Language aptitude in the visuospatial modality: L2 British Sign Language acquisition and cognitive skills in BSL-English interpreting students

Freya Watkins[1], Stacey Webb[2], Christopher Stone[3], Robin L. Thompson[1*]

[1]Multimodal Multilingual Language Processing Lab, School of Psychology, University of Birmingham, Birmingham, UK

[2]School of Social Sciences, Languages & Intercultural Studies, Heriot-Watt University, Edinburgh, UK

[3]School of Social, Historical and Political Studies, University of Wolverhampton, Wolverhampton, UK

* Correspondence:
Robin L. Thompson
r.thompson@bham.ac.uk

1.   Abstract

Sign language interpreting (SLI) is a cognitively challenging task performed mostly by second language learners (i.e., not native signers). SLI students must first gain language fluency in a new visuospatial modality and then move between spoken and signed modalities as they interpret. As a result, many students plateau before reaching working fluency and training program drop-out rates are high. However, we know little about the requisite skills to become a successful interpreter: the few existing studies investigating SLI aptitude in terms of linguistic and cognitive skills lack baseline measures. Here we report a three-year exploratory longitudinal skills assessments study with British Sign Language-English SLI students at two universities (*n*=33). Our aims were two-fold: first to better understand the prerequisite skills that lead to successful SLI outcomes; second to better understand how signing and interpreting skills impact other aspects of cognition. A battery of tasks was completed at four time points to assess skills including but not limited to: multimodal and unimodal working memory, 2-dimensional and 3-dimensional mental rotation and English comprehension. Dependent measures were BSL and SLI course grades, BSL reproduction tests and consecutive SLI tasks. Results reveal that initial BSL proficiency and 2D mental rotation were associated with selection for the degree program, while visuospatial working memory was linked to continuing with the program. 3D mental rotation improved over the course of the degree, alongside some limited gains in auditory, visuospatial and multimodal working memory tasks. Visuospatial working memory and mental rotation were the skills closest associated with BSL and SLI outcomes, in particular those tasks involving sign language production, thus highlighting the importance of cognition related to the visuospatial modality. These preliminary data will inform SLI training programs from applicant selection through to curriculum design.

2.   Introduction

1. Background and situation of SLI training in universities

Sign language interpreting (SLI) is known to be a uniquely challenging task, but few studies have investigated the linguistic and cognitive skills that make a prospective interpreter more likely to succeed. In Great Britain, one of the main routes to becoming a sign language interpreter is completing a three to four-year SLI undergraduate degree, where students acquire the target sign language (British Sign Language; BSL) alongside developing their interpreting skills. Both sign language learning and learning to interpret are challenging and distinct endeavors and likely due to this challenge, SLI degree programs suffer from high drop-out rates (see e.g. Grbić, 2009; Leitner, 2012; Huhtinen, 2014). First, unlike spoken language interpreting, where many interpreters are bilingual in both their working languages from an early age, the majority of SLI students do not enter programs with pre-existing sign language fluency and thus their initial second language (L2) acquisition occurs within a university, and not community, context (Cokely, 1986; Peterson, 1999; Monikowski and Peterson, 2005). Further, there may be special challenges involved in learning an L2 in a different modality. Concerningly, there is evidence that the academic demands of SLI degrees mean that students have fewer opportunities to engage with the deaf community (Pivac, 2014); that SLI students overestimate their sign language fluency and interpreting skills (Beal et al., 2018; Robinson and Henner, 2018); and that many do not attain working sign language fluency even by the end of their degree programs (see e.g. Volk, 2014). Together, these contribute to a gap between SLI training completion and competent practice (Witter-Merithew and Johnson, 2005).

As well as learning a L2 in a new modality during their degree program, SLI students also have to learn how to interpret. SLI is both cognitively and linguistically demanding, involving simultaneous use of two languages in two different modalities (Padden, 2000). However, there is minimala lack of research into whether a cognitive aptitude profile exists for L2 SLI students embarking on study with the aim of reaching professional fluency. Here we follow López Gómez et al. (2007), who found that perceptual-motor skills and cognitive verbal abilities played a greater role than personality in predicting SLI students' sign language proficiency, suggesting that greater focus should be placed on cognitive predictors of signing and interpreting outcomes, as does Stone (2017). While a lot of research on cognitive aptitude for spoken interpreters exists, some of this is modality-specific and only applicable to the spoken-language aspects of SLI, and is less informative regarding the cognitive and linguistic skills required to interpret sign language in the visuospatial modality. Another important difference is that spoken language interpreters work primarily from their L2 into their L1, whereas most signed language interpreters work primarily from spoken L1 into signed L2 (Nicodemus and Emmorey, 2013). Assessing linguistic and cognitive aptitude for SLI prior to entry into interpreter training programs could help reduce drop-out rates, and minimize the rehousing of struggling SLI degree students into Deaf Studies programs (Stone, 2017), and would ultimately save a lot of time, both for instructors and students. According to Grbić (2009), beginner SLI students are often motivated by "social goodwill" but are less aware of the cognitive, social and emotional demands of SLI, which initial pre-screening may help to highlight. Assessment at intake is thus not aimed at discouraging L2 sign language learning or potential SLI students, but instead encourages them to recognize as early as possible that SLI is challenging for various reasons.

Importantly, incomplete L2 signed language acquisition or insufficient skill in interpreting results in a lack of language access for deaf people who use SLI services and hearing people who don't sign. In the UK, demand for SLI services frequently exceeds supply (DWP, 2017). This can in turn lead to the use of unqualified, non-professional language brokers who do not meet the national standards for interpreting (CFA, 2012). Understanding the linguistic and cognitive factors which are important for both successful L2 signed language acquisition and high-level interpreting skill is thus key to

improving access for deaf people. Furthermore, training SLI students involves a significant financial investment. Whether students self-fund their studies or are supported by government grants, it is vital to ensure that the limited resources in SLI training (Webb, 2017) are used sensibly and those students with optimum potential are supported. As Hunt and Nicodemus (2014) point out: "[o]ne of the problems with gatekeeping is timing; by the time a student gives evidence that they are not suitable, they have already invested a great deal of time and money in their degree." While aptitude testing on admission is considered integral and is standard practice in spoken language interpreting courses (see e.g. Timarová and Ungoed-Thomas, 2008), SLI degree programs in the UK do not presently use any research-informed methods to test applicants in terms of their suitability (Stone, 2017). In the US, the assessment of cognitive skills was found to be only a minor aspect of US SLI programs' entrance requirements (Marks, 2014), with most institutions focusing only on American Sign Language (ASL) and English skills, and in Australia, initial SLI degree screening has been found to be informal and often not evidence-based (Bontempo and Napier, 2009). Initial data on which cognitive and linguistic assessments are likely related to SLI aptitude would therefore also be a benefit to prospective students and SLI programs.

2.   Cognitive and linguistic skills in spoken language interpreting

In spoken language interpreting, it has long been known that cognition skills like working memory and cognitive load are vital to the interpreting process (see e.g. Gile, 1997) and thus an important consideration for interpreter educators. As a result, cognitive aptitude for spoken language interpreting has been investigated to a greater extent than for SLI and research now spans several decades (for review, see Russo, 2011). This means that some aptitude test batteries have been validated for their reliability and are now widely used to screen spoken language interpreting trainee candidates or for intensive language training programs. For example, many interpreter training programs use long-standing commercial aptitude batteries like the Modern Language Aptitude Test (MLAT, Carroll and Sapon, 1959; Carroll et al., 2010) or the Pimsleur Language Aptitude Battery (Pimsleur, 1966), which include tasks involving learning new vocabulary and phonetic discrimination.

