



Research Article

¹Nanjing Xiaozhuang University, Nanjing, China; ²Xi'an Jiaotong-Liverpool University, Suzhou, China and ³Bilingual Cognition and Development Lab, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies, Guangzhou, China

Cite this article: Wang, J., & Wei, R. (2023). Is bilingualism linked to well-being? Evidence from a big-data survey. *Bilingualism: Language and Cognition*, 1–11. <https://doi.org/10.1017/S1366728923000603>

Received: 26 September 2022
Revised: 11 July 2023
Accepted: 14 July 2023

Keywords:
bilingualism; well-being; individual difference; foreign language proficiency; Putonghua proficiency

Corresponding author:
Rining Wei;
Email: Rining.Wei@xjtlu.edu.cn

Abstract

In applied linguistics generally and bilingualism research in particular, psychological variables remain a much under-investigated sub-category of individual differences compared with cognitive ones. To better understand the under-researched psychological effects of bilingualism, this study investigated well-being, a psychological construct, based on a big-data survey. Drawing upon a national survey ($N = 12,582$), we examined the influence of bilingualism (operationalised as foreign language (FL) proficiency) and 13 sociobiographical variables (e.g., socio-economic status, SES) on well-being. Among these 14 initial independent variables, perceived social fairness, SES, and health emerged as important predictors for well-being, with FL proficiency and national language (NL) proficiency as potentially important predictors; crucially, FL proficiency was more important than NL proficiency. As the first systematic attempt to link bilingualism with well-being, our study advocates (1) a more holistic perspective towards *language* (including NL and FL(s)) in any bilingual context and (2) fuller use of effect sizes.

Introduction

How foreign/second language (L2) learners' cognitive features, psychological attributes, and other individual differences (IDs) affect target language learning have long intrigued researchers from applied linguistics in general and bilingualism research in particular (Luo & Wei, 2021). In the past five decades, much research has investigated a variety of learners' IDs, cognitive and non-cognitive (e.g., psychological) (Kong & Wei, 2019; Oxford & Ehrman, 1993; Pimsleur, 1966; Robinson, 2002; Rubin, 1975; Skehan, 1989). Psychological variables (e.g., resilience, see Wei et al., 2022b) represent an under-researched sub-category of IDs vis-a-vis their cognitive counterparts, such as working memory (Wen & Skehan, 2021) and aptitude (Li & Zhao, 2021); accordingly, recent calls for more research attention on the former ID sub-category have been made (Luo & Wei, 2021; Walton, 2022; Wei & Gao, 2022).

In the field of bilingualism, there are two different lines of research on psychological IDs (Luo & Wei, 2021). The first line has its roots in theoretical models (e.g., Baker, 1996; Gardner, 1985) that tend to regard IDs as the independent variables (IVs) and language proficiency variables as the dependent ones. In contrast, the second line of inquiry has attracted scholarly attention since the late 2000s (e.g., Dewaele, 2012; Dewaele & van Oudenhoven, 2009); studies along this newer research line (e.g., Grey & Thomas, 2019; Wei & Hu, 2019) tend to treat IDs as the dependent variables (DVs), with language proficiency variables being the independent ones. The present study, focussing on the psychological ID of well-being, contributes to the second research line.

Well-being, conceptualised as a psychological ID (Sari et al., 2018), refers to the psychological state where a person 'subjectively believes his or her life is desirable, pleasant, and good' (Diener, 2009, p.1). The present study represents one of the very few attempts to examine this psychological ID among a nationally representative sample of bilinguals¹ from a big-data survey. Several recent studies (e.g., Kang, 2022) have attempted to explore the impact of language on well-being, generating valuable insights such as 'language makes life better' (Zhang & Cheng, 2022). However, as applied linguistics researchers, we need to probe further: higher proficiency in which language (e.g., a first language? a foreign language (FL)?) makes life better (viz. leading to a higher level of well-being) when it comes to bilinguals? This big question is particularly important, as bilingualism is the norm in most regions of today's world including the People's Republic of China (henceforth 'China'). As will be shown

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

This article has earned badges for transparent research practices: Open Data. For details see the Data Availability Statement.

¹A bilingual refers to a person with bilingualism, viz. 'at least partial mastery in more than one language' (Dewaele & Li, 2013, p. 231). This definition of 'bilingualism' is so flexible as the number of languages involved can be a figure no smaller than two. Hence in the present study, we use bilingualism as an umbrella term in a broad and inclusive sense so that it encompasses trilingualism, multilingualism, and other variants. But in several places in this paper, 'multilingualism' is inevitable because the authors of the cited sources use this wording.



below, extant studies about the potential link² between language and well-being have unfortunately ignored foreign language(s) and focused exclusively on the Chinese language. The present study endeavours to narrow this research gap.

The present study is the first systematic attempt to link language with well-being based on a nationally representative sample from China by considering BOTH the national language AND English (a FL). English is more relevant to the post-reform generation (i.e., people born in 1978 or later) in China (Wei & Su, 2008; Wen & Zhang, 2020). Accordingly, our study focuses upon the respondents belonging to this generation ($N = 3471$) from a recent wave (2017) of the Chinese General Social Survey (CGSS), which utilised a representative national sample ($N = 12,582$).

Besides adding new knowledge to well-being, an under-investigated psychological ID, our study contributes to four further areas. First, in the current big-data era, any research effort to utilise representative samples, especially samples from big-data surveys readily available in the public domain, would be most valuable (Wei et al., 2022a). Specifically speaking, while Zhang and Cheng (2022) attempted to link language with well-being based on data from the CGSS (a big-data survey), replications (especially with an improved design) are needed because the value of replication research (Marsden et al., 2018) is increasingly recognised in applied linguistics generally and bilingualism research in particular. Second, we advocate a more HOLISTIC perspective towards language (including the national language and FL) in any bilingual context. This advocacy for broadening the research scope for the link between LANGUAGE and a focal psychological ID (e.g., well-being) will prove to be worthwhile (see our analysis below). Third, our study can paint a more comprehensive picture of the influence of bilingualism and other sociobiographical factors on well-being by adopting a more refined approach based on hierarchical regression (Wei et al., 2020) supplemented by dominance analysis (Mizumoto, 2023). Fourth, this study increases our understanding of the psychological profiles of bilinguals in China, an under-examined English as a foreign language (EFL) context (Wei & Hu, 2019), where there are more than 390 million English-knowing bilinguals (Wei & Gao, 2022).

Literature Review

Psychological effects of bilingualism

In the past decade, much research has explored the effect of bilingualism on different psychological IDs, be they negative (e.g., anxiety, see Jiang & Dewaele, 2020), neutral (e.g., extroversion, see Chen et al., 2015), or positive (e.g., L2 grit, see Wei et al., 2020). Most of the relevant studies are primarily quantitative; when reviewing quantitative research, we focus on effect size, which is more important than the statistical significance level (p) (Wei & Gao, 2022; Wei et al., 2021).

Several studies (e.g., Dewaele & MacIntyre, 2014) have explored the influence of bilingualism on foreign language enjoyment (FLE), a positive psychological ID similar to our focal variable, well-being (Wang et al., 2021). For instance, based on a sample of 189 high school students in Greater London, Dewaele et al. (2018, p. 684) found (1) number of languages known was 'unrelated' to FLE ($F(5, 185) = 1.6$, *no effect sizes reported*, $p > .05$); and (2) level in the FL (operationalised as self-evaluation on a four-point scale)

had a statistically significant effect on FLE ($F(3, 185) = 4.4$, $p < .005$, $eta^2 = .066$, Cohen's $d = .53$). While it was commendable that these researchers provide more than one effect size measure for the focal effect, unfortunately they did not provide effect sizes for statistically non-significant results; in contrast, the present study will report effect sizes for both statically significant and non-significant results following the recommended statistical practices (Larson-Hall & Plonsky, 2015).

Some studies (e.g., Dewaele & van Oudenhoven, 2009) have explored the influence of bilingualism on tolerance of ambiguity (TA). For instance, Chen's (2004) survey of 193 EFL students in China found that TA was statistically significantly correlated with their English proficiency (a measure of bilingualism) ($r = .407$, $p < .001$). A more recent study in the Chinese EFL context is Wei and Hu's (2019) survey of 260 English-using bilinguals; an important finding was that bilingualism accounted for 1.4% of the TA variance.

