Article

Tests of Learning or Testing for Learning? An Exploratory Study of Motivation and Language Learning Strategies Among HSK Level 1-3 Test-takers in UK

Clare Wright*
University of Leeds, UK

Yang Lu
University of Nottingham, UK

April Jiawei Zhang
Macau University of Science and Technology, SAR China

Lulu Zhang
Ying Zheng
University of Southampton, UK

Abstract

This in-depth study evaluated motivation, learning strategies and demographic factors among early-level Chinese HSK test-takers, predicting differences at higher test levels when Chinese character knowledge is required. 71 UK-based HSK test-takers from beginner (1/2) to intermediate (3) levels were recruited; using quantitative survey and qualitative interviews, test-taker characteristics were analysed. However, findings revealed similar patterns at all three levels, and few effects of demographic factors. Higher levels of intrinsic than extrinsic motivation, and higher use of meaning-focused learning strategies were found across levels, although learners may start with structure-based priorities at beginner levels, especially in female learners. Including pinyin made beginner-level HSK exams seem more accessible, but commitment to learning characters was also evident at beginner stages. Noting limitations in the HSK exam format as a test of linguistic abilities, the findings are evaluated for wider pedagogical implications for effective early-stage L2 Mandarin Chinese language development.

*Corresponding author. Email: C.E.M.Wright@leeds.ac.uk

1 Introduction

The aim of this study is to add to existing understanding of how assessment may play a role in fostering effective L2 Mandarin Chinese (henceforth Chinese) language development. As part of a wider project funded by Chinese Testing International (CTI), known as Hankao International, this study explored factors impacting on progress up from beginner to intermediate level, to see if there are marked differences between levels 1 and 2, which do not require character knowledge for the assessment, compared to level 3, which does.

In line with other studies of learner attitudes in examination-focused language learning settings (Wang et al., 2009; Yu, 2010), three key aspects were identified for this study - first, different kinds of motivation (personal, community, or global); then, the use of different kinds of learning strategies (structure or meaning-based), and finally, challenges in learning written Chinese and using pinyin. By using a mixed methods approach through quantitative surveys and in-depth interviews, the project took a rich in-depth approach to discover what makes for a successfully motivated Hanyu Shuiping Kaoshi (HSK) test-taker, testing assumptions that there would be changes in motivation and strategies between level 1 and 2 test-takers compared to level 3 test-takers. To our knowledge, no such cross-sectional study of early stage anglophone HSK test-takers has been carried out before, though see a sister paper from this project (Zheng et al., 2021), and the study thus aims to provide helpful insights and guidance to learners, teachers, and assessors about the learning journey facing HSK test-takers.

2 Background to Study

Notions of motivation and learning strategies, particularly in relation to exam-based instruction, have long been influenced by Gardner's (1985) early definition of "the extent to which an individual works or strives to learn the language because of a desire to do so and the satisfaction experienced in this activity" (Gardner, 1985, p. 10). Later models of motivation also distinguish between instrumental or extrinsic benefits, e.g. in taking high stakes tests, and integrative or intrinsic benefits, e.g. in developing general language and communicative abilities (Ryan & Deci, 2000). Motivation in this sense is a double-edged sword. On the one hand, it may influence test-taker performance; and on the other, what students learn and how they perform may affect their motivation (Pintrich & Schunk, 2002). How far motivation for test preparation may impact on general learning and vice-versa, and what teachers should do to boost effective learning and test preparation, thus remain vexed questions to operationalise and explore, particularly in relation to HSK exams, which remain relatively underexplored in assessment and strategy-based research (Wang et al., 2022).

2.1 Learning and assessing Chinese

Interest in learning Chinese and take-up of Chinese assessment has grown exponentially in recent years (Wright, 2018; Zhang & Lin, 2017); globally, test-taker numbers of all available Chinese language tests exceeded 6 million in 2017 (Xinhua News, 2017). Currently, HSK exams are the most widely recognised tests of Chinese within China, and are growing in use around the world (Lee et al., 2017; Lu & Song, 2017). Originally introduced in 1984, the new Revised HSK exams were introduced in 2010 and consist of six levels, divided into three broad proficiency bands: Elementary (Levels 1 and 2), Intermediate (Levels 3 and 4) and Advanced (Levels 5 and 6). HSK tests do not include speaking, as there are separate oral exams, Hanyu Shuiping Kouyu Kaoshi (HSKK), at the same three broad levels: Elementary, Intermediate and Advanced. According to the official HSK website, its main principle for the test was to establish a 'test-teaching correlation', underpinning its aim to 'promote learning and teaching through testing'. Hanban (2008) stated that the HSK levels correspond to the Chinese

Language Proficiency Scales for Speakers of Other Languages and the Common European Framework of Reference descriptions of proficiency levels: namely, Levels 1 to 6 as equivalent to CEFR A1, A2, B1, B2, C1 and C2, Levels 1 and 2 are seen as approximating to American Council on the Teaching of Foreign Languages (ACTFL) Novice; level 3 to Novice High-Intermediate Low. HSK 1-2 exams include pinyin alongside characters; HSK 3 uses only characters, so requires good reading knowledge, though no writing is required. HSK 3 can also be used as the basis for competitive study bursaries (e.g. Confucius Institute of Scotland, n.d.). A further update of HSK levels is taking place, but this study focuses on the revised levels described above.

Limitations have been noted in HSK's assumed alignment to external standards, particularly to CEFR levels (Hsiao & Broeder, 2013). This has caused concerns about ensuring valid and reliable criteria for what the HSK tests claim to measure, especially for written components (Lu, 2017; Lu & Song, 2017; Yan & Xie, 2014). As a result, L2 Chinese learners' written language may be assessed less or with test items of much lower levels of difficulty than those for speaking and listening (Lu, 2017). The European Benchmarking Chinese Language project (Lu & Song, 2017) took some steps to address this, for example, regarding CEFR as an overall guideline for describing L2 competence in Chinese but adding specific requirements when applied to the features of the Chinese language in relation to writing, including the competence for writing the script and the Romanised alphabetic phonetic system.

There remain many contradictory claims over effective teaching methods (Guder et al., 2007), and how to ensure effective written functional competence in written Chinese. Influenced by recent research (e.g. Seymour, 2006), taking the notion of word as the basic unit for expressing meaning, some have proposed 语文分进 (*Yǔ Wén Fēn Jìn*, teaching the spoken and written language separately), focusing on oracy first (Zhao, 2010). Others argue that characters are the basic units for expressing meaning (Zhang, 2007), and propose methods such as 随文学语 (*Suí Wén Xué Yǔ*, learning the spoken language through learning the characters). Further, 组合法 (*Zǔ Hé Fǎ*, combination methodology) is recommended for teaching, as compounding or combining is seen as the fundamental rule for words, construction of characters, pronunciation and grammar in Chinese.

As a result perhaps of these contradictions, some learners in their second or third year of a Chinese course are still not able to understand the simplest stories in Chinese texts (Diao, 2013), while some textbooks oversimplify reading comprehension tasks (Guder et al., 2007). Others have proposed that beginners should be only taught to type Chinese characters rather than to write them (Allen, 2009). Pinyin may even be accepted as a parallel 'other script' and even included as part of the competence in written Chinese by the European Benchmark of the Chinese Language (Lu, in press).