More recently, research has focused in particular on the importance of working memory (WM) in spoken language interpreting, however the findings are mixed. Both improvements in WM after interpreting training (Babcock et al., 2017; Chmiel, 2018; Christoffels, de Groot and Kroll, 2006) and WM effects on simultaneous interpreting fluency (Lin, Lv and Liang, 2018) have been reported. Auditory WM has also been shown to be more important than social factors like personality in simultaneous spoken language interpreting (Anssari-Naim, 2021), and L2 auditory WM correlated with consecutive interpreting performance (Dong, Liu and Cai, 2018). However, some studies found no effects of WM: for example, professional interpreters were no different in WM capacity to beginner interpreter students (Liu, Schallert and Caroll, 2004); linguistic factors such as word knowledge in L1 and L2 were more important for interpreting performance than increased WM capacity (e.g. Padilla, Bajo and Macizo, 2005), and interpreting training has been found to improve language processing skills, but not WM (Tzou et al., 2012). In terms of other cognitive skills, spoken language interpreters have been shown to have superior cognitive flexibility over bilinguals with no interpreting training (Yudes, Macizo and Bajo, 2011) as well as superior dual-task attention compared to non-interpreters (Strobach et al., 2015; Morales et al. 2015). Macnamara (2009) also provides a review of cognitive functions and capacities required for interpreting, including chunking, online decision-making and processing speed.

In the present preliminary study, we focus on sign language interpreters who, like spoken language interpreters, make use of auditory WM, but also need to rely on visuospatial WM. According to Baddeley's model of WM (Baddeley, Hitch and Allen, 2009; Baddeley, 2012), auditory memory and visual memory are maintained through separate functional components: the phonological loop, and the visual sketchpad respectively. This suggests that there may be different memory requirements or processes for SLI compared to spoken language interpreting, given the different modalities that must be attended to.

3. Existing work on linguistic and cognitive aptitude for SLI

As with spoken language interpreting, there has been a long-standing interest in establishing what skills are required to be successful in SLI (for review, see Nakano, 2021). Some studies have taken the approach of surveying and studying the attributes of qualified sign language interpreters (e.g. Jones and Quigley, 1974; Shaw and Hughes 2006; Seal, 2004), the latter finding that visual attention while inhibiting distractors was a particularly important skill. Other studies have compared skills in sign and spoken language interpreting students. For example, Shaw (2011) found better visual memory skills and concentration in SLI students than in spoken language interpreting students, both in terms of retention duration of visual information and performance when visual distractors were present. Other linguistic and cognitive aptitudes that have been mentioned in the literature for both spoken language interpreting and SLI include high command of both working languages, verbal fluency, processing speed, good WM and concentration, and for SLI alone a further factor was the capacity to sign and talk simultaneously (Frishberg, 1986; Lara and de los Santos, 2000; cited in López Gómez et al., 2007).

Researchers have also explored the role of WM in SLI, again with mixed results. Wang (2016) found no evidence that auditory WM capacity in English or visual WM in Auslan related to SLI task performance in either direction. However, van Dijk et al. (2012) found that auditory WM span in Dutch and visual WM in Sign Language of the Netherlands were related to the quality of interpretations by SLN-Dutch interpreters. Looking at domain-general cognitive skills beyond WM, Macnamara et al. (2011) found that highly skilled ASL-English interpreters had greater mental flexibility, faster cognitive processing, psychomotor speed and were better at task switching when compared to less skilled interpreters. Haug, Nussbaumer and Stocker (2019a; 2019b) describe initial data from working Swiss German Sign Language (DSGS)-German interpreters on a battery of cognitive tasks, developing their own versions of WM tasks featuring DSGS stimuli. These preliminary results suggest that the cognitive abilities of interpreters on all normed tests are above average, and that weaknesses or average performance in certain cognitive areas may be compensated by strengths in others. Studies on cognition in qualified working interpreters can inform us about the 'end results' of SLI training and professional experience, but are less informative regarding prospective SLI students: it is not clear which skills improved through SLI training or professional experience, or if some skills should ideally be at a certain level at the outset of learning.

In terms of students, some SLI educators have administered broad pre-admission test batteries before SLI training, finding predictive relationships between performance in the screening battery and later SLI outcomes (e.g. Humphrey, 1994; Bontempo and Napier, 2009). However, only the cumulative effect of the test batteries is discussed and not the individual assessments, meaning it is hard to be sure what to attribute predictive power to, nor were the test batteries re-administered at later points in order to understand progression in different areas. To our knowledge, there have only been a few studies that have investigated cognitive aptitude for SLI trainees over the course of their studies (López Gómez et al., 2007; Stone, 2017; Macnamara and Conway, 2016). López Gómez et al. (2007)

~~found that perceptual-motor skills and cognitive verbal abilities played a greater role than personality in predicting sign language proficiency, suggesting greater attention should be paid to cognitive predictors of signing and interpreting outcomes. Specifically, they~~ López Gómez et al. (2007) found that a nonsense-sign repetition test was a good predictor of successful Spanish Sign Language acquisition, whereas for Stone's cohort, the MLAT number-learning test was predictive of students' BSL exam results. Both studies argued these tasks directly or indirectly relate to how the structure of signs is encoded. L2 sign language learners are known to struggle with phonological processing in the new visuospatial modality (e.g. Ortega and Morgan, 2015; Williams and Newman, 2016). This has also been shown in deaf signers who learn a sign language as an L2 or late L1 and experience a 'processing bottleneck' at the phonological level (Mayberry and Fischer, 1989). Another study that has taken a re-test approach to cognition in SLI trainees is Macnamara and Conway (2016), who administered a battery of cognitive tests targeting WM capacity and SLI performance at four points throughout an ASL-English SLI training program. Their main findings were that WM capacity predicted initial SLI performance and was an even stronger predictor of final SLI performance. Additionally, students who performed well initially maintained a high level of performance, whereas those who performed poorly initially benefited more from the SLI training, but not enough to catch up to the higher level. Despite these initial longitudinal investigations into student SLI cognitive aptitude, none of these studies assessed students' baseline skills before the start of their program, meaning that performance could have been already changed by sign language teaching or other factors.

4.   Cognitive adaptations from L2 sign learning

Research has shown that fluent signers outperform non-signers on several measures of visuospatial ability like mental rotation (MR) and image generation (for review, see Emmorey, 1998; 2002), suggesting skills-based enhancements from exposure to language in the visuospatial modality. In particular, MR has been shown numerous times to be improved as a result of sign language experience (e.g. Emmorey et al., 1998; McKee, 1987; Talbot ~~&~~ and Haude, 1993), indicating it is an important skill required during sign language use. As well as needing to mentally rotate to understand e.g. topographical uses of signing space, Watkins, Abdlkarim and Thompson (2018) also suggest that MR skills are crucial to BSL comprehension when perceiving signs from the side, as opposed to a face-to-face orientation. Side-on comprehension is needed in many real-life situations, e.g. group conversations, or in SLI scenarios such as conferences, yet is often neglected in teaching materials and methods.

Some skills-based enhancements in visuospatial WM and MR extend into late L2 signers (e.g. Keehner and Gathercole, 2007; Kubicek and Quandt, 2021). However, like many interpreting aptitude studies, they lack baseline measures before sign learning or critical SLI instruction began. For example, Keehner and Gathercole found improved visuospatial WM in fluent late L2 signers working as BSL interpreters, but the authors acknowledge that their participants may simply have been spatially adept before learning to sign, which in turn may have facilitated successful L2 sign language acquisition. Thus, it is unclear whether such effects derive directly from signing experience, or whether those reaching fluency are predisposed with better cognitive abilities such as MR or visuospatial WM. If pre-existing visuospatial skills are found to be predictive of successful L2 sign language learning, these could then be targeted for specific training, either within a signing context or as a general cognitive skill, within SLI training programs. It has already been shown that deaf and hard of hearing children can improve their MR skills through targeted practice (Passig and Eden, 2001), which then feeds back into sign language fluency. An advantage of a longitudinal approach

with a baseline measure at the start of the SLI program means we can detect both how skills progress across time, and how individual differences in baseline skills relate to this development.