Several recent studies (e.g., Liu & Wang, 2021) have examined the influence of bilingualism on grit, another positive ID variable similar to well-being. Khajavy and Aghaee (2022) focussed on perseverance of effort and consistency of interest – namely, two components of grit, in 226 EFL learners at one private language institute in Iran; one major finding was that perseverance of effort was positively correlated with English proficiency (operationalised as final scores at the end of the semester in the school-based test) ($r = .16$, $p = .014$). Based on a sample of 462 Chinese EFL learners, Wei et al. (2020) found that self-rated English proficiency (a measure of bilingualism) could explain up to 3.9% of the variance in L2 grit, with the effect size range being 1.1%–3.9% in their hierarchical regression models. These researchers utilised Wei and Hu's (2019) effect size benchmark system, where .5%, 1%, 2%, and 9% respectively represent the small, typical (medium), large, and very large cut-offs for the effect size measure R^2 .³ Regarding the influence of bilingualism, its effect size upper limit (3.9%) exceeded the 'large' benchmark and its lower limit (1.1%) exceeded the 'typical' one; hence the effect of bilingualism on L2 grit was deemed important by Wei et al. (2020). Wei and Hu's (2019) benchmark system, which has been adopted for effect size interpretation in recent studies (e.g., Dewaele & Botes, 2020), will also be adopted in the present study.

The effect of bilingualism on well-being

In connection with learning in general, the effect of learning on well-being has been widely discussed both theoretically (e.g., Desjardins, 2008; Wang & Wang, 2008) and empirically (e.g., Cuñado & de Gracia, 2012). As regards language learning in particular, FL learning that leads to the development of bilingualism can become 'a source of well-being' (Proietti Ergün & Ersöz Demirdağ, 2022, p.3). Put differently, bilingualism can exert influence on well-being (to varying degrees in different contexts), as will be illustrated by the studies reviewed below.

Two important lines of inquiry concerning bilingualism and well-being merit attention. The first line of research has a particular focus upon minority groups, such as immigrants (e.g., Kim et al., 2012; Lee et al., 2021) and minority ethnicities (e.g., Shields & Price, 2002) in a particular polity. The other line of inquiry that is emerging does not focus exclusively on minority groups, which will be delineated below.

²For instance, the link between English proficiency and well-being has been 'seldom' explored (Zhang et al., 2021, p. 2) (see also Note 4).

³After R^2 is un-squared, the small, typical (medium), large, and very large cut-offs for the effect size measure r are (roughly) .07, .10, .14, and .30.

Grey and Thomas (2019) investigated the link between bilingualism (operationalised as language proficiency) and well-being among 210 female citizens of the United Arab Emirates. The language proficiency variable was a self-report measure, with a higher score indicating a higher level of Arabic language proficiency. Well-being was gauged with the World Health Organization's well-being index. This cross-sectional correlational study identified a nearly 'very strong' association ($r = .26, p < .01$) between language proficiency and well-being.

Kim et al. (2012) explored the settlement experiences of 46 adult immigrant learners of English from three different first language (L1) backgrounds: Chinese, Arabic, and Vietnamese. All these learners were undertaking English language study in one English programme for adult migrants at the time of investigation. The data, which comprised audio-recordings, transcripts, and field notes of semi-structured interviews, were collected over a one-year period. On the one hand, the researchers rated the participants' English proficiency (an indicator of bilingualism) as the ability to speak this additional language in the semi-structured interviews on a nine-point scale according to the IELTS criteria. On the other hand, these researchers evaluated the participants' perception of well-being by assigning one of the three scores—1 being low, 2 medium, and 3 high—to each selected interview by scrutinising its entire transcript, its field notes, and the socio-biographical profile for the participant in question. This study identified a 'strong' association ($r_s = .34, p = .04$) between bilingualism (operationalised as English proficiency) and well-being. The two studies reviewed above produced relatively large effect sizes (approaching the 'very large' benchmark of .30); however, these results probably over-estimated the strength of association between bilingualism and well-being because bivariate analyses (e.g., correlation) tend to generate inflated effect sizes (Wang et al., 2022; Wei et al., 2022a). One solution to this problem is to employ multivariate analyses that will generate a more comprehensive picture in terms of effect sizes.

Although most studies used (simple) bivariate analyses, some recent studies have started to employ multivariate analyses that will generate a more comprehensive picture. For instance, employing hierarchical regression, Khawaja et al. (2016) investigated the factors (e.g., English proficiency) affecting migrants' well-being among Chinese-speaking Taiwanese migrants who settled in Australia. The participants ($N = 271$) completed a questionnaire battery available in both Chinese and English. Well-being was measured by the Satisfaction with Life Scale (SWLS) (Diener et al., 1985), a five-item instrument on a seven-point Likert-scale (a higher score indicating a high level of well-being). English language proficiency (a measure of bilingualism) was self-rated by the participant (scale unreported). Results indicated that age ($\beta = .131, p = .001$), years of residence ($\beta = .222, p = .004$), social support ($\beta = .440, p = .288$), and bilingualism ($\beta = .104, p = .444$) were positively linked with well-being, accounting for 17% of the well-being variance. Although these authors reported the conventional effect size β , which is largely un conducive to cross-study comparisons, it would be optimal if two or more types of effect sizes (e.g., ΔR^2 for regression) could be reported (Luo & Wei, 2021) so as to facilitate comparisons across studies. To overcome this limitation, in our regression analyses below, we reported the crucial effect size index ΔR^2 , which can facilitate cross-study comparisons, along with the conventional indexes (B and β).

Drawing on the AsiaBarometer Surveys 2006 and 2007, the primary part of Zhang et al. (2021) investigated the effect of English

proficiency (a measure of bilingualism) on subjective well-being; the secondary part of their study was a robustness check 'conducted with an alternative dataset targeted at a specific country among the 14 countries or regions' (p. 12). Their sample comprised 14,811 respondents in 14 East and Southeast Asian countries or regions, including China. These researchers used hierarchical regression to control for six sociobiographical variables (e.g., gender) and ascertain the effect of English proficiency on well-being. They found that (1) English proficiency explained .9% of the well-being variance ($\beta = .127$), and (2) the six control variables altogether accounted for 3.3% of the well-being variance. Finding (2) painted a vague picture of the contribution of each of these six predictors to the variance in well-being, which could be improved by Wei et al.'s (2020) more refined version of hierarchical regression that will be adopted by the present study. Notwithstanding this limitation, it is commendable that Zhang et al. (2021) reported more than two types of effect size indexes (e.g., β and R^2), among which ΔR^2 is most conducive for cross-study comparisons as mentioned above. As will be seen below, Finding (1), which suggested that English proficiency was positively linked to well-being at a level slightly below the typical benchmark (1%), will be usefully compared with results from the present study.

In the Chinese context, a handful of studies have explored the effect of language variables on well-being (e.g., Kang, 2022). Two studies are most relevant. One is the secondary part of Zhang et al.'s (2021) research reviewed above, which used the 2017 CGSS data for the robustness check (see their Table 5). Five control variables (gender, age, marital status, education, and socioeconomic status (SES)), two IVs (English listening proficiency and English speaking proficiency), one mediator (income satisfaction), and one DV (happiness⁴) from merely 4,032 participants were used in regression analyses 'after screening the missing and abnormal values of the selected variables' (p. 12). Some interesting findings were: (1) English listening proficiency was statistically significantly correlated with happiness ($\beta = .063, p < .01$) and the explanatory power of this language variable was weaker than that of SES ($\beta = .201, p < .01$ see Model 11 in their Table 5); (2) English speaking proficiency was also statistically linked with happiness ($\beta = .062, p < .01$) and again the explanatory power of this language variable was weaker than that of SES ($\beta = .201, p < .01$, see Model 14 in their Table 5).

Despite their use of two effect size measures (e.g., β and R^2), three major limitations remain. Firstly, a prerequisite for mediation analyses is that the links respectively between the IV, DV, and mediator are established (Agler & Boeck, 2017; Judd & Kenny, 1981)⁵. Unfortunately, Zhang et al. (2021) failed to take into account this prerequisite; it appears that they rushed to apply mediation analyses before establishing the link between IV (English proficiency) and DV (well-being)⁶. Secondly,

⁴Happiness was measured by the questionnaire item 'Do you think you live a happy life?'. This measure was identical with that used to gauge well-being by Zhang and Cheng (2022). Hence throughout the manuscript we used the wording 'well-being' rather than 'happiness' unless in direct quotes from Zhang et al. (2021).