Teachers can thus remain in a dilemma as to which approach to adopt with their learners. Mastering written Chinese remains one of the key challenges, particularly for western learners, and exam-based assessments of Chinese literacy, or an overly-high focus on structural accuracy, can seem divorced from the real-world requirements of communicative Chinese (Wright, 2019). In addition, due to exam washback, classes may turn into exam preparation sessions, rather than incorporating the kind of combined methodologies suggested above (Cheng & Curtis, 2012; Green, 2006). Therefore, understanding how test-takers are motivated to develop their knowledge of Chinese, and what learning strategies they prefer to adopt, particularly in relation to HSK 3 where written Chinese is required, will help shed light on this debate, and provide teachers with a clearer sense of what learners do - or do not - find effective.

2.2 Motivation

Extensive research has examined the variety of motivations in taking tests, given evidence that test-taker beliefs about tests and motivation were found to have positive impacts on score gains (Reeve & Lam, 2007), whether tests were taken for intrinsic or extrinsic reasons (Ryan & Deci, 2000). Research

generally suggests that type of test may have an impact, in that tests for extrinsic, instrumental or global purposes e.g. as required for competitive college entry, high stakes professional employability, or international study may be more motivating than generalised language tests (Cheng, 2008; Gao et al., 2004; Li, 2016). A study conducted by Cheng et al. (2014) explored the relations between test-takers' motivation, test anxiety and test performance by examining three high-stakes language tests, i.e. the Canadian Academic English Language (CAEL) Assessment in Canada, the College English Test (CET) and the General English Proficiency Test (GEPT). The findings in the study of Cheng and her team described different purposes for taking the tests in each testing contexts and indicated that the CAEL was seen as much higher value to the test-takers, while the usefulness of the CET test was perceived as more dominant than that of the GEPT. Peng et al. (2014) also reported that students would be more prepared to sit high-valued tests deemed as of higher importance to stakeholders; higher self-efficacy and effort expenditure would be placed on those high-valued tests.

Intrinsic integrative or personal motivation may also be recognised as a key motivation to learn a second language, envisioning an ideal multi-lingual L2 self (Csizér & Dörnyei, 2005), being proficient enough to use the language communicatively (Yu, 2010). Research has also shown some students using test-taking as a way of monitoring their own progress: Jin (2000) explored students' motivation in taking the College English Test (CET), and found students keen to improve their ability to use English communicatively. The complex nature of motivation for learning Chinese, including motivational changes at different stages of Chinese learning among UK university students and their perceptions of developing a multi-lingual ideal-self, has also been noted (Mayumi & Zheng, 2021).

Test-taking motivations can thus be seen as representing intrinsic attitudes (for personal or community value), as well as extrinsic attitudes to the value of the language on a global stage. Given the extrinsic value placed on high-stakes test-taking for competitive study or international opportunities (such as HSK3), it seems logical to suggest that as learners progress from beginner to intermediate levels of Chinese would move from more intrinsic levels of motivation to more extrinsic levels of motivation.

2.3 Learning strategies

Another factor interconnected with test-taking is use of learning strategies (Cohen & Macaro, 2007; Oxford, 1990, 2011). Put simply, strategy types may differentiate between cognitive, metacognitive and social strategies. Cognitive strategies relate to specific activities to learn, remember and recall language items, including whether to focus on structure (bottom-up) or meaning (top-down) (Whong, 2013). Metacognitive strategies refer to the attitudes and skills around learning more generally, while social strategies refer to how to engage with others to help with problems and motivate a shared sense of learning.

Effective strategy use is generally believed to play a positive role in language learning success and performance, though evidence can be mixed (Sheorey & Mokhtari, 2001). The challenge in mastering Chinese, noted above, suggests that an effective use of wide range of strategies, including for writing purposes, could benefit learners, e.g. leading to greater confidence and better performance (Wang et al., 2009). Strategy use may have a different impact at different course levels; Griffiths (2003) found a significant relationship between strategy use and course level, in that higher level students made highly frequent use of communication strategies. Students at intermediate level may make wider use of learning strategies than beginner and even advanced levels (Hong-Nam & Leavell, 2006), particularly in use of metacognitive (managing self-learning, checking self-learning) and social strategies (help-seeking and correction-seeking). Strategy use may also impact on test type - among those taking the College English Band (CET-4), Zhang and Zhang (2013) found that test-takers who used more metacognitive strategies got higher test scores in the reading part of CET-4. Different types of strategies, such as bottom-up reliance on L1 or top-down reliance on problem-solving may also relate to test level - Tavakoli et al. (2012)

found that test-takers who got higher scores in IELTS listening tests would frequently use problem-solving and directed attention strategies, whereas test-takers who got lower scores tended to use mental translation strategies.

A gap remains in seeing how far existing findings may apply to L2 Chinese, though Shen (2005) is among a handful of studies investigating strategy use, demonstrating the value for successful students in adopting a wide range of orthographic knowledge-based cognitive strategies. In view of the challenges in learning Chinese and preparing effectively for the different HSK levels, it can be assumed that learners would require specific strategies both for structural knowledge such as characters, and meaning-based communicative skills; evaluating empirical evidence of how learners develop such strategies at different levels was therefore one of the main drivers for this study.

Research into students' motivation and learning strategies also indicates the potential impact of demographic factors such as students' background on motivation, learning strategies and their test performance, as outlined below. The impact of test-taker background characteristics on language ability and test performance has long been recognized as influential in the field of language testing, but seems less so for research into strategy development, and hardly at all among HSK Chinese test-takers, which prompted the focus in the present study.

Bachman (1990) argued that test-taker demographic characteristics could include, among other factors, gender, age, length and quality of educational exposure; there is a large volume of published studies on standardized tests (e.g. IELTS, TOEFL) exploring the relationship between such test-taker background characteristics and language ability. Stricker and Rock (1995) examined the complex interrelationships between test-takers' background (including, among others, parental background, age, college type) and US-based GRE (Graduate Record Examinations) test performance. They concluded that parental education had strongest association with test performance, as well as indicating that those who had better educated parents would probably be younger and have attended higher qualification colleges. Gender is not consistently seen as an influence on test performance but evidence on this is mixed (Dass & Akmar, 2007). Educational background, including length of learning (LOL), has been found to be a significant factor, which can impact different test components or language skills in different ways. In analyzing students taking the TOEFL iBT test, Manna and Yoo (2015) found that test-takers with fewer learning hours in English achieved lower means in test results; meanwhile, for those with greater exposure to English, especially living in an English-speaking country, there was a clear positive effect on speaking results, and smaller but still positive effects on reading, listening and writing scores. However, it remains to be confirmed how test-taker characteristics may play a role in assessment performance and general language learning development, particularly in relation to Chinese.

To address some aspects of these gaps, an initial investigation by Zheng et al. (2021) was carried out as a baseline for the current study, by conducting a quantitative analysis from a cohort of approximately 1833 UK-based HSK test-takers of key demographic factors including gender, age, and scores on each component of the test. The analysis revealed no significant score differences compared by gender; however, some effect of age was found, with higher scorers being older test takers. A plausible interpretation for this is that HSK content/topics are more suitable for older learners. However, the nature of the database did not allow for potential co-variation of age with length of learning, which could be a confounding factor, in that length of learning could also lead to higher scores. The study also noted some skill component difficulties, since listening was identified as a big challenge for level 1 test-takers while reading was seen as more of a challenge for level 2 and level 3.