5. Present study and domains of investigation

Longitudinal studies are a good approach to these questions, as they allow the identification of developmental milestones and any individual differences in performance, as well as the time points at which learners begin to perform at the level of working interpreters.

One open question is the link between WM and modality in SLI (for further discussion, see also Wilson and Emmorey, 2003; Hall and Bavelier, 2011; Wang and Napier, 2013; Williams, Darcy and Newman, 2015). Existing studies have not always distinguished between multimodal dual-attention WM tasks requiring both auditory and visuospatial attention, and unimodal WM tasks where only one modality must be attended to. Furthermore, most of the aforementioned research on SLI and WM only includes a single WM task in just one modality, and very few studies overall have employed multimodal WM tasks. One exception is Bontempo and Napier (2009), who included a dual-task memory exercise (divided attention) in their pre-SLI-degree test battery, but they did not describe the results at task level, only discussing the composite test battery as a whole. Here we attempt to test three measures of WM: one auditory, one visuospatial and one dual WM task with multimodal input to attend to (auditory and visuospatial) simultaneously (see Fig. 1).

To our knowledge, none of the existing studies into cognitive aptitude for sign language, whether based on students or working interpreters, have specifically looked at MR (despite being highlighted as a possible predictor by López Gómez et al., they did not include a rotation task in their test battery). Here we use two different tasks: one where 2D shapes simply need to be rotated in a circle (Fig. 2a), and another that uses 3D-rendered blocks that must be rotated around the vertical axis only (Fig. 2b). It is MR in this plane which should be most like the way sign language input must be rotated while comprehended (Watkins, Abdlkarim and Thompson, 2018).

We include two linguistic tasks that assess English skills. Initial English vocabulary knowledge was found to predict self-rated ASL proficiency after 1 semester of ASL instruction (Williams, Darcy and Newman, 2017). English reading comprehension was also shown to improve during the SLI degree by Stone (2017), and we repeat the same measure here. Summarizing or 'paraphrasing' is known to be an important skill in spoken language interpreting (e.g. Moser-Mercer 1985; Russo and Pippa 2004; Russo 2014). Being able to comprehend and summarize complex spoken English before interpreting it into 'chunks' of BSL is also required of SLI students and thus we created a task to assess this paraphrasing ability. We also include the MLAT Number Learning task used by Stone (2017) as a measure of phonological encoding, which was predictive of later BSL grades (see also Martinez and Singleton, 2019, who found sign learning and word learning to be highly correlated in hearing non-signers).

As Macnamara et al. (2011) point out, cognitive skills do not exist in a vacuum, and there have been a range of studies exploring SLI aptitude in terms of more social factors like disposition and personality (Stauffer and Shaw, 2006; Bontempo 2008; Bontempo et al., 2014). Hence, we also include one personality measure of risk-taking that was predictive of continuation on the SLI degree in Stone (2017), the Barratt Impulsiveness Scale, version 11 (BIS-11; Patton, Stanford and Barratt, 1995). Specifically, Stone found that SLI students remaining on the degree program were significantly more impulsive, i.e. more likely to take risks, than students who were rehoused in a Deaf Studies program (due to choice or poor BSL/interpreting exam performance). Macnamara et al. (2011) also found that a different measure of risk-taking (Behavioral Inhibition System; Carver and

White, 1994) differentiated between highly skilled and less-skilled interpreters Lastly, we also used a non-verbal reasoning task as a control measure which we do not predict will change over time, nor impact BSL or SLI performance.

The present preliminary study asks three main research questions: (1) Do any of the cognitive and linguistic assessments predict being selected for, or continuing on, the SLI degree? (2) How do the cognitive and linguistic assessments change in time over the course of the SLI degree? (3) Are any of the cognitive or linguistic assessments associated with BSL and SLI performance outcomes? In sum, we aim to see if we can replicate previous results in some domains by using identical or similar tasks as in previous longitudinal studies, as well as investigating some new areas (MR, and manipulating the modality of WM tasks). Together, these elements could help us get a step closer to understanding which initial cognitive and linguistic skills indicate the potential of an L2 sign language learner and a successful SLI student.

3.   Materials and Methods

1.   Participants

SLI students (*n*=33) were recruited from the degree program 'MA (Hons) BSL (Interpreting, Translating and Applied Language Studies)' at Heriot-Watt University (*n*=23 total) or the degree program 'BA (Hons) Interpreting (BSL/English)' at the University of Wolverhampton (*n*=10). Two consecutive year groups of Heriot-Watt students were tested (HW1: *n*=11, HW2: *n*=12; see Table 1). Across all cohorts, 3 participants were heritage signers of BSL and were thus excluded from later analyses. The remaining L2 participants mostly had only limited exposure to BSL before beginning their interpreting program (e.g. introductory courses). A further group of prospective candidates for the first HW cohort (HW0: *n*=19) were tested at interview, but were not selected for entry into the degree program. Therefore, data only exists for this group at session 1.

Table 1: Demographics of the cohorts of SLI students who took part in the longitudinal study.

Institution	Heriot-Watt University			Uni. Wolves
Cohort	rejected	1st cohort	2nd cohort	single cohort
Cohort code	HW0	HW1	HW2	WV
N	19	11	12	10
Mean age at session 1 (yrs;mths)	24;2	22;8	27;1	26;3
Mean prior BSL exposure (y;m)	3;1	1;8	1;4	4;10
Session 1 (pre-course)	in-person	in-person	in-person	in-person
Session 2 (end of year 1)	n/a	in-person	online	online
Session 3 (end of year 2)	n/a	online	online	online
Session 4 (midway year 3)	n/a	n/a	online	online

2.   Longitudinal study design

Participants were tested on a battery of cognitive and linguistic assessments at four sessions, approximately one year apart. Session one was before beginning the program and thus, for most participants, before critical exposure to BSL had begun. Session two was at the end of the first year

of the course and the third session was at the end of the second year, before students began their placement years. The fourth and final session was halfway through this placement year. Due to disruption related to the Covid-19 pandemic (e.g. lack of equipment or suitable space for remote participation at home) as well as students dropping out of their courses or not being interested in further participation at later test sessions, it was not possible to re-test all participants at test sessions two to four. Furthermore, the online testing at later sessions was spread out more than planned compared to pre-pandemic testing, which was carried out in-person on the Heriot-Watt and Wolverhampton University campuses.

3. Description of Test Battery

First, the cognitive and linguistic aptitude assessments comprising the predictor variables are described, followed by the BSL and SLI assessments which make up the outcome variables.

1. Predictor Variables

1. *Working memory tasks*

As a measure of multimodal WM, participants completed the Psychology Experiment Building Language (PEBL; Mueller, 2014; Mueller and Piper, 2014) test battery version of the dual n-back task (Jaeggi et al., 2008). In the task, participants must simultaneously recall a sequence of letters presented auditorily as well as the spatial location of a sequence of squares presented visually on a grid. Participants press a button when the letter or square location matches the letter or square location presented *n* trials ago. The task has one block each of 1-back, 2-back and 3-back trials. The dependent measure was combined accuracy (average accuracy across both letter-matching and spatial matching).

For visuospatial WM, participants completed the PEBL version of the Corsi block-tapping task (Corsi, 1972; Kessels et al., 2000). In the task, participants must memorize the order in which a sequence of blocks change color and then click the blocks in the same order. The sequence gets progressively longer as the task goes on. The dependent measure was the number of correct responses.