⁵Put differently, previous studies and/or extant theories need to lend support to the three links in a mediation model: the link between the IV and DV, the link between the IV and mediator, and the link between the mediator and DV. In connection with the first link, if there are insufficient empirical and/or theoretical evidence to support the association between the IV and DV, more studies are needed before pursuing any meaningful mediation analyses.

⁶Zhang et al. (2021) observed that 'previous studies have seldom explored the linkage between one's English proficiency and happiness' (p. 2, emphasis added). Under this circumstance, shouldn't the researchers conduct more empirical studies (similar to the present study) first to explore this linkage rather than rushing to introduce/involve potential

although the valid questionnaires in the 2017 CGSS totalled 12,580, only 4,032 were used in the second part of Zhang et al. (2021), who failed to account for the information from 8450 questionnaires (61.17%, over two-thirds of the original sample)⁷. This practice was problematic because it incurred significant information loss and cast down on the trustworthiness of the design of the relevant item⁸. Thirdly, although it is commendable that Zhang et al. (2021) reported more than one type of effect size indexes (e.g., β and R^2), they only utilised β s in discussing the link between English proficiency and happiness. However, the effect size index β is inconducive to comparisons across different studies; in contrast, ΔR^2 is much more useful in cross-study comparisons (Wang et al., 2022). To overcome this particular limitation, one useful approach is calculating effect size ranges based on hierarchical regression analyses (Wei et al., 2020) and/or ranking the relative importance of each predictor in regression models via dominance analysis; the present studies endeavoured to attempt both methods (see **Analytic Strategy** for details).

The other most relevant study is Zhang and Cheng's (2022) paper entitled 'Language Makes Life Better: The Impact of Mandarin Proficiency⁹ on Residents' Subjective Well-Being' that drew upon two earlier waves (2012 and 2015) of the CGSS. As the title suggested, this study focused on Putonghua proficiency, although it aspired to respond to the big question 'is *language* linked to well-being'. It turned out that these researchers ONLY used Putonghua proficiency data from the CGSS surveys despite the availability of English proficiency data. To address this limitation, the present study, which draws upon the 2017 CGSS, will not confine *language* data to the national language *only*. Put differently, our study endeavours to address the (potential) role of FL(s) in the data collection process¹⁰.

mediators? Regarding when and to what extent researchers should involve potential mediators for the focal link between the IV and DV, Agler and Boeck (2017) warn that 'there are simply too many alternative explanations to consider' in real-life data analysis (p.5). Considering the need for parsimony and a desire to avoid false positives, we suggest that researchers examine the link between the IV and DV with sufficient research effort prior to focusing on generating additional explanations with mediators. All in all, concurring with Zhang et al.'s (2021) above observation, we used it as a starting point for our study.

⁷After checking the CGSS dataset, we found over 8000 valid questionnaires have missing values on income satisfaction, which was selected as the mediator in Zhang et al.'s (2020) analysis. It was unclear why these authors used this, out of so many potential factors (see the 13 IVs in Zhang and Cheng (2022), as mediator. It was also questionable whether the time was ripe to conduct mediation analyses. If these researchers had picked another variable, a much larger sample would have been used in their analysis, which would make full use of information rather than causing significant information loss.

⁸The item used to measure income satisfaction reads 'To what extent do you agree with the following: I'm satisfied with my family's income.'. It has eight options on a six-point Likert scale (including 1 = 'totally disagree' and 6 = 'totally agree') and the other two 'I don't know.' and 'I refuse to answer.'. Despite the inclusiveness of these options, over 8000 questionnaires suffered from missing values on this item. Hence some critical issues surrounding this single-item measurement, such as how resistant the respondents were when invited to answer this question, are yet to be fully explored. This situation is different from the single-item measurement for well-being, which did not incur missing values and was shown to be effective and satisfactory (e.g., Abdel-Khalek, 2006; see also **The Present Study** section).

⁹This wording was used in Zhang and Cheng's (2022) own translation of their paper title. 'Mandarin' goes by different names. It is called *Putonghua* (the common language) in the Chinese mainland, *guoyu* (national language) in China's Taiwan region, and *huayu* (Chinese language) in Singapore (Chen, 1999). We stick to the wording Putonghua throughout this paper.

¹⁰The inclusion of FL(s) was partly inspired by recent studies on psychological IDs similar to well-being which considered FL(s) in their research design. For example, Yin (2021) examined the impact of language proficiency on psychological integration among 569 immigrants in one second-tier city in China; she considered proficiency in the national language (Putonghua), the local dialect, and English (FL).

A most relevant finding from Zhang and Cheng (2022) was that Putonghua proficiency was a statistically significant predictor for well-being. However, it was only reported that the language variable TOGETHER WITH the other 13 non-language predictors accounted for (at most) 19.8% and 20.4% in the variance of well-being respectively for the 2012 and 2015 waves. Such reporting practice reflected one limitation with data analysis, which was also identified in Zhang et al.'s (2021) study reviewed above: Zhang and Cheng (2022) unfortunately did not attempt to ascertain the unique contribution of the predictors (e.g., Putonghua proficiency) to the well-being variance. To overcome this limitation with the data analysis process, our study adopts Wei et al.'s (2020) more refined version of hierarchical regression that helps to gauge each IV's unique contribution.

The links between other sociobiographical variables and well-being

Our study aims to ascertain the extent to which bilingualism, vis-à-vis other sociobiographical variables, is linked to well-being.

As indicated above, in a recent study most relevant to ours, Zhang and Cheng (2022) examined the influence from up to 13 non-language sociobiographical variables (e.g., SES) on well-being. As indicated above, these 13 predictors altogether explained about 20% of the well-being variance in the 2012 and 2015 CGSS waves. Besides this result concerning effect size, three major findings concerning statistical significance included (1) in both CGSS waves, gender, health¹¹, marital status, perceived social fairness, SES and income emerged as statistically significant predictors ($p < .01$) of well-being consistently; (2) in both waves, ethnicity and resident type turned out to be statistically non-significant predictors; and (3) age, years of education, religion and household registration were unstable predictors as the statistical significance for their prediction was not consistent in both waves. As our study also drew upon the CGSS, all of the non-language predictors in Zhang and Cheng (2022) were considered in the analysis below.

The Present Study

Research questions (RQs)

RQ1: To what extent does bilingualism (operationalised as English proficiency) predict well-being?

RQ2: To what extent does bilingualism, vis-à-vis the selected sociobiographical variables (e.g., SES), predict well-being?

Data

The data source for the present study was the 2017 wave of the CGSS, which was the latest available at the time of writing. The CGSS is 'the first' continuous survey project run by academic institutions in the Chinese mainland (Hu & Li, 2019, p. 156). It draws on a nationally representative sample and collects data at the multiple levels of society, community, family, and individual. Although it was described as 'an annual survey started since 2003' (Chen et al., 2021, p. 2), in the past few years it has been conducted every two years; for instance, following the 2015 wave

¹¹This variable, also called 'health condition' in Zhang and Cheng (2022) was an ordinal variable requiring each participant to self-evaluate his/her health conditions on a five-point Likert scale (1 = highly unhealthy, 5 = highly healthy). This variable was retained in the present study as one of the selected sociobiographical variables.

(see e.g., Luo & Wei, 2021), the 2017 wave of CGSS released the collated data to the public in October 2020, whereas the data of the next wave were not released prior to the submission of our paper for peer review. We suggest that using the 2017 wave of CGSS be useful because of two major¹² considerations. First, our 'prototype paper' (i.e., the paper being replicated), Zhang and Cheng (2022) utilised two EARLIER waves (2012 and 2015) of the same big-data survey. Second, although the secondary (and less important) part of Zhang et al. (2021) utilised the 2017 CGSS only to some extent, the primary part of Zhang et al. (2021) utilised the AsiaBarometer Surveys 2006 and 2007. Put differently, the 2017 CGSS wave was used in a supplementary way to support Zhang et al.'s (2021) focus on another big-data survey (similar to the CGSS) conducted over 10 years ago. Conducted through individual interviews and structured questionnaire surveys (Liu et al., 2021), the CGSS endeavours to systematically monitor the changing relationships between social structure and quality of life in urban and rural China (Luo & Wei, 2021). Adopting a random sampling method¹³, the 2017 CGSS questionnaire solicits essential information on more than 700 variables including well-being, English proficiency (a measure of bilingualism), Putonghua proficiency, and a wide range of sociobiographical characteristics (e.g., SES) from respondents. The 2017 CGSS national sample comprised 12,582 Chinese citizens aged 18 or above. Following scholars who examine Chinese generations based on key historical events (Egri & Ralston, 2004; Tang et al., 2017), we defined the group of people born in 1978 or later as the post-reform generation as the year 1978 was regarded as the beginning of China's modernization (Lu & Alon, 2004; Tang et al., 2017). For the purpose of the present study, the focal sample was confined to the respondents born in 1978 or later ($N = 3471$) in the above 2017 CGSS national sample.