In order to explore these general trends further, the current study was designed to investigate key motivational attitudes, comparing personal, community and global motivational profiles, and key language learning and character learning strategies relating to focus on structure, or meaning, and character learning. The study compared learners across the first three levels of HSK test-takers, to see if motivation and strategies may differ between levels (beginner level, at HSK 1-2, compared to

intermediate level, at HSK3). Due to the research team's familiarity with the Chinese learning context in the UK, the focus is on test-takers in the UK. A mixed-methods approach of quantitative data from online survey questions and qualitative data from individual interviews was used, and research questions were framed to identify potential groupwise and individual patterns, which could provide useful insights for teachers and test designers beyond general statistical trends.

2.4 Research questions

1. Do HSK test-takers show patterns of motivation in relation to personal, community or global factors?
2. Do HSK test-takers show patterns in strategy use for general language learning in relation to strategies used for communication (meaning) and examination preparation (structure)?
3. Do HSK test-takers show patterns in strategies for learning written Chinese?

The following assumptions were tested:

A) that test-takers would demonstrate a shift between lower (level 1-2) to intermediate level (level 3) in emerging functional communicative competence, through high levels of personal motivation compared to community or global motivation, more meaning-focused learning strategies and higher use of character-based learning strategies and less reliance/use of pinyin for written Chinese;

B) higher scores on strategies even within each level may reveal some common demographic associations with demographic factors, e.g. Length of Learning (LOL), age, gender;

C) individual differences (e.g. individual motivational views and specific strategies) may be more prevalent than general group differences.

3 Methodology

3.1 Data collection instruments

An online survey was designed in five sections based on standard self-report questionnaire design (Dörnyei, 2007). The first section captured key demographic data, including gender, age, LOL, educational level, L1, any heritage background. The second and third parts related to motivation (18 items), then on general learning strategies (20 items). These sections were the same for all participants. The fourth section was on strategies for learning written Chinese. Due to the different nature of written Chinese knowledge required at the different levels, this section was specifically designed for each group (20 items for the beginner levels, 17 items for the intermediate level). The second, third and fourth sections used a self-report Likert scale design. Participants were asked to read the statements in each section and click the box to respond (1, 2, 3, 4, or 5) that indicated how much they agreed with each statement as it related to them, where 1 was Strongly disagree, and 5 was Strongly agree. The final section was an open question allowing students to contribute any comments, including on how challenging they found the HSK exam and its relation to using Chinese more broadly. A semi-structured interview protocol was also designed, aligned to the survey's general themes, including the overall impression of HSK, the motivation to learn Chinese and to take HSK exams, the difficulties in HSK, the methods of learning Chinese and preparing for HSK, and the possible suggestions to improve HSK exams, allowing individuals to discuss the themes in more depth. Invitations to participate in the project were distributed via Confucius Institute testing centres testing centres to those who had registered at those centres to conform to data protection (GDPR) rules. The online surveys were completed voluntarily and anonymously using Googleforms. The research team cross-checked survey and interview questions for clarity and to avoid duplication, and all test instruments were piloted in March 2019. Data collection took place in November 2019 to February 2020 (the survey questions are provided in an Appendix, full information on all instruments are available by application to the researchers).

3.2 Participants

Participants in the survey were 71; 40 at level 1, 13 at level 2 (henceforth counted together as 53, grouped as beginner levels), 18 at level 3 (intermediate level). The majority were currently in education, 11 at secondary school, 33 were university students. The remainder were university staff, retired or in other work. Participants were recruited via purposive sampling via six HSK testing centres in Southampton, Sheffield, Manchester, Nottingham, London and Leeds. 21 were recruited for interviews, including 14 level 1 test-takers, three level 2 test-takers and four level 3 test-takers, which were conducted face-to-face in convenient university locations in Sheffield, Southampton, or online for recruits from other cities in the UK, lasting 26-92 minutes. Interviewees were offered a coffee voucher of £10 to recompense them for their time.

Biodata: 36 were female, 35 male. Ages ranged from 14 years to 59 years (mean 26.8 years, SD 11.01). The length of learning (LOL), measured in months, ranged from 1 month to 60 months, with one outlier at level 3 reporting more than 10 years. At level 1-2, the range was 1-60 months; mean LOL was 12.13 (SD 12.54) months; at level 3 the range was 6-48 months (without the single outlier); mean LOL was 20.67 (SD 11.35) months. The majority reported themselves as British or US English L1 (11 others were from around the EU); 8 were reported as heritage learners. Given this very small proportion of non-English L1 speakers, and of heritage learners, L1 and heritage background were removed as factors in subsequent analyses.

All survey responses were entered into SPSS; many of the items showed non-normal distribution, particularly in the smaller intermediate cohort, so non-parametric measures are used - Wilcoxon Signed Rank (WSR) or Mann-Whitney (U) tests of difference and Spearman (r) tests of association (Larson-Hall, 2015).

4 Results and discussion

4.1 Motivation patterns and themes (RQ1)

For the first RQ, three types of motivations were tested: personal, community and global (aligned to intrinsic, mixed and extrinsic motivation). Initial analysis assumed that personal motivation will be higher than other types within-group, but also that motivation scores across all types would be higher for intermediate (level 3), linked to efforts needed for approaching the more challenging level of exam; global motivation was expected to be for level 3, compared to lower levels, linked to the extrinsic rewards for reaching level 3, such as eligibility for study grants in China.

Quantitative survey data: Mean scores are given for the three motivation components - Personal Motivation (PM), Community Motivation (CM), Global Motivation (GM). There were 18 items in total, 6 items per component, with a maximum score of 5 per item. Across the whole cohort, it seems that PM was ranked highest, CM was lowest, as seen in Table 1 below.

Table 1

Test-taking Motivation by Level and by Gender (Mean and SD)

Level & Gender	Personal Motivation (PM)	Community Motivation (CM)	Global Motivation (GM)
Overall (n=71)	4.37 (.49)	2.52 (.67)	3.69 (.83)
Beginner (n=53)	4.38 (.51)	2.44 (.71)	3.53 (.83)
Intermediate (n=18)	4.34 (.44)	2.74 (.46)	4.16 (.65)
Female (n=36)	4.48 (.39)	2.48 (.65)	3.71 (.82)
Male (n=35)	4.27 (.57)	2.56 (.69)	3.67 (.85)

Comparing within-group analyses of difference, at each level, PM and CM rankings were significantly different (WSR=-7.27, $p<.001$). CM and GM were ranked significantly differently (WSR=-7.14, $p<.001$). PM and GM were ranked significantly differently (WSR=-5.01, $p<.001$).

To test if key demographic factors played a role, motivation factors were compared for significant differences by level and by gender, and further tested for association with age or LOL. At beginner levels PM was highest, and was significantly different to CM (WSR=-6.34, $p<.001$), and to GM (WSR=-4.99, $p<.001$); GM and CM were also significantly different (WSR=-6.1, $p<.001$). At intermediate level, PM was significantly higher than CM (WSR=-3.63, $p<.001$), and GM was also significantly higher than CM (WSR=-3.74, $p<.001$), but PM and GM were not significantly different. However, between-group analysis shows that the two levels were not significantly different from each other, apart from GM (significantly higher at intermediate level than beginner level - WSR=852, $p<.01$). By gender, different types of motivation were also similar, as shown in Table 1, with no significant differences.

For Age effects across the whole cohort, there were negative correlations with general CM (r=-.307, $p<.01$) and with GM (r=-.336, $p<.01$), suggesting possibly that the younger (including some school-age) learners were more motivated in wider aspects of motivation than older learners, who had higher levels of personal motivation.