As a measure of auditory WM, participants completed the PEBL test battery version of the Digit Span task (Croschere et al., 2012). In this task, participants must remember a sequence of numbers presented auditorily number-by-number, and then type in the sequence of numbers as it was heard. The sequence of numbers gets progressively longer as the task goes on. The dependent measure was accuracy of responses.
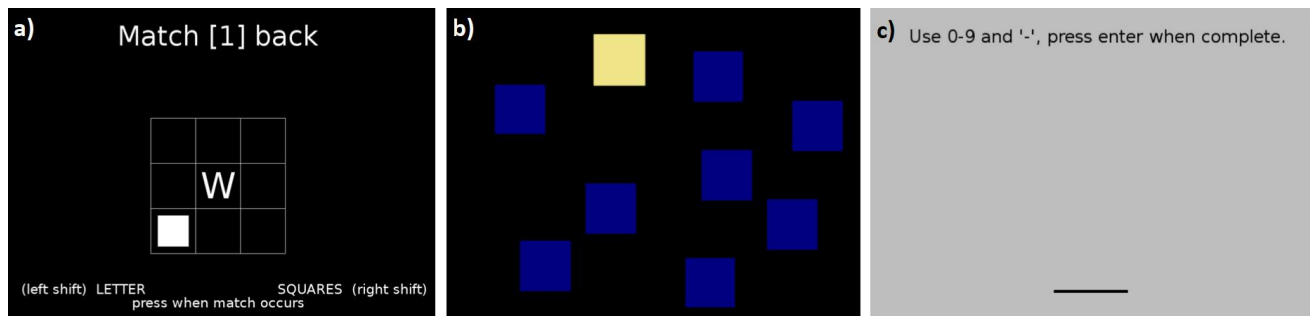

Figure 1: Example stills from the working memory tasks used in the study: a) Dual n-back task (multimodal WM); b) Corsi blocks task (visuospatial WM); c) Digit span task (auditory WM).

2. Mental rotation tasks

At sessions one and two, HW1 participants completed the simple 2D PEBL mental rotation task (Shepard and Metzler, 1988; Berteau-Pavy, Raber and Piper, 2011). The HW2 and WV groups also completed this task at sessions three and four. In this task, participants must decide if two shapes presented side-by-side on screen are the same or different by mentally rotating them. The dependent measure was a speed-accuracy trade off-score which combines RT and accuracy measures, and was calculated using the Balanced Integration Score (Liesefeld, Fu and Zimmer, 2015; Liesefeld and Janczy 2019).

For 3D mental rotation, participants completed a 3D block rotation task comprising 96 three-dimensional stimuli validated by Ganis and Kievit (2015). The task was created using PsychoPy/Pavlovia (Peirce et al., 2022) and is freely available at https://pavlovia.org/freyawatkins/block_rotation. The task is a measure of MR around the vertical axis, where participants must decide whether two shapes presented side-by-side on screen are identical or mirror images by mentally rotating them as quickly and accurately as possible. The dependent measure was again a speed-accuracy trade-off score, combining RT and accuracy into a single measure.
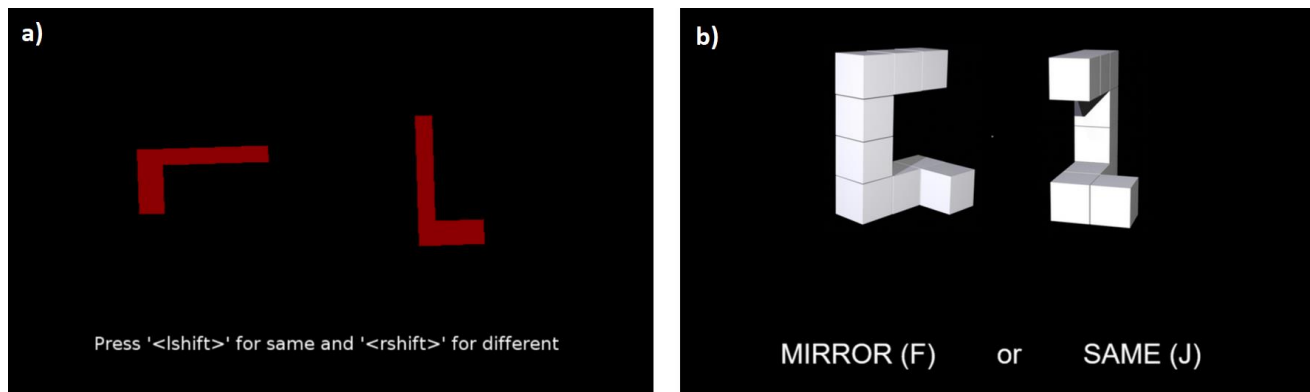
Figure 2: Example stills from the two mental rotation tasks used in the study: a) simple circular 2D shape rotation task showing an example 'same' pair rotated 90°; b) 3D-rendered block rotation task around vertical axis only, showing an example 'mirror' pair rotated 150°.

3. Phonological encoding

At session one, the HW2 and WV groups completed the Number Learning subtest of the MLAT (Carroll and Sapon, 1959; Carroll et al., 2010), which is a measure of 'auditory alertness' and phonological encoding. Participants are taught a number system in a made-up language through auditory input and tested by being asked to translate new combinations of numbers from the made-up language back into English numerals. The test features 43 items. The dependent measure was response accuracy.

4. Linguistic tasks

As a measure of English reading comprehension, participants completed the revised Kirklees version of the Vernon Warden Reading Test (Hedderly, 1996; Warden, 1956). In the task, participants must complete 42 sentences by selecting the most appropriate word to fill a gap from the five options provided. Participants were given ten minutes to complete the task, which was completed via pen and paper at face-to-face testing sessions and using a digitized version at online testing sessions. The dependent measure was test accuracy.

As a measure of summarizing ability, students had to listen to a short presentation (a TED talk on climate change and food, Ebi, 2019) and afterwards suggest a title for the presentation and summarize the presentation in five key bullet points. The dependent measure was accuracy of the summary.

5. *Non-verbal reasoning*

Participants completed the Matrices subtest from the Kaufman Brief Intelligence Test-2 (KBIT-2; Kaufman and Kaufman, 2004). In this task, participants are presented with visual stimuli with a specific rule or relationship, which participants must understand and then select the picture or pattern from the options provided that best fits that relationship or rule. Participants were given ten minutes to complete the task, which was completed via pen and paper at face-to-face testing sessions and using a digitized version and spreadsheet answer key at online testing sessions. The dependent measure was test accuracy.

6. *Impulsivity*

Participants completed the BIS-11 (Patton, Stanford and Barratt, 1995), a 30-item questionnaire, where they self-rate the frequency of their behavior and preferences regarding impulsivity. The dependent measure was the total score.

2. Outcome Variables

1. *BSL performance*

All students at session one completed a copy-sign task, which was a measure of BSL reception and production. The task consists of ten BSL signs and three short BSL sentences which were presented twice by a Deaf L1 signer and had to be reproduced by the students as accurately as possible. Accuracy of BSL production was coded by a sign language interpreter, with marks for phonological parameters (handshape, movement, location, orientation, non-manual features) articulated correctly for individual signs as well as phrasing and prosody for sentences.

At sessions three and four students were tested on the BSL Sentence Reproduction Test (BSL-SRT; Cormier et al., 2012). The test involves viewing BSL sentences of increasing complexity and reproducing them as accurately. The test is therefore an assessment of both BSL comprehension and production. Videos of BSL production were coded for accuracy by an SLI instructor, with one mark for each sentence reproduced correctly.

Module grades from BSL modules were collected from both semesters for the first two years of the course. For HW1 and HW2 students, grades included two first-year intensive practical modules in BSL and two second-year modules in Advanced BSL. For WV students, grades were made up of the first-year modules 'Intermediate BSL Enhancement for Interpreters A, B and C' and the second-year modules 'Advanced BSL Enhancement for Interpreters A, B and C'[1]. Third-year grades mostly relate to student placements and are not included in analyses. Grades follow the standard UK university grading system (0-100; whereby most marks fall between the range 40-80, see e.g. Yorke et al., 2000).