Dependent variable

The outcome variable well-being¹⁴ was measured by one item: 'Do you think you live a happy life?' Responses were originally coded on a five-point Likert scale ranging from 1 to 5, where 1 indicated 'very unhappy' and 5 'very happy'. A higher score reflected a higher level of well-being.

This single-item measure has both advantages and disadvantages. But overall speaking, it is 'reliable, valid, and viable' in survey-based studies (Abdel-Khalek, 2006, p.129). The single-item measurement for well-being has been adopted in a series

¹²The two studies illustrated below, which discussed the 'seldom'-explored linkage between language and well-being (Zhang et al., 2021, p. 2), were highly relevant to our study hence they represented our 'major' considerations. There are many examples in social science disciplines other than language-related fields; for instance, the 2013 CGSS data were the focus of Qi et al. (2023); Yan et al. (2023) utilised the 2017 CGSS data, claiming that this wave provided 'the *most recent* data from the CGSS' (p. 4, emphasis added).

¹³This 2017 CGSS data utilised multi-stage stratified random sampling involving three stages. In the first stage, 105 primary sampling units (100 counties/districts and five large cities) were randomly drawn from the national population. In the second stage, four village/neighbourhood committees per county/ district and 80 per large city were chosen sampled from each primary sampling unit. In the third stage, households were sampled from each village/neighbourhood committee; and then one individual was randomly selected from each sampled household for a questionnaire-based interview, resulting in a total sample size of 12,582 individuals (Wang et al., 2021).

¹⁴For this DV, except 3 (.08%) and 3(.08%) participants respectively selecting 'I don't know.' and 'I refuse to answer.' 40 (1.15%) selected '1 = very unhappy', 192 (5.5%) '2', 448 (12.9%) '3', 2177 (62.7%) '4', and 608 (17.5%) '5 = very happy'. In other words, this single-item measurement for well-being did not incur missing values (see also Note 8).

of analyses based on the CGSS (e.g., Ding et al., 2021; Qi et al., 2023); most recently, Yan et al. (2023) reported that the single question on well-being from the 2017 CGSS is 'reliable and effective' (p. 4) in the Chinese context.

Independent variables

There were 14 initial IVs in the present study. These 14 variables were selected for two major reasons. First, the present study was exploratory in nature in that 'no previous study has ever focused on the English-well-being linkage' (Zhang et al., 2021, p.2). Any exploratory study should prioritise selecting factors, which should not incur significant information loss, from a myriad of potential influencing factors; specifically, the present study managed to avoid selecting some variables (e.g., income satisfaction, see the critique of Zhang et al., (2021) in the **Literature Review** section) that unfortunately caused serious information loss from the 700+ variables covered by the CGSS). Second, the present study represented a replication of Zhang and Cheng (2022). A common strategy for variable selection in any replication study is to retain all (if not most) of the IVs from the prototype paper for our (partial) replication; hence our study chose to (1) keep all of the 14 IVs¹⁵ from Zhang and Cheng (2022) and (2) add another factor (English proficiency) because of our adoption of a more HOLISTIC perspective towards language.

Two were our focal IVs (viz. factors of our main interest): English proficiency and Putonghua proficiency. The two items (on a 5-point Likert scale) respectively measuring the respondent's proficiency in English listening and speaking were added up to indicate the overall English proficiency; this newly created variable was a proxy for bilingualism, with a higher score reflecting a higher level of bilingualism. Similarly, the second focal variable, Putonghua proficiency, was measured by adding up the scores from the respondents' ratings on a 5-point Likert scale of their proficiency in both Putonghua listening and speaking; again, a higher score of this newly created variable indicated a higher Putonghua proficiency level.

The other 12 IVs (see Table 1), which might affect well-being (see the Literature Review section), were also considered to facilitate comparison with previous research (e.g., Zhang & Cheng, 2022). For instance, SES was assessed with a five-point Likert scale in response to the question; 'in your opinion, which SES does your family belong to' (1 = 'far below the average level of SES' and 5 = 'far above the average level of SES'); the higher the score, the higher the SES; this measure of SES was the same as used in previous studies (e.g., Zhang et al., 2021; Zhang & Cheng, 2022).

Analytic strategy

RQ1, which explored the association between bilingualism and well-being, was dealt with via simple linear regression. RQ2 examining the influence of the 14 initial IVs on well-being was

¹⁵We acknowledge that it would be most useful for readers if authors could attempt to offer justifications for variable selection. However, neither the authors of this prototype paper nor Zhang et al. (2021) made such attempts. Zhang and Cheng (2022) just reported that they 'considered demographic, economic, social, religious, and health-related factors'; similarly, Zhang et al. (2021) simply wrote that 'some socio-demographic variables and some variables that may influence the data results' (p. 5), without specifying what those 'variables that may influence the data results' are, let alone attempting to justify variable selection. In other words, these authors unfortunately failed to include some justifications for variable selection.

Table 1. Participant Profile

Non-binary Variable	Range	Valid <i>N</i>	Mean (<i>SD</i>)
Well-being	1-5	3465	3.90 (.79)
Perceived social fairness	1-5	3458	2.99 (1.03)
Health	1-5	3418	4.07 (.86)
SES	1-5	3449	2.68 (.68)
English proficiency	2-10	3467	4.18 (1.93)
Putonghua proficiency	2-10	3470	8.01 (1.65)
Years of education	0-19	3469	12.32 (3.80)
Ln income	0-16.12	3081	11.04 (1.52)
Age	18-39	3471	29.61 (5.99)
Binary Variable	Valid <i>N</i>	Valid percentage (frequency)	
Rural (Household registration)	3459	52.67 (1822)	
Married (Marital status)	3471	62.00 (2152)	
Female (Gender)	3471	52.46 (1821)	
Urban (Resident type)	3471	74.30 (2579)	
Han (Ethnicity)	3471	91.76 (3185)	
Atheists (Religion)	3471	91.01 (3159)	

Note: Following Chen et al. (2021), we took the logarithm of 'income' and generated a new variable called 'Ln income' to reduce collinearity.

addressed via hierarchical regression supplemented with dominance analysis. Hierarchical regression helps to ascertain the unique contribution of each IV to the DV-variance (Kong & Wei, 2019; Wei et al., 2020); the order for entering the IVs into regression models is of paramount importance because the DV-variance explained by each IV may vary significantly with the entry order (Wang et al., 2022; Wei et al., 2020); accordingly, researchers should attempt all possible entry orders and offer a range of effect sizes (rather than one single effect size) for each IV unless there are well-established theories to guide the (ideal) entry order (see Wang et al., 2022; Wei et al., 2022a). Dominance analysis helps to 'accurately determine predictor importance in multiple regression' (Mizumoto, 2023, p.195, emphasis added); this analysis generates the R^2 change which occurs when adding one IV to all possible subset regression models and hence identifies the contribution of the IV 'by itself and in combination with other predictors' (Tonidandel & Lebreton, 2011, p.2). In this analysis, the average of R^2 change produced in all possible subsets is called dominance weight; the sum of dominance weights is equal to the total DV-variance-explained (reflected by the effect size R^2) (see Table S2 in Supplementary Material). A reader-friendly version of dominance weights is called rescaled weights; the rescaled dominance weight for each IV reflects the percentage of the DV-variance it contributes to the total DV-variance-explained, and the rescaled values add up to 100% (see Table 4).