Individual item analysis (see Appendix 1) revealed similarities in rankings of the most highly rated statements, but with a slightly different emphasis between groups. The three highest ranked statements at beginner level all indicated high personal motivation: *I find learning Chinese interesting (PM)* (Mean 4.85); *I want to use the exam as a goal to show I am making progress (PM)* (Mean 4.47); *I can imagine myself speaking Chinese with Chinese people (PM)* (Mean 4.45).

For the intermediate group, the top 3 were similar on two PM items, but with one global motivation item, suggesting the emergence of broader global interest in Chinese: *I find learning Chinese interesting (PM)* (Mean 4.83); *I study Chinese because I am interested in Chinese culture (PM)* (Mean 4.67); *I study Chinese because it will allow me to travel in Chinese speaking countries more easily (GM)* (Mean 4.61).

In sum, it seems that both groups show similar patterns of higher personal motivation (intrinsic) compared to community or global (extrinsic), with few distinct findings relating to age, LOL or gender.

Qualitative data: The qualitative data supports the findings of high levels of general enthusiasm and intrinsic motivation for learning Chinese from the start, though extrinsic motivation for the HSK exam was clear even among beginners. The HSK exam was seen as serving as a 'paper proof' of the test-takers' language ability, which can show their language ability in scholarship applications, university application, and job applications, or to indicate what to work on next. Being easy-to-pass was another reason for beginner learners to take the exams, along with being inexpensive at the beginner levels and well supported by the Confucius Institute in the UK. As one participant stated, '*It's like just having proof that you've reached a certain level' (level 1 RF)*. Another noted '*HSK can provide some sort of objective measure of my current level ... [the exam can] give me an idea basically of what area I need to work more on' (level 2 JC)*. Participants were interested in China, and Chinese culture, not only to study but also to work. Business trips to China, scholarship applications to study in Mainland China or Taiwan, working opportunities at Confucius Institutes in the UK, internships in China, and possible future careers in China were important reasons for the test-takers to study Chinese and gain HSK certificates, even at beginner levels.

There was some sense of exam washback: 81% of the interviewees used HSK to monitor their learning, including all the HSK 2 and HSK 3 test-takers, noting that long-term learning could feel less motivating or felt less successful during class. One of the interviewees explained that '*what I studied for the test, I learn more in five days before the exam revising for the exam than in the past, maybe four weeks in class' (level 1 JYF)*.

However, learning Chinese was an exciting challenge, as a '*foreign language, a more exotic one' (level 1 HB)*. Interviewees believed that '*it is more important to be fluent than to have passed the exam' (level*

2 RM). Learners showed a passion for language learning and multilingual communication, including two learners who could speak more than five languages. Both of these individuals were proud of their language background and explained that language learning has been '*part of my identity' (level 1 JH)*. Chinese was difficult, but learners '*wanted to get to know and learn more about Chinese society and culture, the culture of the past culture, the present' (level 1 JH)*.

The notion of Chinese as challenging, here, is turned into a positive motivational goal for these individuals, and worth trying to master, as noted elsewhere (Wright, 2019). In this cohort, intrinsic personal motivation to start learning Chinese was high from the outset, though with some extrinsic value too, and little further shift by level. The exam was a useful benchmark of motivation, either personally or for global benefits. Teachers would therefore benefit from anticipating high levels of motivation initially, but may need to work to maintain motivation among their students as time goes on.

4.2 General learning strategy patterns and themes (RQ2)

In this section of the survey, learning strategy use was separated into communicative vs. structure-based strategies (10 items in each type, scored on a scale of 1-5), presented first for the whole cohort to see if there are any general differences. Next data are presented by group, assuming that beginners (level 1-2) would score higher on structure-based strategies, but with no particular associations with motivation, LOL, age or gender. By contrast, intermediates (level 3) were predicted to show higher mean scores for communicative-based strategies, presuming a greater interest in functional communicative abilities by level 3; strategy use was predicted to show clear patterns of associations with motivation, and with LOL, but not with age or gender. However, for the whole cohort, there were no differences between strategy type (communicative mean 3.79, SD=.51; structure-based mean 3.71, SD=.60).

Split by level and then by gender, there were small but largely non-significant differences as shown in Table 2 below, though beginner levels' use of communicative strategies was slightly higher (ns) than structure-focused strategies, against predictions. The intermediate group showed no difference by type. By gender, there was a significantly higher use of structure-based strategies amongst intermediate-level female participants, compared to male participants (WSR=237, $p<.01$).

Table 2

Use of Strategies by Level and by Gender

Level	Gender	Strategy type	Mean (SD)
Beginner	Female (n=26)	Communicative	3.64 (.50)
		Structure-based	3.57 (.57)
	Male (n=27)	Communicative	3.87 (.48)
		Structure-based	3.77 (.54)
Intermediate	Female (n=10)	Communicative	3.98 (.48)
		Structure-based	4.22 (.23)
	Male (n=8)	Communicative	3.70 (.60)
		Structure-based	3.36 (.82)

For links with motivation, a significant association was found between communicative strategies and GM (r=.509, p=.031) but not with PM or CM. Associations were not found with LOL. However, there was found to be a significant negative association for age with use of structure-based strategies (r=-.726, p=.017) in the female cohort, interpreted as a higher use of structure strategies found among younger learners, plausibly due to a school exam context.

Looking at the top 3 ranked individual items, these again were similar when compared by level, as shown below. In both groups, there were similar items, indicating a common degree of autonomy and proactive engagement in mastering the language, including a commitment on learning to read characters even at beginner level. Top items for beginners were: *In reading tasks, I look for characters I have learned to help me get the meaning (Structure)* (Mean 4.38)*; I use online materials or websites to help improve my listening and reading (Communicative)* (Mean 4.36)*; I notice my mistakes and use that information to help me when preparing for the exam (Structure)* (Mean 4.32). For intermediate level, the only different item was: *I use apps, e-dictionaries to help me in learning words (Communicative)* (Mean 4.67).

Overall, there was no confirmation in the quantitative data that there would be higher use of communicative-focused learning strategies at intermediate level, or more consistent patterns at this higher level of associations between communicative strategies and motivation patterns. There was also no consistent pattern of association between strategy use and LOL, age or gender.

Qualitative data: Evidence in the interviews also showed a clear balance of attitudes in favour of communicative approaches to learning, even at the beginner levels, including an eagerness to engage in authentic language activities. Clear expectations of communication were embedded from the start without any obvious difference between groups.

Interviewees generally believed that communicative strategies were the best way of learning both spoken and written Chinese. Authentic conversation with Chinese native speakers was seen as the best way of learning speaking, using instant feedback from daily conversation to help pick up the language. Friends, relationships, reality shows, Chinese songs and dramas, were all seen as appropriate sources of the language, though listening could be hard if sources '*speak too fast, so it was quite hard (to listen to)' (level 3 AB)*. Another important way of learning through communication with friends, where effective learning meant they would not *'let friends down' (level 2 RM and level 3 AW)*. Interviewees also reported that '*I tend to write in a diary like every other day' (level 1 AH)* and *'I hope to be able to write a letter in Chinese before I will take the HSK 2 test' (level 1 MAB)*.

Apps and online support were used to build both communicative use and structure-based learning, e.g. personalising vocabulary practice via flashcard apps. However, one participant noted '*it gives you the illusion that you are making a lot of progress, but then it does not translate to anything... It is easy to understand characters on flashcards. It gets a lot harder to read them in actual text…We do not really see them in many times' (level 2 JC)*.