---

[1] Details of the BSL and interpreting modules at the institutions in the present study are at the links below. Heriot-Watt University: https://web.archive.org/web/20220124142725/https://www.hw.ac.uk/uk/study/undergraduate/british-sign-language-interpreting-translating-and-applied-language.htm University of Wolverhampton: https://web.archive.org/web/20220311152939/https://www.wlv.ac.uk/courses/ba-hons-interpreting-british-sign-languageenglish/

2. *SLI performance*

At session four, HW2 and WV participants completed two consecutive interpreting tasks hosted on GoReact (GoReact, 2022). In the first, they had to sequentially (not simultaneously) interpret a four-minute story from BSL into English. The story, signed by a L1 Deaf signer in an informal register, was about receiving their Covid-19 vaccine. In the second task, students interpreted a three-and-a-half-minuten instructional video explaining the procedure for visiting a Covid-19 ward in a hospital from English into BSL, with a more formal register. Students were allowed to pause the video to 'chunk' their interpreting as they saw fit, this 'chunking' skill was also part of the assessment of the interpretation. These tasks were assessed by an SLI educator and graded like university assignments (0-100). The total score for each consecutive interpreting task was calculated based on poise, style, consecutive management, comprehension, conceptual rendition, vocabulary, accuracy, repairs and an overall mark.

Module grades from second-year interpreting modules were also collected. For HW1 and HW2 students, these comprised the 'Introduction to Translation and Interpreting skills' module and for WV students, the modules 'Consecutive Interpreting 1 and 2'. Third-year grades mostly relate to student placements and are not included in analyses.

Ethics approval statement

For the initial HW1 cohort, ethical approval was gained from Heriot-Watt University. For the HW2 and WV cohorts, approval was gained from the University of Birmingham Science, Technology, Engineering and Mathematics Ethical Review Committee (ERN_18-1170). Updated ethical approval was gained for later testing sessions which took place online due to the COVID-19 pandemic. All subjects gave written informed consent in accordance with the Declaration of Helsinki, as well as video consent for tasks where participants were filmed while signing.

4. Test Procedure

Table 2: All skills assessed in the longitudinal study and the test/re-test procedure by cohort.

	Session			
	pre-degree	**1st year**	**2nd year**	**3rd year**
Skill (assessment)	*n*=33	*n*=13	*n*=11	*n*=9
Multimodal WM (Dual N-Back)	All	All	All	HW2, WV
Visuospatial WM (Corsi Blocks)	All	All	All	HW2, WV
Auditory WM (Digit Span)	HW2, WV	-	HW2, WV	HW2, WV
2D mental rotation (Shape rotation)	HW1	HW1	All	HW2, WV
3D mental rotation (Block rotation)	HW2, WV	HW2, WV	All	HW2, WV
Phonological encoding (Number Learning)	HW2, WV	-	-	-
English comprehension (Kirklees Reading)	All	All	All	-
Non-verbal reasoning (KBIT-2 Matrices)	All	All	All	-
Summarizing (TED talk task)	HW2, WV	-	-	-
Impulsivity (Barratt Impulsiveness Scale)	-	-	All	HW2, WV
BSL sign repetition (copy-sign)	All	-	-	-
BSL sentence repetition (BSL-SRT)	NA	-	All	HW2, WV
BSL (BSL module grades)	NA	All	All	placement
SLI (BSL to English task)	NA	NA	-	HW2, WV
SLI (English to BSL task)	NA	NA	-	HW2, WV

SLI (SLI module grades)	NA	NA	X	placement

Within each testing session, the order of assessments was randomized. This was dependent in part on availability of researchers to run an assessment at any given time and on computer availability for online testing.

5.   Statistical Analysis

1.   *Analysis plan*

Our exploratory analyses are divided into three sections. First, we use stepwise backward logistic regression models to determine whether any assessments were predictive of (1) being selected for, and (2) continuing on the SLI degree. The second analysis looks at changes in the linguistic and cognitive assessments over time, looking at each predictor in turn. Here we use linear mixed models with random effects structure for participants per task, with initial BSL proficiency as a covariate. Finally, the third analysis looks at correlations between the linguistic/cognitive assessments and (1) BSL performance and (2) SLI performance (grades and BSL/SLI tasks). Here we begin by looking at whether there are any relationships between the predictors at the pre-degree testing session and later outcome variables, and then look at any relationships when assessments are repeated at later testing points. We calculate R^2 values and also fit linear mixed models with initial BSL proficiency as a covariate.

2.   *Data availability and reproducibility*

We report all data exclusions (if any), all manipulations, and all measures in the study, and we attempt to follow JARS (Kazak, 2018). This study's design and its analyses were not pre-registered. In line with standards of reproducible research, the scripts and data (excluding video data and possibly identifying variables such as age) are made available with this publication and can be retrieved on the following publicly accessible repository: https://osf.io/kjctg. We used R version 4.0.5 (R Core Team, 2021) plus the packages {lme4} v1.1.27.1 (Bates et al., 2015), {sjPlot} v2.8.10 (Lüdecke, 2021) and {blorr} v0.3.0 (Hebbali, 2020) for the regression/mixed effects model analyses and output reported below, plus {effsize} v0.8.1 (Torchiano, 2020) to calculate effect sizes. For data processing and visualization we used the package {tidyverse} v1.3.1 (Wickham, 2019), for file organization {here} v1.01 (Müller, 2020); and for plotting details we used {scales} v1.1.1 (Wickham and Seidel, 2020), {PupillometryR} v0.0.3 (Forbes, 2020), {sdamr} v0.1.0 (Speekenbrink, 2021), {plotrix} v3.8.1 (Lemon, 2006) and {patchwork} v1.1.1 (Pederson, 2020).

3.   *Attrition and missing data*

Our longitudinal study was subject to high levels of attrition over time, for a combination of known and unforeseen reasons. The aforementioned high drop-out rate from SLI degrees was a known factor that we expected would greatly reduce participation at later sessions. Missing data from these participants could be considered 'missing-at-random', where degree program drop-out relates to poor grades or difficulty with the content of the program. However, the outbreak of the Covid-19 pandemic introduced a large number of reasons for data to be 'missing-not-at-random', which we could not account for with auxiliary variables (e.g. socioeconomic background, income-to-needs ratio, participant disability, parental education). For example, the pandemic necessitated online testing at sessions 2, 3 and 4. This was not equally accessible to all participants, due to differences in access to equipment, technological knowledge, time and space to participate remotely, etc.

Furthermore, pandemic-related illness also prevented some participants from repeating tasks at specific time points. While multiple imputation of missing data is possible, even when data are missing-not-at-random for reasons such as these (see e.g. Madley-Dowd et al., 2019), this approach is only considered appropriate when auxiliary variables which may have correlated with missingness are also present in the dataset. Since we did not collect initial data on these factors, imputing missing data would not have produced less biased estimates.

4.   Results

1.   Predictors of selection and remaining on SLI degree program

1.   *Predictors of selection*

The first sub-analysis examines whether any of the cognitive and linguistic assessments at session one (pre-degree) were predictive of being selected for the degree program, using data from the HW1 cohort, plus the candidates who were assessed at interview but not selected (HW0). We fit a backward stepwise logistic regression model to identify possible predictors of the binary outcome variable 'selected' (0, 1) out of the following predictor variables: multimodal WM, visuospatial WM, non-verbal reasoning, English vocabulary, 2D-MR and initial BSL self-rating. At each step, variables were chosen based on p-values, and a default p-value threshold for backward stepwise regression of 0.1 was used to set a limit on the total number of variables included in the final model. The stepwise regression reduced the predictors to just 2D-MR score ($z=1.78$, $p=0.075$) and BSL self-rating ($z=1.91$, $p=0.056$), whereby higher 2D-MR scores and higher BSL self-ratings respectively were both marginally significant predictors of associated with being selected for the degree. While we did not have enough data from the copy-sign task coded to include in the model, self-rated BSL proficiency and copy-sign task scores were significantly positively correlated ($r^2=0.38$, $t=2.4$, $p=0.043$).