We conducted all the statistical procedures (except for dominance analysis) with SPSS 27; the supplementary dominance analysis was performed via a web application developed by Fan (2023). In language-related disciplines, it is only very recently that researchers become increasingly aware of dominance analysis as a viable alternative for estimating predictor importance in regression models (Mizumoto, 2023).

For the sake of brevity, in what follows, we set the statistical significance level at the conventional cut-off ($\alpha = .05$, non-directional). Following the recommended practice of statistics reporting (Sun et al., 2010; Wei et al., 2019), we reported exact p values (with very small p s being reported as $p < .0005$).

Results

RQ1. The link between bilingualism and well-being

A simple linear regression analysis, with bilingualism as the IV and well-being as the DV, generated a statistically significant model ($R^2 = .026$, adjusted $R^2 = .026$, $F = 91.779$, $p < .0005$), suggesting that bilingualism (i.e., English proficiency) explained 2.6% of the well-being variance. The effect size exceeded the 'strong' benchmark (2%) in Wei and Hu's (2019) effect size interpretation system.

RQ2. The influence of the selected sociobiographical variables on well-being

Prior to running hierarchical regression, we conducted two rounds of data checking. The first round was a preliminary analysis, aiming to ascertain which of the initial 14 IVs could be included into regression analysis. The inclusion criterion was that the strength of the association between a predictor and well-being should be stronger than the typical benchmark (viz. $r = .1$) in bivariate analyses. This criterion helps to ensure the principle of parsimony (Leech et al., 2014), which has also been adopted in some recent studies (e.g., Wang et al., 2022; Wei et al., 2022b).

The preliminary analysis that consisted of several rounds of bivariate analyses confirmed that the eight italicised variables in Table 2 were suitable for inclusion into regression analysis. Specifically, seven correlation analyses yielded the effect sizes (in descending order, see Table 2) for the links between the non-binary variables and well-being; one independent-samples t -test revealed the statistically significant difference ($p < .0005$) in well-being between respondents with rural household registration and those with non-rural household registration ($r = .110$, slightly exceeding the 'typical' benchmark); the latter ($M = 3.99$, $SD = .71$, $N = 1635$) had a higher level of well-being than their counterparts ($M = 3.82$, $SD = .84$, $N = 1818$).

The second round of data checking was conducted to ensure that the assumptions (e.g., normality and homoscedasticity) for regression were met. For example, when checking for potential outliers, we conducted several rounds of casewise diagnostic analyses and deleted 53 cases that had a standardized residual greater than 3 or smaller than -3 . Hence the initial sample size ($N = 3471$) was slightly reduced to 3418.

Then we used this revised sample in a series of hierarchical regression analyses with the eight italicised variables (see Table 2) as predictors. Each predictor was entered, one by one, into each of the eight models (or 'blocks' as called in SPSS). A total of eight predictors will generate 40,320 ($8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$) possible entry orders and hence produce up to 40,320 different scenarios (see Table 3 for one sample scenario); put differently, for each predictor, there could be 40,320 different effect size values. All in all, our hierarchical regression analyses generated two crucial findings.

The first finding was that the eight predictors, regardless of the entry orders, explained a total of 11.4% in the well-being variance ($R^2 = .117$). The second crucial finding included the ranges of the effect size ΔR^2 for the eight predictors: perceived social fairness (3.29 - 4.58%), SES (2.08 - 4.86%), health (1.60 - 3.65%), English

Table 2. Links between the 14 initial independent variables and well-being

Variable	Effect size r_s (p)
Non-Binary Variable	
Perceived social fairness	.215(*)
Health	.193(*)
SES	.190(*)
English proficiency	.159(*)
Putonghua proficiency	.146(*)
Years of education	.135(*)
Ln income	.134(*)
Age	-.047(.006)
Binary-variable	
Household registration	.110(*)
Marital status	.079(*)
Gender	.063(*)
Residence type	.046(.089)
Ethnicity	.012(.470)
Religion	.009(.604)

Note: * indicates $p < .0005$. Three digits numbers following the decimal point are kept except for the need to reveal a more nuanced difference.

proficiency (.24 – 2.41%), Putonghua proficiency (.23 – 2.15%), years of education (.03 – 1.68%), income (.01 – 1.18%), and household registration (.03 – .95%). Regarding the first and second predictors, their maximum and minimum effect sizes exceeded the large benchmark (2%), suggesting that they are very important predictors. Regarding the second and third predictors, SES and health, their effect size minimums exceeded the typical benchmark (1%) and their maximums the large benchmark (2%), indicating that they were important predictors. Regarding English proficiency and Putonghua proficiency, their effect size maximums also exceeded the large benchmark (2%), although their effect size minimums dropped below the small benchmark (.5%); this meant that they could be important predictors for well-being. Regarding the last three predictors, their effect size minimums fell below the small benchmark (.5%), and at the same time their effect size maximums did not exceed the large benchmark (2%), suggesting that they might exert negligible effect on well-being and hence were relatively unimportant. Furthermore, these crucial results (e.g., effect size ranges) were confirmed with dominance analysis (see the last column in Table 4). Additionally, the results of dominance analysis were visualized in Figure S1 of Supplementary Material.

Table 3. Hierarchical regression predicting well-being: Model Summary

	Model 1 English proficiency	Model 2 Putonghua proficiency	Model 3 Years of education	Model 4 Household registration	Model 5 Ln income	Model 6 SES	Model 7 Health	Model 8 Perceived social fairness
R^2	.024	.032	.032	.033	.036	.066	.084	.117
ΔR^2	.024	.008	.001	.000	.003	.030	.018	.033
ΔF	73.741	23.641	1.723	1.308	10.370	95.1977	58.6605	110.864

Note: For Models 1-8, the variable underneath 'Model' indicates that it was the newly added predictor in this particular model.

Table 3 provides the key information of one example from the 40,000+ regression scenarios predicting well-being. In this scenario, English proficiency was entered into the first block, Putonghua proficiency the second, years of education the third, resident type the fourth, income the fifth, SES the sixth, health the seventh, and perceived social fairness the eighth. Each block statistically significantly ($p < .0005$) added to the prediction of well-being; in Model 8 (see Table 3), the eight predictors altogether accounted for 11.4% of the variance in well-being ($R^2 = .117$). The ΔR^2 column in Table 3 contains the most important information: (1) English proficiency, SES and perceived social fairness respectively explained 2.4%, 3.0% and 3.3% of the well-being variance, which exceeded the large effect size benchmark (2%), (2) health accounted for 1.8% of the well-being variance, which was higher than the typical effect size benchmark (1%), (3) Putonghua proficiency contributed 0.8% in the well-being variance, which exceeded the small effect size benchmark (0.5%), and (4) the unique contribution to the well-being variance respectively from years of education (.1%), household registration (.0%), and Ln income (.3%) fell below the small effect size benchmark (.5%) and hence could be deemed negligible. It was noteworthy that in this particular regression scenario English proficiency emerged as a more important predictor for well-being than Putonghua proficiency.

Discussion

RQ1 examines to what extent bilingualism is linked to well-being. A concise answer to RQ1 is that higher bilingualism was statistically significantly associated with a higher well-being level and the strength of this association $r = .16$ (i.e., the unsquared version of $R^2 = .026$) slightly exceeded the strong benchmark (.14). The former part of our answer concerning statistical significance echoed previous studies (e.g., Grey & Thomas, 2019; Kim et al., 2012). Similarly, the latter part of our answer concerning the strength of association fell within the range of effect sizes identified in earlier research; specifically, our effect size (.16) was higher than Zhang et al.'s (2021) result (.09) and lower than Grey and Thomas' (2019) finding (.26).