Participants also reported using structure-based strategies in learning speaking and writing, and some evidence of exam-focused washback or traditional drill approaches in classes. Shadowing the teacher in the class was the most direct and common way of learning: '*I just write down what the teacher writes on the whiteboard and I just follow him' (level 1 HB)* or '*we just go through the HSK book' (level 3 AW)*. Some learners did not enjoy class teaching. One participant pointed out that '*the teachers become reliant on the PowerPoint, but [when] they follow this formula, it is ineffective... I feel like they need to be more interactive and have a loose plan for the teaching' (level 1 RF)*. Strategies like writing characters out on paper or in the air were '*very repetitive, but it did work' (level 3 AW)*. '*I have experienced have the mindset where we just repeat repeat repeat, and it will eventually go into your head … I am personally against this style of learning, it works, but I do not like, I do not enjoy it' (level 1 RF)*.

Two of the HSK 3 test-takers mentioned a sudden step-change in their ability to remember characters, as if all of a sudden, they found Chinese easy to read. This may have been specific to these learners. However, the possibility of a "breakthrough moment" in understanding Chinese easily should be encouraged amongst learners, providing motivation for learners at HSK levels if they are about to enter a new and more demanding stage with characters.

In sum, just as for RQ1, there was insufficient evidence to support predictions that strategy use would change as LOL and test-level increased, nor that strategy use would be associated with particular

motivation factors, or show other bio-data effects such as age or gender. Instead, the quantitative and qualitative data suggest that even beginner learners can be highly engaged with proactive meaning-focused learning strategies, and intent on mastering characters. However, learners do not always manage to develop consistent strategies, including using more meaning-focused strategies at higher levels of ability. This was particularly true of older female learners who seemed to be more reliant than others on using structure-based strategies. There was also evidence of discouragement due to washback in levels of teacher-led emphasis on accuracy. Greater awareness of the value of a wide range of strategies should help learners and teachers across a range of language learning needs, and that teachers can do more to assist learners in developing a more consistent attitude to using strategies effectively.

4.3 Writing strategy patterns and themes (RQ3)

For RQ3, the survey items aimed to uncover patterns of effective strategies for developing written literacy, or evidence of maintaining pinyin. Differences were expected to be clear in this section, given the requirement to be functionally literate in Chinese for the HSK3 exam but not for levels 1 and 2. The survey therefore used slightly different items for the two groups.

For the beginner group, twenty items tapped participants' attitudes to using written Chinese vs. pinyin – ten items related to learning characters (LC), and reading and writing skills more broadly (RW), then ten items focused on attitudes to pinyin, and understanding its relation to spoken pronunciation and tones (PIN/SPK). For the level 3 group, there were seventeen items, similar to the beginner level themes but with added items on word learning - ten items related to Chinese (learning characters (LC), reading and writing (RW) and word learning (WL)), and seven items related to pinyin and tones (PIN/SPK). The overall trends in strategy use were reviewed in parallel, to test the assumption in this project that by level 3 participants would feel more confident about learning characters, reading and writing Chinese, and would rely less on pinyin than at the lower levels.

Results grouped by Chinese or pinyin strategy type, split by gender, are summarised in Table 3 below; however, they reveal little significant difference between levels. Overall, data confirm that beginner learners had similar interest in mastering written Chinese as intermediates, but both groups were inconsistent in knowing or using the best strategies. Split by gender, one significant difference was that beginner-level males ranked use of pinyin strategies significantly more highly than females (Mann-Whitney U= 63.5, $p<.05$).

Table 3
Use of Chinese or Pinyin Writing Strategies

Gender	Strategy (with level)	N	Mean (SD)
Female	Chinese (level 1/2)	26	3.23 (.39)
	Chinese (level 3)	10	3.34 (.46)
	Pinyin (level 1/2)	26	3.70 (.51)
	Pinyin (level 3)	10	3.37 (.63)
Male	Chinese (level 1/2)	27	3.37 (.47)
	Chinese (level 3)	8	3.49 (.21)
	Pinyin (level 1/2)	27	3.84 (.39)
	Pinyin (level 3)	8	3.88 (.39)

Analysing individual items in both groups provided further indication of sustained reliance on pinyin even at level 3. For the beginner group, the top 3 strategies were: *Pinyin has made it easier for me to*

learn Chinese (Mean 4.62)*; Pinyin is the pronunciation of the Chinese language not the written language* (Mean 4.47)*; I can achieve higher marks if exams have both pinyin and characters* (Mean 4.09)*.* For the intermediate group, the top 3 strategies emphasised the importance of prior knowledge, and reaffirmed some reliance on pinyin: *I have tried to remember the characters I have learned in the ways I find effective* (Mean 4.5)*; I can guess the meaning of a Chinese word made of two characters I have learned (WD1)* (Mean 4.33)*; Pinyin has made it easier for me to learn Chinese* (Mean 4.17)*.*

Strategy scores were analysed for associations with other co-variables (HSK level, LOL, gender, motivation, learning strategy use). Few associations were significant across the group as a whole, though some weak-moderate significant results were found when split by level. In the beginners group, a moderate negative correlation was found between LOL and use of Chinese strategies (r=-.309, $p<.05$). Positive correlations were also found between use of Chinese strategies and use of structure-focused strategies (r=.354, $p<.01$). Community motivation (CM) was associated with use of both Chinese strategies (r=.330, $p<.05$), and pinyin strategies (r=.290, $p<.05$). Global motivation (GM) was also associated with use of pinyin strategies (r=.352, $p<.01$).

For the beginner-level learners, split by gender, several associations with motivation and learning strategies were found; for males, use of pinyin strategies correlated significantly with GM (r=.466, $p<.05$), while use of Chinese strategies correlated with CM (r=.406, $p<.05$); their structure-based strategies were associated with use of Chinese strategies (r=.382, $p<.05$), and use of pinyin strategies (r=.467, $p<.05$). In the female beginner group, pinyin strategies correlated with use of structure-based strategies (r=.390, $p<.05$), reaffirming the gender effect on structure-based strategies found earlier. However, gender-focused differences for Chinese vs. pinyin learning strategies were not sustained to intermediate level. Here the only evident pattern was a sustained reliance on pinyin strategies, which scored highest in both males and females (Table 3).

Qualitative findings: Interview comments echoed the tension between pinyin and written Chinese noted in the quantitative findings above. Statements reflected both personal (intrinsic) and instrumental (extrinsic) value in using the exam as evidence of their own progress with language knowledge, including being able to read characters, or for future employability options where written Chinese would be required. The distinction between pinyin and written script was clearly understood, even from the start, *'obviously from HSK 3 I think that pinyin is gone, and you are expected to know characters. But even if driven into it, I see the … importance of characters, to drill you from an early stage' (level 1 AH)*. Many were aware of the limitations of reliance on pinyin, but also believed that skills in typing characters could be a significant advantage, not just in exams but since typing is what most people do in daily tasks.

At both levels, test-takers showed a range of individual motivations, strategies and learning needs. Apps seemed to be the most common way to engage with characters, and they showed high levels of motivation to learn by themselves, rather than rely on class. They mentioned flash card apps and online dictionaries (Anki, Memrise, Quizlet, Pleko), online chat (HelloTalk, WeChat), storytelling websites, dramas and podcasts, especially those which integrated sound and text, *'where you can read it and listen to at the same time. so that's quite helpful' (level 3 RS)*. They also used traditional teachers' techniques for handwriting, such as repetition of correct stroke order. *'I try to learn how they are written and practice those characters using my hand rather than relying on digital input' (level 1 MAB),* but others were aware of a perceived "risk" of using apps and other digital formats which might lead to forgetting this: *'the limitation for this is that if you just use the pinyin keyboard, you might not remember the stroke order. I am guilty of this sometimes' (level 3 AB)*.