2.   *Predictors of continuation*

Our second sub-analysis looks at students across the three cohorts who were selected for the degree program, i.e. excluding those who were not successful in gaining a place at interview. Here we ask whether any of the initial assessments carried out pre-degreeat session 1 predicted whether students were continuing with the degree program at the time of writing (in final or penultimate year). We again fit a backward stepwise logistic regression model to identify possible predictors of the binary outcome variable 'continuing' (0, 1) out of the following predictor variables: multimodal WM, visuospatial WM, non-verbal reasoning, English vocabulary, and initial BSL self-rating. At each step, variables were chosen based on p-values, and a default p-value threshold for backward stepwise regression of 0.1 was used to set a limit on the total number of variables included in the final model. The stepwise regression reduced the predictors of continuing with the SLI degree to just visuospatial WM ($z=-1.77$. $p=0.078$), whereby a higher visuospatial WM score significantly predicteds were marginally associated with continuing on the degree. As an additional analysis, we also modeled impulsivity at the end of second year, but this did not predict continuation in the degree.

2.   Changes in cognitive and linguistic skills during SLI program

In our second set of analyses, we look at all the cognitive and linguistic assessments in turn and examine their change over time, as students progress through the SLI program. All linear models have initial BSL skill as a covariate and random effects for participants, and use a standard p-value

threshold for linear models of 0.05. Models of English comprehension also include age at initial testing session as a covariate.

1. *Working memory tasks over time*

Multimodal WM (Dual N-Back) was significantly improved after two years of study compared to pre-degree ($t=2.32$, $p=0.02$, d=0.56; Fig. 3a). However, this improvement did not hold at the final session ($t=0.12$, $p=0.91$) compared to pre-degree.
Performance on the visuospatial WM task (Corsi Blocks) was significantly improved after one year of study compared to pre-degree ($t=2.01$; $p=0.044$, d=0.29; Fig. 3b). However, this improvement did not hold for the latter two sessions (second-year: $t=-0.06$, $p=0.95$; third-year: $t=1.52$, $p=0.13$). Auditory WM performance (Digit Span) was not significantly improved by the second -year compared to pre-degree ($t=0.53$, $p=0.59$), but third-year accuracy was significantly higher than pre-degree ($t=2.72$, $p=0.007$, d=1.57; Fig. 3c).

2. *Mental rotation tasks over time*

Performance on the 2D shape rotation task was significantly improved after one year of study compared to pre-degree ($t=2.71$, $p=0.007$[2]). Interestingly, the covariate initial BSL skill played a negative role in predicting 2D-MR performance overall ($t=-2.21$, $p=0.027$). Reaction times were significantly faster than pre-degree at all subsequent sessions (first year: $t=-4.04$, $p<0.001$, d=1.07; second-year: $t=-3.88$, $p<0.001$; third-year: $t=-2.12$, $p<0.001$). On the 3D block rotation task, speed-accuracy trade-off scores were significantly higher than pre-degree at all subsequent sessions (first year: $t=2.51$, $p=0.012$, d=1.32; second-year: $t=4.6$, $p<0.001$, d=0.91; third-year: $t=5.31$, $p<0.001$, d=1.68; Fig. 3d) compared to pre-degree performance. Likewise, reaction times on the 3D-MR task were significantly faster than pre-degree at all subsequent sessions (first year: $t=-2.2$, $p=0.028$, d=1.97; second-year: $t=-5.3$, $p<0.001$, d=1.33; third-year: $t=-5.69$, $p<0.001$, d=2.53) compared to performance pre-degree, although accuracy alone was only significantly improved at the final session ($t=-2.16$, $p=0.031$, d=0.57).

---

[2] Effect sizes could only be calculated for 2D-MR where enough participants completed the task at both of the sessions being compared.
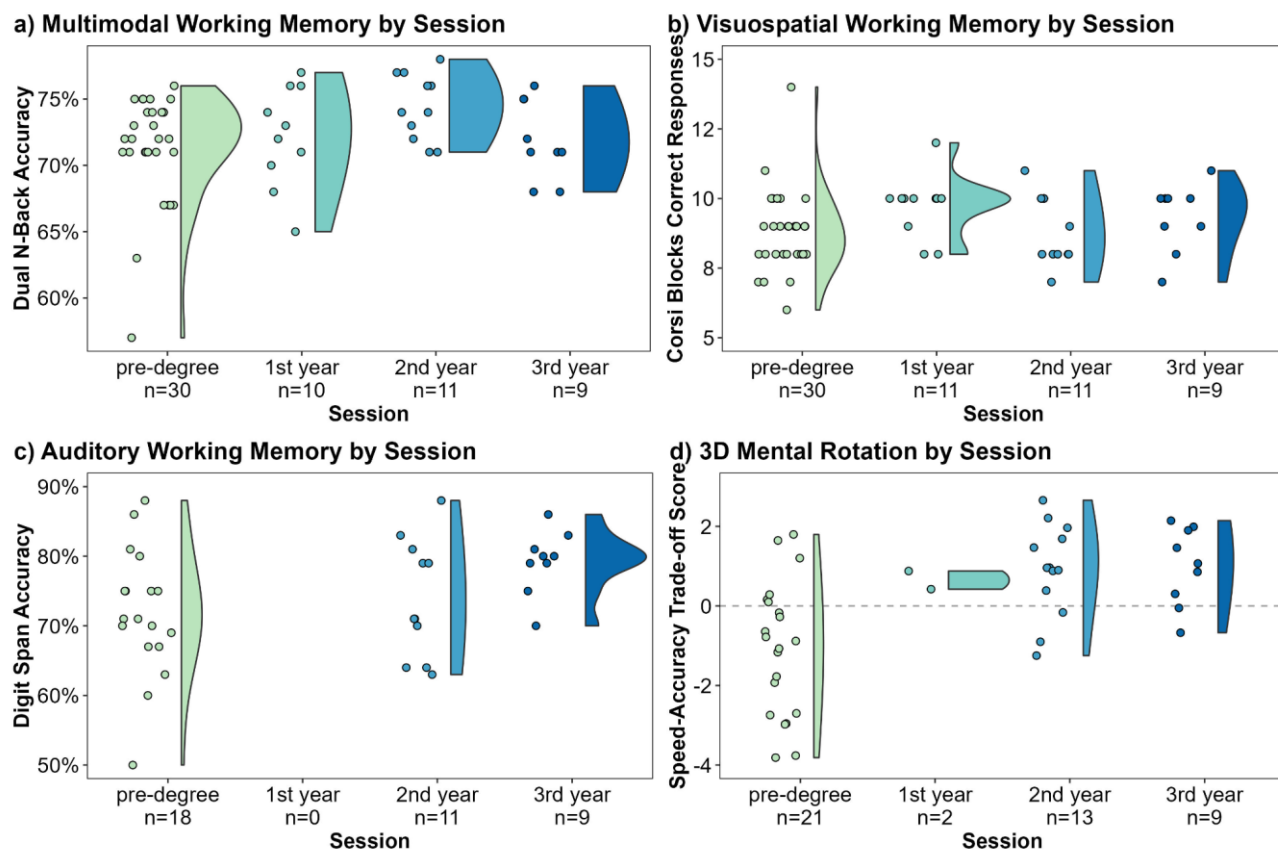
Figure 3: ChangeImprovements in SLI student performance on four cognitive assessments over time:- a) multimodal WM was significantly improved after two years of study compared to pre-degree (t=2.32, p=0.02, d=0.56); b) visuospatial WM was significantly improved after one year of study compared to pre-degree (t=2.01; p=0.044, d=0.29); c) auditory WM was significantly improved after three years of study compared to pre-degree (t=2.72, p=0.007, d=1.57); d) 3D mental rotation was significantly improved compared to pre-degree at all subsequent sessions (vs. first year: t=2.51, p=0.012, d=1.32; vs. second-year: t=4.6, p<0.001, d=0.91; vs. third-year: t=5.31, p<0.001, d=1.68).