The mechanism behind this link between bilingualism and well-being (an indicator of a better life) may be attributed to both (language) learner external (e.g., job requirements) and internal factors (e.g., an open and curious attitude towards the world, see Luo & Wei, 2021). For example, in China, an adult with FL-based bilingualism (Wei et al., 2022a) manages to find a highly-paid job because of his/her higher English language proficiency compared with the other competitors, which leads to a higher level of well-being. We need to acknowledge that there may be alternative explanations and applying mediator analyses could be one strategy in future research to ascertain the

Table 4. Predictor importance

A more refined version of hierarchical analysis Dominance analysis						
Variables	<i>r</i>	<i>b</i>	<i>p</i>	ES range (% of DV variance)	Importance	Rescaled dominance weight (%)
Fairness	.22	.183	<.0005	[3.29, 4.58]	very important	32.40
SES	.22	.156	<.0005	[2.08, 4.86]	very important	26.35
Health	.20	.132	<.0005	[1.60, 3.65]	important	20.36
English	.16	.064	.005	[.24, 2.41]	potentially important	6.92
Putonghua	.16	.056	.005	[.23, 2.15]	potentially important	6.70
Income	.11	.011	.560	[.01, 1.18]	negelible	2.80
YoEdu	.15	-.025	.298	[.03, 1.68]	negelible	2.46
HouReg	.10	.019	.350	[.03, .95]	negelible	2.02

Note: Fairness = perceived social fairness; English = English proficiency; Putonghua = Putonghua proficiency; YoEdu = Years of education; HouReg = Household registration

mechanism behind the bilingualism-well-being linkage. However, the current state of knowledge does not warrant the application of mediation analyses; more empirical investigations are required to establish the extent of the bilingualism-well-being linkage BEFORE meaningful mediation analyses are conducted (see also Note 5).

As results from bivariate analyses (e.g., simple linear regression) tend to produce inflated effect sizes (Wang et al., 2022; Wei et al., 2022a), more attention should be paid to the discussion of RQ2 below, which was addressed by multivariate analyses (e.g., hierarchical regression) that could paint a more accurate picture than bivariate analyses.

RQ2 probes the extents to which bilingualism and other selected variables were linked to well-being. Our succinct answer to RQ2 is that perceived social fairness (effect size $\Delta R^2 = 3.4 - 4.7\%$), SES (2.08 – 4.86%), and health (1.60 – 3.65%) were important predictors for well-being, while English proficiency (0.3 – 2.5%) and Putonghua proficiency (0.2 – 2.0%) were potentially important predictors. Three crucial observations can be made here. Firstly, in terms of statistical significance, our results (e.g., see Model 8 in Table S1 of Supplementary Material) were consistent with previous studies (e.g., Zhang et al., 2021); specifically, for instance, perceived social fairness, health, SES, and Putonghua proficiency were statistically significant predictors for well-being, as were the case in Zhang and Cheng (2022). Secondly, the effect size ranges from the present study cannot be compared with previous findings, partly because researchers using multiple regression tended to rely only upon one effect size for the contribution of each predictor (e.g., Khawaja et al., 2016; Zhang & Cheng, 2022), or just one single overall effect size for a block of predictors (e.g., Sun et al., 2016; Zhang et al., 2021). Thirdly, our study represents the first systematic attempt to consider English proficiency alongside Putonghua proficiency as correlates of well-being based on a big-data survey. Our endeavour to address the big question ‘is language linked to well-being’ and broaden the research scope on well-being turns out to be worthwhile: both the upper and lower limits of the effect size for English proficiency were higher than their counterparts for Putonghua proficiency, suggesting that English proficiency was a more important predictor for well-being than Putonghua proficiency. Put differently, it was problematic for the prototype paper (viz. Zhang & Cheng, 2022) to overlook English proficiency and focus on Putonghua proficiency. Accordingly, we suggest that

future studies pursuing the above big question adopt a more HOLISTIC perspective towards language, which includes the national language and FL.

Conclusion

Partly motivated by the seldom-explored linkage between one’s FL proficiency and well-being (Zhang et al., 2021), the present study, based on the empirical data from a nationally representative big-data survey, has found that higher bilingualism is linked to a higher level of well-being, among others. We argue that to address the big question ‘is language linked to well-being’, ‘language’ should be inclusive enough rather than confined to the national language. Based on the finding, we hypothesise that bilingualism (operationalised as FL proficiency) is linked to well-being at a strength level (at least) comparable to the link between proficiency in the national language and well-being. This hypothesis will need to be further tested, modified, or falsified in further studies; before sufficient empirical data are accumulated in this regard, the time may not yet be ripe for the development of a theory focussing on the extent of the bilingualism-well-being linkage, which integrates bilingualism and other sociobiographical correlates (e.g., SES). Accordingly, we call for further replication studies to generate more empirical data, preferably based on representative samples such as the big-data sample used in our study. With empirical evidence from (partial) replications, it is then feasible to generate sufficient theoretical insight which helps pave the way for developing a theory; just as Bollier and Firestone (2010, p.8) rightfully point out, the more data there are, the better chances of ‘finding the “generators” for a new theory’.

In terms of methodological contributions, two major points merit attention. Firstly, thanks to the benefits¹⁶ of using big data, the present study represents one of the first attempts (e.g., Luo & Wei, 2021; Wei et al., 2022a) to utilise publicly available big-data surveys which are usually not designed for language-focused research purposes. We urge colleagues to explore and mine relevant data from those surveys to address issues of interest in the field of applied linguistics generally and in bilingualism research in particular. Secondly, we have made fuller use of effect

¹⁶Big data surveys have many inherent advantages; for instance, Yae and Yoon (2017) discuss the authoritative, economical, holistic, and timely features of big data.

sizes (in terms of both reporting and interpreting), compared with most previous studies. On the one hand, regarding effect size reporting, we provided not only a range of effect sizes via a more refined version of hierarchical regression, which could be usefully supplemented with dominance analysis, but also *two or more types of effect sizes* to facilitate cross-study comparisons. On the other hand, regarding effect size interpreting, we adopted an interpretation system that is more appropriate for survey-based research (Walton, 2022; Wei & Hu, 2019), rather than systems (e.g., Plonsky & Oswald, 2014) that are more relevant to experiment-based studies and/or have inflated benchmarks; one consequence of using those benchmark systems is that the results may have ‘discouraged researchers from delving further’ (Dewaele, 2012, p. 43) into links of interest, such as the link between bilingualism and well-being in the present study. Responding to recent calls for stronger methodological rigour (Li, 2022; Wei & Hu, 2021), we suggest that colleagues make fuller use of effect sizes in both reporting and interpreting the results.

Despite its substantive and methodological contributions, this study has three major limitations. First, despite the many advantages (e.g., sample representativeness) of using extant big-data surveys (e.g., the CGSS), one apparent disadvantage is that such surveys are not designed to satisfy all of the intended purposes of a particular research (Luo & Wei, 2021). In future, studies similar to the present one will benefit from leveraging both big data and small data (e.g., experimental evidence generated from a small sample) (Wei et al., 2022a). Second, besides the measure of bilingualism in our study, there are other useful measures, including Dewaele and Li’s (2013) operationalisation of bilingualism (or what they call ‘a global measure of multilingualism’). A different measure of bilingualism may generate a different set of results and interpretations. Third, although our study utilising a big data survey involved people from different occupations, it aimed to paint a general picture and hence did not probe into particular occupation groups. Given the recent calls for more research attention to non-student populations (e.g., Mercer, 2021; Wei & Su, 2015; Wei et al., 2022b), it will be useful for future studies based on non-big-data samples to focus on a particular occupation group such as business professionals and teachers (see e.g., Alqarni, 2021). These issues merit continued research effort.

Supplementary Material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1366728923000603>

Data availability. The data that support the findings of this study are openly available from <https://osf.io/zd9k5/>.

Acknowledgements. The authors would like to extend their thanks to the anonymous reviewers and the handling editor for their constructive comments on an earlier version of this paper. All remaining inadequacies are the authors’ responsibility. The writing of this paper was supported by the Bilingual Cognition and Development Lab, Center for Linguistics and Applied Linguistics, Guangdong University of Foreign Studies (No. BCD202009), the Chinese Society for Multilingualism and Multilingual Education affiliated to the International Association of Multilingual Education, and Xi’an Jiaotong-Liverpool University [grant number REF-19-02-01]. We also wish to thank our research assistants, Mr Hongzhong Chen, Mr Yuansheng Li, and Ms Yangxu Shen for their technical support (e.g., formatting).

References

Abdel-Khalek, A. M. (2006). Measuring happiness with a single-item measurement. *Social Behavior and Personality: An International Journal*, 34(2), 139–150.