Interviewees also highlighted the importance of using speaking and authentic interaction to support learning, whether in daily life or via digital sites; there remained some reluctance to engage with written Chinese even at level 3, along with frustrations in mastering structural knowledge required for reading well. Ambiguity in interpreting written Chinese at sentence level remained a barrier, as item-based word knowledge did not always equate to ability to understand meaning:

> *I know the characters. But when I have to put them into a context, I … sometimes …didn't realise that was a new word…. I know all these characters separately, but now I have to learn them with their match together, which is sort of the next step (level 3 AW).*

They acknowledged that exam format played a role - use of pinyin made the Level 1 and 2 tests *'relatively easy' (level 1 MAB)* and *'it would have been impossible to get the first foothold [without Pinyin]' (level 1 JH)*. The lack of writing in the exam at level 3 was also noted - it was '*kind of strange in the writing that you don't write with your Hanzi … I know it makes sense from a marking perspective, but it just feels like much, much easier' (level 3 RS)*.

In sum, test-takers revealed in general they had clear, if varied, goals and strategies for learning written Chinese, and displayed an encouraging degree of self-motivation and initiative. To some extent, working to HSK requirements helped achieve these goals, although there was also an element of exam washback effects in driving a focus on structural accuracy, since the exam was seen as requiring item-based recognition of words, rather than being relevant to real-world use of written Chinese. However, as in the quantitative findings, there does not seem to be evidence to support a general assumption of increased mastery of written Chinese and reduced reliance on pinyin by level 3, nor any increased consistency in preferred strategy use.

These findings are valuable for teachers and assessors to understand the range of individualised approaches to mastering written Chinese, in line with other research into learner strategies (e.g. Jiang & Cohen, 2012). Teachers are urged to develop more consistent effective teaching strategies to help learners, to build confidence in learning written Chinese using as much authentic task input and literacy practice as possible. Also, acknowledging that typing in pinyin will remain a high priority for learners is important in class and in HSK assessment, given the ease with which this opens up access to using Chinese. Teachers could therefore reduce expectations of accuracy on handwriting characters at the start, but encourage what appears to be a strong intrinsic interest in mastering characters, and incentives for character recognition, so learners could then develop greater character knowledge and accuracy in writing in a more linguistically scaffolded way. Furthermore, we call for further research on how to improve and support teacher strategies as well as learner strategies, as approaches to effective Chinese language pedagogy continue to evolve (Wright, 2019; Wu & Duan, 2016; Zhao, 2016).

5 Conclusion

This study investigated whether people taking the Chinese HSK language test at levels 1 to 3 showed specific patterns of motivation, learning strategies, and goals for mastering written Chinese as they progressed from beginner to intermediate level. Using both surveys and interviews, the quantitative and qualitative data in fact revealed a complex mix of individual goals and learning practices. Learners were found to have very high motivation in choosing to start to learn Chinese in the first place, but a high barrier to mastering written Chinese and a sustained reliance on pinyin remained evident at level 3. This may not be helped by a frequent classroom emphasis on structural accuracy over meaningful communicative abilities, despite this being one of the strong motivations for the majority of our cohort. The HSK exam was also found to create some washback effects - both beneficial in motivating learning and as a marker of progress, but also negative effects by focusing on item-based traditional word learning rather than supporting real-world authentic Chinese. Given that digital Chinese usage (using pinyin keyboards) may well be the way of the future, it is timely to review the way teachers teach and the way the HSK exams assess learners' grasp of written Chinese at these early levels. The need is emphasised for better, more consistent teaching practices, using the best training in using task-based approaches, and awareness of the latest scholarship on teaching written Chinese, to ensure consistent communicative yet accurate learning. Further research in and out of the classroom is therefore needed to clarify how best to teach and examine Chinese in the years ahead.

Note

Appendix: Survey questions and codes

Motivation questions - PM: Personal Motivation; CM : Community Motivation ; GM: Global Motivation

PM1	I find learning Chinese interesting.
PM2	I study Chinese because I am interested in Chinese culture.
PM3	I can imagine myself speaking Chinese with Chinese people.
PM4	I think I have a good ability to learn Chinese.
PM5	I put a lot of effort into practicing Chinese, even outside of the classroom.
PM6	I want to use the exam as a goal to show I am making progress.
CM1	My parents encourage me to study Chinese.
CM2	I want to take this exam because if I don't, I feel I will be letting other people down.
CM3	I work hard in my Chinese classes because my teachers encourage me to do well.
CM4	I want to take the exam because my friends also take the exam.
CM5	I study Chinese because I must take at least one foreign language course while at school.
CM6	I take this exam to gain an extra qualification which will help me stand out in university, college or work applications.
GM1	I study Chinese because I think it will be useful in finding a good job in a global market.
GM2	As China becomes more important internationally, I feel I should try to learn more about China.
GM3	Learning Chinese is important to me, as I would like to study in China in the future.
GM4	I study Chinese because it will allow me to travel in Chinese speaking countries more easily.
GM5	This exam is important to show an employer I know more than one language.
GM6	I study Chinese because it has been promoted by others, such as the UK Government or education authorities, as a useful skill.

Learning Strategy questions - Communicative (Comm) or Structure-Based (SB)

COMM1	I use apps, e-dictionaries to help me in learning words.
SB1	I repeat words until I have learned their Pinyin and tones.
COMM2	I think of a visual relationship between characters and their Pinyin.
COMM3	I try to use new Chinese words or expressions in a sentence so I can remember them.
SB2	I have practised using mock exams (at least 3 times).
SB3	I write every new Chinese word or grammar point at least 3 times to help me remember them.
COMM4	I try and find ways to use the Chinese language I know outside class.
COMM5	I try not to translate word for word from English to Chinese.
COMM6	I first read over a Chinese passage quickly, then go back and read carefully.
COMM7	I use online materials or websites to help improve my listening and reading.

SB4	I look for words or grammar in my own language that are similar to Chinese to help me remember.
COMM8	To understand new Chinese words, I guess from the context.
COMM9	I don't worry if I don't understand every word.
COMM10	In listening tasks, I try to guess what the other person will say next in Chinese.
SB5	In reading tasks, I look for characters I have learned to help me get the meaning.
SB6	I always check with the teacher to make sure I've understood words or grammar points I will need for the exam.
SB7	I have clear goals for improving my Chinese skills for the exam.
SB8	I find it easy to make time to practice for the exam.
SB9	I will give myself a reward if I do well in the exam.
SB10	I notice my mistakes and use that information to help me when preparing for the exam.

Written Chinese strategies - LC: Learning Characters; RW: Reading Words; PIN: Pinyin; SPK: Speaking; WD Word Learning (level 3)

Level 1&2

LC1	I do not find Chinese characters difficult to learn.
LC2	We cannot learn Chinese well if we do not learn the characters.
LC3	Remembering the structure (e.g., top-bottom, left-right, etc.) of a Chinese character is difficult.
LC4	Remembering the strokes in a character is difficult.
LC5	The radicals in characters helps me to remember their meanings.
LC6	I do not find writing characters is more difficult than pronunciation, grammar, listening, reading and speaking.

