate languaging resource. With fast technological development and important theoretical breakthroughs in applied linguistics, Chinese language teaching and testing have space to find more creative solutions to transform our current knowledge from a monolingual to a multilingual view, from handwriting to typewriting, and from the pen-and-paper mode to multimodal computer-mediated composition. Learners of Chinese as L2 should be allowed to experiment with the new script without fear of being penalised for not producing the desired quality of the script as L1 speakers might. Teaching in the digital context suggests that we treat students as content producers, creators, developers, performers, builders, and community participants, rather than as end-users passively receiving knowledge and skills in a process they are not empowered to participate in. Due to the deeply entrenched native-speakerism in Chinese teaching, translanguaging remains unknown to many Chinese teachers (Wang 2019b, 2020), resulting in such creative insights and practices being overlooked, discouraged, and dismissed. However, as Schissel et al. (2019: 377) argued, it is critical for scholarship in L2 teaching “to keep pace with changes in society and to align with the knowledge base in Applied Linguistics.”

Without a doubt, many teaching professionals will have questions about the significance of adopting the translanguaging approach in their teaching and assessment. It is anticipated that such an approach will appear to be rather bold or even debatable, or even be resisted or heavily criticised in mainstream language education, particularly for Asian languages. There is a need for more research into how translanguaging and creative assessment design can be used in second language teaching, and more importantly, how to transform language teachers and teacher educators’ perceptions in second language teaching.

This study is limited by its focus on a single case and the possible influence of power relations in eliciting students’ opinions of the translanguaging design. The survey was integrated into the Canvas Quizzes, and students’ comments were linked to their writing responses. It is important to stress again that this evaluation was undertaken first and foremost as an exploration of a pilot procedure, data from which would primarily inform the first author’s future teaching, learning and assessment practices. Future research on students’ feedback might be administered anonymously in the context of more conventional research and can be undertaken in a more fine-grained way.
Overall, this evaluation offers several important implications for future research on integrating translanguaging into language teaching and assessments. We contend that the discussion on integrating the translanguaging approach into language assessment should move beyond the classroom and the instructional sphere (Baker and Hope 2019) and should inaugurate enhanced pedagogies and new assessment designs.

Acknowledgements: This publication was supported by the Marsden Fund Council from New Zealand Government funding, managed by Royal Society Te Apārangi. The project number is UOA1925.

References