- Agler, R., & Boeck, P. De. (2017). On the Interpretation and Use of Mediation: Multiple Perspectives on Mediation Analysis. *Frontiers in Psychology*, 8, 1–11. <https://doi.org/10.3389/fpsyg.2017.01984>
- Alqarni, N. A. (2021). Well-being and the perception of stress among EFL university teachers in Saudi Arabia. *Journal of Language and Education*, 7(3), 8–22. <https://doi.org/10.17323/JLE.2021.11494>
- Baker, C. (1996). *Foundations of bilingual Education and bilingualism*. Multilingual Matters.
- Bollier, D., & Firestone, C.M. (2010). *The promise and peril of big data*. Washington, DC: Aspen Institute, Communications and Society Program.
- Chen, P. (1999). *Modern Chinese: History and Sociolinguistics*. Cambridge University Press.
- Chen, W. (2004). A study of the correlation between English majors’ scores of TEM4 and tolerance of ambiguity. *CELEA Journal*, 27(1), 3–6.
- Chen, Y., Dies, Y. J., Uni, S., & Mu, Z. (2015). A survey study: The correlation between introversion/extroversion and oral english learning outcome. *Journal of Language Teaching and Research*, 6(3), 581–587. <https://doi.org/10.17507/jltr.0603.14>
- Chen, N., Shen, Y., Liang, H., & Guo, R. (2021). Housing and adult health: Evidence from Chinese General Social Survey (CGSS). *International Journal of Environmental Research and Public Health*, 18(3), 1–13. <https://doi.org/10.3390/ijerph18030916>
- Cuñado, J., & de Gracia, F. P. (2012). Does education affect happiness? Evidence for Spain. *Social Indicators Research*, 108(1), 185–196. <https://doi.org/10.1007/s11205-011-9874-x>
- Desjardins, R. (2008). Researching the links between education and well-being. *European Journal of Education*, 43(1), 23–35. <https://doi.org/10.1111/j.1465-3435.2007.00333.x>
- Dewaele, J.-M. (2012). Personality: Personality traits as independent and dependent variables. In M. Sarch, R. Stephen, & W. Marion (Eds.), *Psychology for Language Learning: Insights from Research, Theory and Practice* (pp. 42–58). https://doi.org/10.1057/9781137032829_4
- Dewaele, J.-M., & Botes, E. (2020) Does multilingualism shape personality? An exploratory investigation. *International Journal of Bilingualism* 24(4), 811–823.
- Dewaele, J.-M, & Li, W. (2013) Is multilingualism linked to a higher tolerance of ambiguity? *Bilingualism: Language and Cognition*, 16(1), 231–240.
- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237–274. <https://doi.org/10.14746/ssl.2014.4.2.5>
- Dewaele, J.-M., & van Oudenhoven, J. P. (2009). The effect of multilingualism/multiculturalism on personality: no gain without pain for Third Culture Kids? *International Journal of Multilingualism*, 6(4), 443–459. <https://doi.org/10.1080/14790710903039906>
- Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, 22(6), 676–697. <https://doi.org/10.1177/1362168817692161>
- Diener, E. (2009). Introduction—The science of well-being: Reviews and theoretical articles by Ed Diener. In *The Science of Well-being The collected works of Ed Diener* (pp. 1–10). Springer.
- Diener, E., Emmons, R. A., Larsen, R. J., Griffin, S., Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). *The Satisfaction With Life Scale*. 49(1), 71–75. <https://doi.org/10.1207/s15327752jpa4901>
- Ding, J., Salinas-Jiménez, J., & Salinas-Jiménez, M. D. (2021). The impact of income inequality on subjective well-being: The case of China. *Journal of Happiness Studies*, 22(2), 845–866. <https://doi.org/10.1007/s10902-020-00254-4>
- Egri, C. P., & Ralston, D. A. (2004). Generation cohorts and personal values: A comparison of China and the United States. *Organization Science*, 15(2), 210–220. <https://doi.org/10.1287/orsc.1030.0048>
- Fan, L. (2023). *Dominance Analysis Calculator [Web application]*. <https://virgentop.shinyapps.io/DACalculator/>
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Arnold.
- Grey, I., & Thomas, J. (2019). Nationalidentity, implicit in-group evaluation, and psychological well-being among Emirati women. *Journal of*