3. *Linguistic and other assessments over time*

English reading comprehension (Kirklees) showed no effect of session when comparing pre-degree accuracy to first-year performance ($t=-1.19$, $p=0.24$) or to second-year performance ($t=1.38$, $p=0.17$). Our control measure, non-verbal reasoning (KBIT-2 Matrices), also showed no effect of session when comparing pre-degree accuracy with first-year performance ($t=-1.28$, $p=0.20$) or second-year performance ($t=0.5$, $p=0.62$). Due to time constraints on in-person testing and limited data from ~~some~~ online testing sessions, the tasks assessing phonological encoding, summarizing and impulsivity were only conducted at one session each and therefore no change-over-time analyses were conducted.

3.

4. **Predictors of BSL and SLI performance**

Our final set of analyses examines whether any of the cognitive and linguistic assessments we~~a~~re associated with (1) BSL and (2) SLI performance. Firstly, we look at whether any assessments were related to BSL measures, such as grades in BSL modules and BSL Sentence Reproduction Test

scores. Due to the high level of attrition rate across the longitudinal study, we do not attempt to fit a large mixed model with all predictors for outcome variables at the final session. Instead we report correlations and individual regression analyses, modeling BSL and SLI measures as a function of predictor assessments, with initial BSL proficiency as a covariate. Models of English comprehension also include age at initial testing session as a covariate.

1. *Working memory tasks and BSL performance*

As a reminder of our hypotheses: we did not predict that the WM assessments would have an impact on any BSL outcomes, other than visuospatial WM. There was no relationship between initial multimodal WM (Dual N-Back) and second-year BSL grades ($r^2=0.006$), nor did multimodal WM relate to third-year BSL-SRT scores ($r^2=0.002$). The~~re~~ relationship ~~was a weak association~~ between third-year multimodal WM skill and SRT scores~~, but this~~ was not ~~a~~ significant ~~relationship~~ ($r^2=0.12$, $t=0.83$, $p=0.44$). We found no relationship between initial visuospatial WM (Corsi Blocks) and second-year BSL grades ($r^2<.001$), nor with third-year BSL-SRT score~~s~~ ($r^2=0.10$). However, third-year SRT scores were significantly positively associated with visuospatial WM skill ~~performance~~ in ~~both~~ second-year ($r^2=0.609$, $t=2.73$, $p=0.041$; Fig. 4a), but the positive correlation with ~~and~~ third-year WM was not significant ($r^2=0.301$, $t=1.55$, $p=0.17$; Fig. 4b)~~, the former being a significant relationship~~. In terms of auditory WM, initial digit span scores were not~~only weakly~~ associated with first- ($r^2=0.12$) or~~and~~ second-year BSL grades ($r^2=0.14$), nor with ~~but not~~ third-year BSL-SRT scores ($r^2=0.001$).
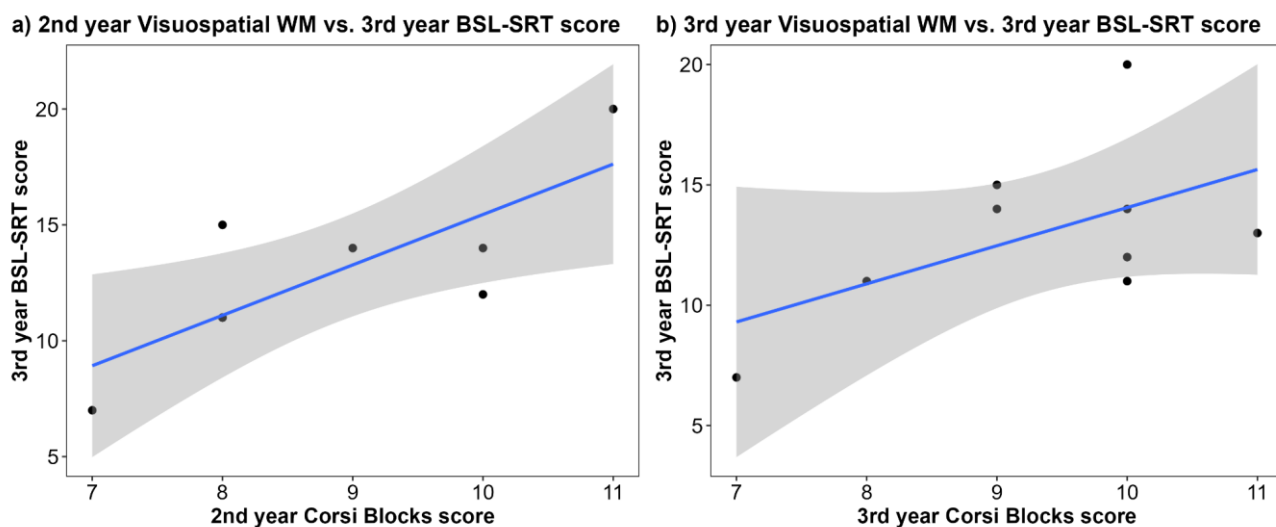
Figure 4: Correlations ~~Associations~~ between SLI student performance on working memory tasks~~visuospatial WM~~ and BSL measures~~3rd-year BSL-SRT scores~~: a) 2nd-year visuospatial WM was a significant predictor of 3rd-year BSL-SRT scores (n=7; r2=0.609, t=2.73, p=0.041); b) 3rd-year visuospatial WM was positively correlated with 3rd-year BSL-SRT scores, but this relationship was insignificant (n=9; r2=0.301, t=1.55, p=0.17).

2. *Mental rotation tasks and BSL performance*

There was no relationship between initial 2D rotation and second-year BSL grades ($r^2=0.001$). The~~re~~ relationships~~was a moderate positive correlation~~ between second-year 2D-MR scores and second-year BSL grades ($r^2=0.27$, $t=1.64$, $p=0.14$), as well as between third-year rotation scores and third-year BSL-SRT scores ($r^2=0.19$, $t=1.5$, $p=0.18$) showed mode, but these were also not significant. Likewise, for 3D-MR, there was no relationship between initial 3D rotation skill and second-year