RW1	My teacher has taught us the stroke order of a character.
RW2	My teacher frequently checks through exercises if we have remembered the characters we have learned.
RW3	The HSK exam is not difficult because I don't have to write characters.
RW4	Writing skills in Chinese is not as important as skills of listening, speaking and reading.
PIN1	Pinyin has made it easier for me to learn Chinese.
PIN2	Reading sentences or short texts in Pinyin is much easier than reading in characters.
PIN3	I usually put Pinyin together with the new vocabulary to help me to remember them.
PIN4	I prefer typing Pinyin on my phone or computer to handwriting characters.
PIN5	Learning Pinyin is still the priority in my Chinese classes.
PIN6	If I can't write characters in a test, Pinyin is accepted by my teacher.
PIN7	Chinese exams for my level do NOT need to have Pinyin with the characters.
PIN8	I can achieve higher marks if exams have both Pinyin and characters.
SPK1	Tones are difficult to learn and get them correct.
SPK2	Pinyin is the pronunciation of the Chinese language NOT the written language.

Level 3

LC7	I find it easy to remember the structure and the strokes in a character.
LC8	Writing characters following stroke orders helps me to remember them.

LC9   Radicals in Chinese characters helps me to guess and remember their meanings (e.g., etc.).
LC10   Hand-writing character is most difficult compared to other skills.

WD1   I can guess the meaning of a Chinese word made of two characters I have learned.
WD2   I have tried to remember the characters I have learned in the ways I find effective.
WD3   Hand-writing Chinese character is not so important nowadays.
RW5   My teacher insists that we remember the meanings of the characters we have learned and how to write them.

References

Allen, J. R. (2009). Why learning to write Chinese is a waste of time: A modest proposal. *Foreign Language Annals, 41*(2), 237-251. https://doi.org/10.1111/j.1944-9720.2008.tb03291.x

Bachman, L. F. (1990). *Fundamental considerations in language testing.* Oxford University Press.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing, 25*(1), 15-37. https://doi.org/10.1177/0265532207083743

Cheng, L., & Curtis, A. (2012). Test impact and washback: Implications for teaching and learning. In C. Coombe, B. O'Sullivan, P. Davidson, & S. Stoynoff (Eds.), *Cambridge guide to second language assessment* (pp. 89-95). Cambridge University Press.

Cheng, L., Klinger, D., Fox, J., Doe, C., Jin, Y., & Wu, J. (2014). Motivation and test anxiety in test performance across three testing contexts: The CAEL, CET, and GEPT. *TESOL Quarterly, 48*(2), 300-330. https://doi.org/10.1002/tesq.105

Cohen, A. D., & Macaro, E. (2007). *Language learner strategies: Thirty years of research and practice.* Oxford University Press. Confucius Institute for Scotland. (n.d.). Scholarships to China. Retrieved January 26, 2020, from https://www.confuciusinstitute.ac.uk/education/scholarships-to-china/

Csizér, K., & Dörnyei, Z. (2005). The internal structure of language learning motivation and its relationship with language choice and learning effort. *The Modern Language Journal, 89*, 19-36. https://doi.org/10.1111/j.0026-7902.2005.00263.x

Dass, L., & Akmar, S. (2007). Test taker characteristics: factors that affect performance. In Y. L. K. Ambigapathy Pandian, P. Kell (Eds.), *Innovation and Intervention in ELT– Pathways & Practices* (pp. 372-384). Universiti Putra Malaysia Press.

Diao, L. (2013). *Learning to Read in School Chinese Classes: processes and strategies of German pupils*. Logos Verlag Berlin.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies.* Oxford University Press.

Gao, Y., Zhao, Y., Cheng, Y., & Zhou, Y. (2004). Motivation types of Chinese university undergraduates. *Asian Journal of English Language Teaching, 14*, 45-64.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Edward Arnold.

Green, A. (2006). Watching for washback: Observing the influence of the international English language testing system academic writing test in the classroom. *Language Assessment Quarterly, 3*(4), 333-368. https://doi.org/10.1080/15434300701333152

Griffiths, C. (2003). Patterns of language learning strategy use. *System, 31*(3), 367-383. https://doi.org/https://doi.org/10.1016/S0346-251X(03)00048-4

Guder, A., Jiang, X., & Wan, Y. (2007). *The cognition, learning and teaching of Chinese characters*. Beijing Language and Culture University Press.

Hanban. (2008). *Chinese Proficiency Test (HSK)*. http://english.hanban.org/node_8002.htm

Hong-Nam, K., & Leavell, A. G. (2006). Language learning strategy use of ESL students in an intensive English learning context. *System, 34*(3), 399-415. https://doi.org/https://doi.org/10.1016/j.system.2006.02.002

Hsiao, Y. P., & Broeder, P. (2013). Applying the writing scales of the Common European Framework of Reference for Languages to the new HSK test of proficiency in Chinese: Realities, problems and some suggestions for Chinese language teachers and learners. *Language Learning in Higher Education, 2*(1), 59-74. https://doi.org/10.1515/cercles-2012-0004

Jiang, X., & Cohen, A. D. (2012). A Critical Review of Research on Strategies in Learning Chinese as Both a Second and Foreign Language. *Studies in Second Language Learning and Teaching*, *2*(1), 9-43.

Jin, Y. (2000). The washback effects of College English Test-Spoken English Test on teaching. *Foreign Language World, 118*(2), 56-61.

Larson-Hall, J. (2015). *A guide to doing statistics in second language research using SPSS and R*. Routledge.

Lee, J., Liu, M., Lam, C. Y. , Lau, T. O., Li, B., & Li, K. (2017). Automatic difficulty assessment for Chinese texts. *Proceedings of the IJCNLP 2017, System Demonstrations* (pp.45-58).

Li, Q. (2016). Changes in the motivation of Chinese ESL learners: A qualitative investigation. *English Language Teaching, 10*, 112-122. https://doi.org/10.5539/elt.v10n1p112

Lu, Y. (2017) Exploring the criterion-validity of HSK Level 3 and 4: are assessment and CEFR standards related. In Y. Lu (Ed.), *Teaching and learning Chinese in higher education: Theoretical and practical issues* (pp. 13-34). Routledge.

Lu, Y. (in press). *Assessing L2 competence in Chinese*. Routledge.

Lu, Y., & Song, L. (2017). European benchmarking Chinese language: Defining the competences in the written language. In Y. Lu (Ed.), *Teaching and learning Chinese in higher education: Theoretical and practical issues* (pp. 13-34). Routledge.

Manna, V. F., & Yoo, H. (2015). Investigating the relationship between test taker background characteristics and test performance in a heterogeneous English as a Second Language (ESL) test population: A factor analytic approach. *ETS Research Report Series, 2015*, 1-23.

Mayumi, K. & Zheng, Y. (2021). Becoming a speaker of multiple languages: an investigation into UK university students' motivation for learning Chinese. *The Language Learning Journal.* 1-15. https://doi.org/ 10.1080/09571736.2021.1996621

Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Heinle & Heinle

Oxford, R. (2011). *Teaching & researching: Language learning strategies*. Longman.

Peng, Y., Hong, E., & Mason, E. (2014). Motivational and cognitive test-taking strategies and their influence on test performance in mathematics. *Educational Research and Evaluation, 20*(5), 366-385. https://doi.org/10.1080/13803611.2014.966115

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications*. Prentice Hall.