- Cross-Cultural Psychology*, 50(2), 220–232. <https://doi.org/10.1177/0022022118812131>
- Hu, K., & Li, X. (2019). The effects of media use and traditional gender role beliefs on tolerance of homosexuality in China. *Chinese Sociological Review*, 51(2), 147–172. <https://doi.org/10.1080/21620555.2019.1595567>
- Jiang, Y., & Dewaele, J. (2020). The predictive power of sociobiographical and language variables on foreign language anxiety of Chinese university students. *System*, 89(59), 102207. <https://doi.org/10.1016/j.system.2020.102207>
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review* 5(5), 602–619.
- Kang, H. (2022). The influence of putonghua ability on farmers' subjective well-being: From the perspective of rural revitalization[Putonghua nengli dui nongmin zhuguanxingfugan deyingxiang]. *Chinese Journal of Language Policy and Planning [Yuyan Zhanlue Yanjiu]*, 1(37), 48–60.
- Khajavy, G. H., & Aghaee, E. (2022). The contribution of grit, emotions and personal bests to foreign language learning. *Journal of Multilingual and Multicultural Development*, 1–15. <https://doi.org/10.1080/01434632.2022.2047192>
- Khawaja, N. G., Yang, S., & Cockshaw, W. (2016). Taiwanese migrants in Australia: An investigation of their acculturation and wellbeing. *Journal of Pacific Rim Psychology*, 10. <https://doi.org/10.1017/prp.2016.1>
- Kim, S. H. O., Ehrich, J., & Ficorilli, L. (2012). Perceptions of settlement well-being, language proficiency, and employment: An investigation of immigrant adult language learners in Australia. *International Journal of Intercultural Relations*, 36(1), 41–52. <https://doi.org/10.1016/j.ijintrel.2010.11.010>
- Kong, M., & Wei, R. (2019). EFL learners' attitudes toward English-medium instruction in China: The influence of sociobiographical variables. *Linguistics and Education*, 52, 44–51. <https://doi.org/10.1016/j.linged.2019.03.005>
- Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning*, 65(S1), 127–159. <https://doi.org/10.1111/lang.12115>
- Lee, J., Niu, A., & Yang, H. S. (2021). Language proficiency and subjective well-being: Evidence from immigrants in Australia. *Journal of Happiness Studies*, 0123456789. <https://doi.org/10.1007/s10902-021-00474-2>
- Leech, N. L., C. Barret, K., & Morgan, G. A. (2014). *IBM SPSS for Intermediate Statistics: Use and Interpretation* (fifth edit). Routledge.
- Li, X. (2022). A review of CLIL outcomes: Calling for fuller use of effect sizes and less reliance upon *p*. *International Journal of English for Academic Purposes: Research and Practice*, 2021, 205–212. <https://doi.org/10.3828/ijeap.2021.15>
- Li, S., & Zhao, H. (2021). The methodology of the research on language aptitude: A systematic review. *Annual Review of Applied Linguistics*, 41, 25–54. <https://doi.org/10.1017/S0267190520000136>
- Liu, E., & Wang, J. (2021). Examining the relationship between grit and foreign language performance: Enjoyment and anxiety as Mediators. *Frontiers in Psychology*, 12, 1–10. <https://doi.org/10.3389/fpsyg.2021.666892>
- Liu, P., Cao, J., Nie, W., Wang, X., Tian, Y., & Ma, C. (2021). The influence of internet usage frequency on women's fertility intentions—the mediating effects of gender role attitudes. *International Journal of Environmental Research and Public Health*, 18(9). <https://doi.org/10.3390/ijerph18094784>
- Lu, L., & Alon, I. (2004). Analysis of the changing trends in attitudes and values of the Chinese: The case of Shanghai's young & educated. *Journal of International and Area Studies*, 11(2), 67–88.
- Luo, X., & Wei, R. (2021). Is multilingualism linked to a higher tolerance of homosexuality? Evidence from a national survey. *Journal of Multilingual and Multicultural Development*. <https://doi.org/10.1080/01434632.2021.1874002>
- Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, 68(2), 321–391. <https://doi.org/10.1111/lang.12286>
- Mercer, S. (2021). An agenda for well-being in ELT: An ecological perspective. *ELT Journal*, 75(1), 14–21. <https://doi.org/10.1093/elt/ccaa062>
- Mizumoto, A. (2023). Calculating the relative importance of multiple regression predictor variables using dominance analysis and random forests. *Language Learning*, 73(1), 161–196. <https://doi.org/10.1111/lang.12518>
- Oxford, R., & Ehrman, M. (1993). Second language research on individual differences. *Annual Review of Applied Linguistics*, 13, 188–205.
- Pimsleur, P. (1966). *The Pimsleur Language Aptitude Battery*. Harcourt, Brace, Jovanovic.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 Research. *Language Learning*, 64(4), 878–912. <https://doi.org/https://doi.org/10.1111/lang.12079>
- Proietti Ergün, A. L., & Ersöz Demirdağ, H. (2022). The relation between foreign language enjoyment, subjective well-being, and perceived stress in multilingual students. *Journal of Multilingual and Multicultural Development*. <https://doi.org/10.1080/01434632.2022.2057504>
- Qi, W., Xu, W., Qi, X., & Sun, M. (2023). Can environmental protection behavior enhance farmers' subjective well-being? *Journal of Happiness Studies*, 24(2), 505–528. <https://doi.org/10.1007/s10902-022-00606-2>
- Robinson, P. (2002). *Individual differences and instructed language learning*. John Benjamins.
- Rubin, J. (1975). What the “good language learner” can teach us. *TESOL Quarterly*, 9(1), 41–51.
- Sari, B. T., Chasiotis, A., van de Vijver, F. J. R., & Bender, M. (2018). Parental culture maintenance, bilingualism, identity, and well-being in Javanese, Batak, and Chinese adolescents in Indonesia. *Journal of Multilingual and Multicultural Development*, 39(10), 853–867. <https://doi.org/10.1080/01434632.2018.1449847>
- Shields, M. A., & Price, S. W. (2002). The English language fluency and occupational success of ethnic minority immigrant men living in English metropolitan areas. *Journal of Population Economics*, 15(1), 137–160. <https://doi.org/10.1007/PL00003836>
- Skehan, P. (1989). *Individual differences in second language learning*. Edward Arnold.
- Sun, S., Pan, W., & Wang, L. L. (2010). A Comprehensive review of effect size reporting and interpreting practices in academic journals in education and psychology. *Journal of Educational Psychology*, 102(4), 989–1004. <https://doi.org/10.1037/a0019507>
- Sun, H., Steinkrauss, R., Tendeiro, J., & De Bot, K. (2016). Individual differences in very young children's English acquisition in China: Internal and external factors. *Bilingualism: Language and Cognition*, 19(3), 550–566. <https://doi.org/10.1017/S1366728915000243>
- Tang, N., Wang, Y., & Zhang, K. (2017). Values of Chinese generation cohorts: Do they matter in the workplace? *Organizational Behavior and Human Decision Processes*, 143, 8–22. <https://doi.org/10.1016/J.OBHP.2017.07.007>
- Tonidandel, S., & Lebreton, J. M. (2011). Relative importance analysis: A useful supplement to regression analysis. *Journal of Business Psychology*, 26, 1–9. <https://doi.org/10.1007/s10869-010-9204-3>
- Walton, H. (2022). Investigating a new psychological variable in second language acquisition: Comparing social identity across Canadian French education programs. *Frontiers in Psychology*, 7, 1–12. <https://doi.org/10.3389/feduc.2022.874287>
- Wang, B., & Wang, P. (2008). Contemporary educational functions in the view of GNH[GNH shiyuzhong de dangdaijiaoyu]. *Educational Research[Jiaoyu Yanjiu]*, 3(338), 24–28.
- Wang, Yan, Zhang, R., & Peng, S. (2021). Cognitive differences and influencing factors of chinese people's old-age care responsibility against the ageing background. *Healthcare (Switzerland)*, 9(1). <https://doi.org/10.3390/healthcare9010072>
- Wang, Yongliang, Derakhshan, A., & Zhang, L. J. (2021). Researching and practicing positive psychology in second/foreign language learning and teaching: The past, current status and future directions. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.731721>
- Wang, J., Ying, B., Liu, H., & Wei, R. (2022). Exploring L2 engagement: A large-scale survey of secondary school students. *Frontiers in Psychology*, 13, 1–12. <https://doi.org/doi:10.3389/fpsyg.2022.868825>
- Wei, R., & Gao, X. (2022). New research trends in TESOL: An introduction. *Chinese Journal of Applied Linguistics*, 45(2), 220–223.
- Wei, R., & Hu, Y. (2019). Exploring the relationship between multilingualism and tolerance of ambiguity: A survey study from an EFL context. *Bilingualism: Language and Cognition*, 22(5), 1209–1219. <https://doi.org/10.1017/S1366728918000998>
- Wei, R., & Hu, J. (2021). TESOL in the New Era: Introduction to the 2021 Global English Education China Assembly Special Issue. *International Journal of TESOL Studies*, 3, 1–7. <https://doi.org/10.46451/ijts.2021.12.01>

- Wei, R., & Su, J. (2008). An analysis of foreign language use in the Chinese Mainland [Zhongguo neidi waiyushiyong qingkuang diaocha fenxi]. *The Journal of Chinese Sociolinguistics [Zhongguo Shehui Yuyanxue]*, 2, 9–24.
- Wei, R., & Su, J. (2015). Surveying the English language across China. *World Englishes*, 34(2), 175–189.
- Wei, R., Hu, Y., & Xiong, J. (2019). Effect size reporting practices in applied linguistics research: A Study of one major journal. *SAGE Open*, 9(2). <https://doi.org/10.1177/2158244019850035>
- Wei, R., Liu, H., & Wang, S. (2020). Exploring L2 grit in the Chinese EFL context. *System*, 93. <https://doi.org/10.1016/j.system.2020.102295>
- Wei, R., Jiang, H., & Kong, M. (2021). Attitudes toward trilingualism: a survey study of Chinese Mongolian university students. *Journal of Multilingual and Multicultural Development*, 42(3), 291–306. <https://doi.org/10.1080/01434632.2019.1689245>
- Wei, R., Reynolds, B. L., Kong, M., & Liu, Z. (2022a). Is bilingualism linked to national identity? Evidence from a big-data survey. *Journal of Multilingual and Multicultural Development*. <https://doi.org/10.1080/01434632.2022.2085282>
- Wei, R., Wang, Y., & Li, X. (2022b). Examining resilience in EFL contexts: a survey study of university students in China. *International Review of Applied Linguistics in Language Teaching*. <https://doi.org/10.1515/iral-2021-0221>
- Wen, Z., & Skehan, P. (2021). Stages of acquisition and the P/E model of working memory: Complementary or contrasting approaches to foreign language aptitude? *Annual Review of Applied Linguistics*, 41, 6–24. doi:10.1017/S0267190521000015
- Wen, Q., & Zhang, H. (2020). China going global: Challenges and responses in English as a foreign language teaching and teacher education. In A. B. M. Tsui (Ed.), *English Language Teaching and Teacher Education in East Asia: Global Challenges and Local Responses* (pp. 1113–1134). Cambridge University Press.
- Yae, J., & Yoon, S. (2017). The compatibility condition for expressives revisited: A big data-based trend analysis. *Language Sciences*, 64, 69–102. <https://doi.org/10.1016/j.langsci.2017.07.003>
- Yan, Y., Deng, Y., Igartua, J. J., & Song, X. (2023). Does Internet use promote subjective well-being? Evidence from the different age groups based on CGSS 2017 data. *International Journal of Environmental Research and Public Health*, 20(4). <https://doi.org/10.3390/ijerph20042897>
- Yin, Y. (2021). Language use and Urban integration of new intellectual immigrants in second-tier cities. *Chinese Journal of Language Policy and Planning*, 6(3), 35–44. doi:10.19689/j.cnki.cn10-1361/h.20210303.
- Zhang, W., & Cheng, S. (2022). Language makes life better: The impact of mandarin proficiency on residents' subjective well-being[yuyan rang shenghuo gengmeihao : outpnghuanengli dui jumin zhuguanxingfugan de yingxiang]. *Journal of Shandong University(Philosophy and Social Sciences) [Shandongdaxuexuebao (Zhexueyushehuikexueban)]*, 2, 118–133. <https://doi.org/10.19836/j.cnki.37-1100/c.2022.02.012>
- Zhang, H., Dai, Y., & Liu, H. (2021). English proficiency and happiness: The mediation of income satisfaction and leisure satisfaction and the moderation of the national economy. *SAGE Open*, 11(4). <https://doi.org/10.1177/21582440211054482>