Reeve, C. L., & Lam, H. (2007). The relation between practice effects, test-taker characteristics and degree of g-Saturation. *International Journal of Testing, 7*(2), 225-242. https://doi.org/10.1080/15305050701193595

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54-67. https://doi.org/10.1006/ceps.1999.1020

Seymour, P. H. (2006). Theoretical framework for beginning reading in different orthographies. In R. J. Malateshal & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 441-462). Lawrence Erlbaum Associates.

Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of Chinese. *System, 33*(1), 49-68. https://doi.org/https://doi.org/10.1016/j.system.2004.11.001

Sheorey, R., & Mokhtari, K. (2001). Differences in the metacognitive awareness of reading strategies among native and non-native readers. *System, 29*, 431-449. https://doi.org/10.1016/S0346-251X(01)00039-2

Stricker, L. J., & Rock, D. A. (1995). Examinee background characteristics and GRE general test performance. *Intelligence, 21*, 49-67.

Tavakoli, M., Shahraki, S. H., & Rezazadeh, M. (2012). The relationship between metacognitive awareness and EFL listening performance: Focusing on IELTS higher and lower scores. *The Journal of Language Teaching and Learning, 2*, 24-37.

Wang, J., Spencer, K., & Xing, M. (2009). Metacognitive beliefs and strategies in learning Chinese as a foreign language. *System, 37*(1), 46-56. https://doi.org/10.1016/j.system.2008.05.001

Wang, J., Wright, C., An, N., & Wang, S. (2022) Is OPI effective? A comparative study of the ACTFL OPI and non-OPI [Manuscript submitted for publication]. *Foreign Language Annals.*

Whong, M. (2013). A linguistic perspective on communicative language teaching. *The Language Learning Journal, 41*(1), 115-128. https://doi.org/10.1080/09571736.2011.625097

Wright, C. (2018). Effects of time and task on L2 Mandarin Chinese language development during study abroad. In C. Sanz & A. Morales-Front (Eds.), *The Routledge handbook of study abroad research and practice* (pp. 166-180). Routledge.

Wright, C. (2019). Developing communicative competence in adult beginner learners of Chinese. In C. Shei, M. Zikpi, & D.-L. Chao (Eds.), *Routledge handbook of Chinese language teaching* (pp. 134-148). Routledge.

Wu, Y., & Duan, W. (2016). 后方法时代的教学研究：不同认知风格的汉语 教师在课堂教学策略运用上的差异 [Teacher research in the post-method era: Different applications of classroom teaching strategies for CSL/CFL teachers with different cognitive styles].*语言教学与研究[Language Teaching and Linguistic Studies]*, 2, 40–52.

Xinhua News. (2017). 数据：2017年各类汉语考试考生人数达650万 *(Data: In 2017, the number of test takers for various Chinese tests reached 6.5 million)* https://baijiahao.baidu.com/s?id=1601846353711220976&wfr=spider&for=pc

Yan, P., & Xie, X. (2014). 欧洲汉语能力基准项目能做说明出探——以阅读理解为例 [On the Can-Do statements of European Benchmarking Chinese Language Project: Taking receptive writing as an example]. *考试研究 [Examination Research], 44*, 17-23.

Yu, B. (2010). Learning Chinese abroad: the role of language attitudes and motivation in the adaptation of international students in China. *Journal of Multilingual and Multicultural Development, 31*(3), 301-321. https://doi.org/10.1080/01434631003735483

Zhang, D., & Lin, C.-H. (2017). *Chinese as a second language assessment*. Springer Singapore.

Zhang, L., & Zhang, L. J. (2013). Relationships between Chinese College test takers' strategy use and EFL reading test performance: A structural equation modeling approach. *RELC Journal, 44*, 35-57. https://doi.org/10.1177/0033688212463272

Zhang, P. (2007). 语文分开、语文分进的教学模式 [The teaching model of separate Chinese]. *汉字文化[Chinese Characters], 75*(1), 64-68.

Zhao, J. M. (2010). 初级汉语教学的有效途径——"先语后文"辩证 [An effective way of elementary Chinese teaching-dialectics of "language first, text later"]. *世界汉语教学 [Chinese Teaching in the World] 25*(3), 376-387.

Zhao, Q. (2016). 来华留学预科生汉语口语水平调查及其教学启示 [An Investigation into Chinese Spoken Language Proficiency for Preparatory Students in China]. *世界汉语教学[Chinese Teaching in the World]*, *30*(3), 368–378

Zheng, Y., Zheng, Y. & Zhang, J. (2021). HSK试卷架构对 1-3级考生成绩的影响:以英语母语者为例 [The impact of HKS exam paper costruction on test-takers' performance: among English speakers]. *International Chinese Language Education, 6*(3), 50-59. https://navi.cnki.net/knavi/journals/GZZY/detail

Clare Wright is Associate Professor in Linguistics and Language Teaching at the University of Leeds. Her principal research focuses on pedagogy for early-level learners of Mandarin for communicative purposes.

Yang Lu has been an assistant professor of Chinese Language at the University of Nottingham. Her work focuses on the relationship between learners' discourse competence and their overall L2 competence, especially in assessment.

April Jiawei Zhang is Assistant Professor in the University International College at Macau University of Science and Technology. Her research interests lie in language testing and assessment, teacher education, and English for Academic Purpose.

Lulu Zhang is a Ph.D. student in the Modern languages and Linguistics Programme at the University of Southampton. Her research focuses on language testing and assessment, and teacher education.

Ying Zheng is Associate Professor in Modern Languages and Linguistics at the University of Southampton. Her main research interests include language test psychometric validation and language teachers professional development.

以测验学还是以测助学？英国 HSK 1-3 级考生动机与语言学习策略的探索性研究

克莱尔·莱特
利兹大学，英国

路扬
诺丁汉大学，英国

张家维
澳门科技大学，中国

张露露
郑英
南安普顿大学，英国

摘要

这项研究深入评估了中国 HSK 初级考生的动机和学习策略。我们使用定量调查和定性访谈，招募了 71 名英国 HSK 初级（HSK 1/2）到中级（HSK 3）考生，探究考生对 HSK 备考的态度，以及普遍使用的汉语学习策略，同时检测人口统计因素（如性别）是否在学习过程中具有影响。由于从 HSK3 级开始需要使用汉字，因此本文预测该级别考试与 1/2 级存在差异，然而，研究结果没有显示显著差异。尽管学习者，特别是女性，可能会开始于形式化的学习策略，但在 HSK（1-3）各级中发现，学习者更多地使用以意义为核心的学习策略，且内在动机的评分普遍高于外在动机。在 HSK 考试初级纳入拼音让备考看似更容易，但初学者也同时表达了学习汉字的决心。这些发现能够有效地发展针对初学者的对外汉语教学，具有广泛的教学意义。

关键词

动机，语言学习策略，HSK 1-3 级，应试者特征

克莱尔·莱特博士是利兹大学语言学和语言教学副教授。她的主要研究重点是以交际目的的初级汉语学习者的教学法。

路扬博士是诺丁汉大学中文系助理教授。她的研究专注于学习者的语篇能力与 L2 整体能力之间的关系，尤其是两者在评估中的关系。

张家维博士是澳门科技大学国际学院助理教授。她的研究兴趣是语言测试与评估，教师发展和学术英语。

张露露，英国南安普顿大学在读博士，研究领域包括语言测试与评估，教师发展。

郑英博士是南安普顿大学现代语言和语言学系副教授。她的主要研究领域包括语言测试中的量化验证及语言教师的职业发展。