BSL grades (r^2<.001). ~~However, i~~Initial 3D-MR scores were positively ~~correl~~associated with third-year BSL-SRT scores, but this correlation was not significantly different from zero (r^2=0.37; t=1.81, p=0.12; Fig. 5a). Second-year 3D-MR ~~was~~ and ~~positively associated with~~ second-year BSL grades, ~~a relationship that was~~ were also strongly correlated, but the effect was marginally insignificant (r^2=0.39, t=2.21, p=0.054; Fig. 5b). Third-year 3D rotation scores were ~~only weakly associated with~~ unrelated to third-year BSL-SRT scores (r^2=0.12).
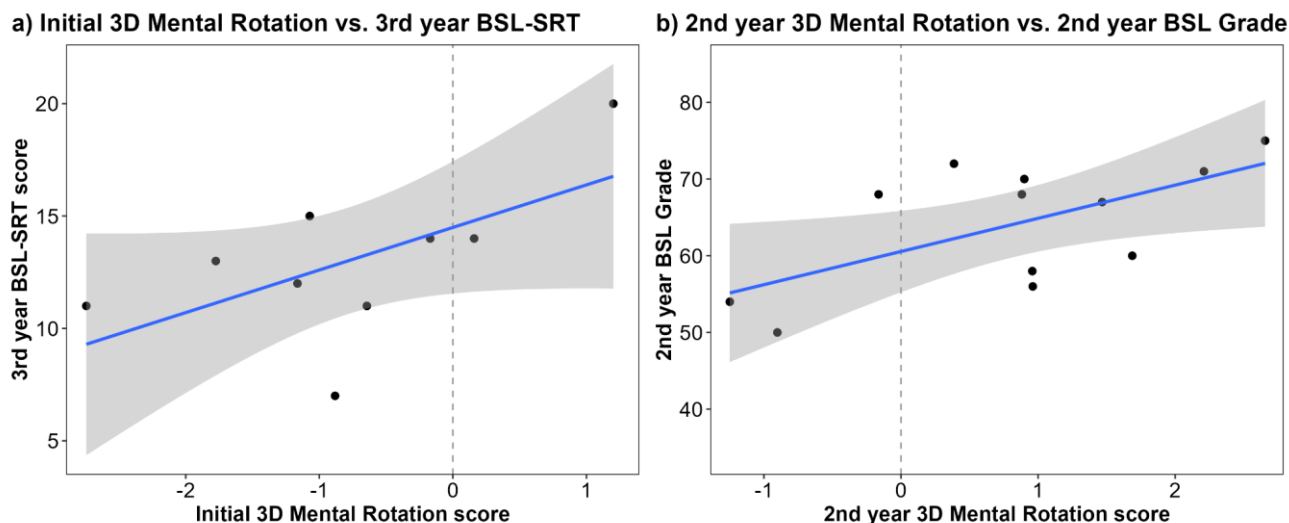
Figure 5: Correlations ~~Associations~~ between SLI student performance on mental rotation tasks and BSL measures: a) pre-degree 3D-MR skill was positively correlated with ~~and~~ 3rd-year BSL-SRT scores, but this relationship was not significant (n=9; r2=0.37; t=1.81, p=0.12); b) 2nd-year 3D-MR was positively correlated with ~~and~~ 2nd-year BSL grades, but this relationship was marginally insignificant (n=12; r2=0.39, t=2.21, p=0.054).

3. *Linguistic and other assessments and BSL performance*

There was a significant positive relationship between initial English reading comprehension (Kirklees) and first-year BSL grades (r^2=0.19, t=2.90, p=0.008; Fig. 6a), but not with subsequent BSL grades (r^2=0.02). The relationship ~~We also found a weaker positive association~~ between initial reading comprehension and third-year BSL-SRT score showed a moderate positive correlation, but this was not significant (r^2=0.237, t=1.56, p=0.17; Fig. 6b). In terms of the summarizing task, as predicted there was no impact on first- (r^2=0.002) or second-year BSL grades (r^2=-0.06), nor was there a relationship with third-year BSL-SRT score (r^2=0.01).

We found no relationship between initial phonological encoding ability (MLAT Number Learning) and second-year BSL grades (r^2=0.033), nor was there a relationship with third-year BSL-SRT score (r^2<.001). There was no association between impulsivity (BIS) and second-year BSL grades (r^2=0.009). Impulsivity had a slight~~was~~ negative~~ly associated~~ correlation with third-year BSL-SRT scores, but this relationship was not significant (r^2=-0.27, t=-1.45, p=0.19). There was no relationship between our control measure (non-verbal reasoning; KBIT-2 Matrices) and second-year BSL grades (r^2=0.001). Non-verbal reasoning also did not correlate with third-year BSL-SRT scores (r^2=0.004).
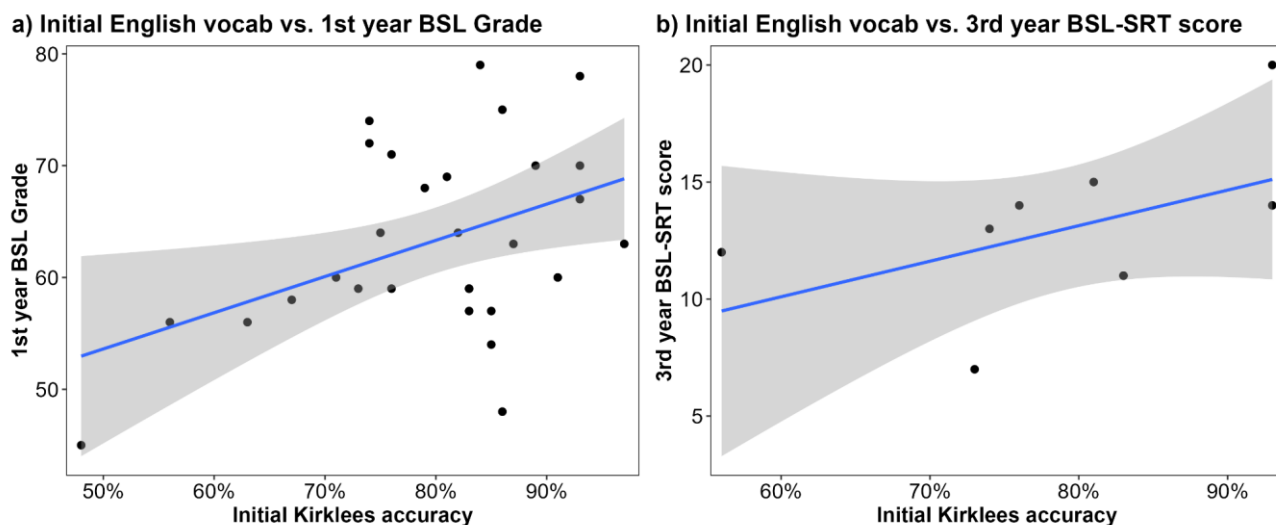
Figure 6: Correl~~Associ~~ations between initial SLI student performance on English reading comprehension and later BSL measures: a) pre-degree English comprehension was a significant predictor of ~~and~~ 1st-year BSL grades (n=29; r2=0.19, t=2.90, p=0.008); b) pre-degree English comprehension was moderately positively correlated with ~~and~~ 3rd-year BSL-SRT scores, but this relationship was not significant (n=9; r2=0.237, t=1.56, p=0.17).

Now we turn to look at the relationships ~~associations~~ between cognitive and linguistic assessments and measures of SLI performance, such as SLI module grades and the two final consecutive SLI tasks.

4.   *Working memory tasks and SLI performance*

There was ~~only a slight positive association~~no significant relationship between initial multimodal WM scores and second-year interpreting grades (r^2=0.11), nor with performance on the third-year English-to-BSL consecutive interpreting task (r^2=0.12). This correlation was slightly stronger for third-year multimodal WM scores, but again not significantly so (r^2=0.16, t=0.7, p=0.51). Initial multimodal WM did not affect scores on the consecutive interpreting task from BSL to English (r^2=0.06).

In terms of visuospatial WM, there was no link between initial performance on the Corsi Blocks task and second-year interpreting grades (r^2=0.008), nor did a relationship emerge with visuospatial WM at later testing sessions. There was a stronger yet insignificant correl~~weak positive associ~~ation between initial visuospatial WM and scores on the third-year English-to-BSL consecutive interpreting task (r^2=0.29, t=1.27, p=0.25), and for second-year visuospatial WM ~~however~~ this relationship was only marginally ~~just missed~~ insignificant~~ce for second-year visuospatial WM~~ (r^2=0.57, t=2.48, p=0.056; Fig. 7a). Second-year visuospatial WM was also positively correl~~associ~~ated with performance in the third-year consecutive interpreting task from BSL to English, but again the correlation was insignificant (r^2=0.46, t=1.82, p=0.14).

There was a promising positive correlation between initial auditory WM (Digit Span) scores and SLI grades in second-year which was marginally insignificant (r^2=0.24 t=1.96, p=0.076; Fig. 7b). Auditory WM at the final testing session was not related to~~weakly associated with~~ the final English-to-BSL consecutive interpreting task (r^2=0.12), nor the BSL-to-English task (r2=0.03).
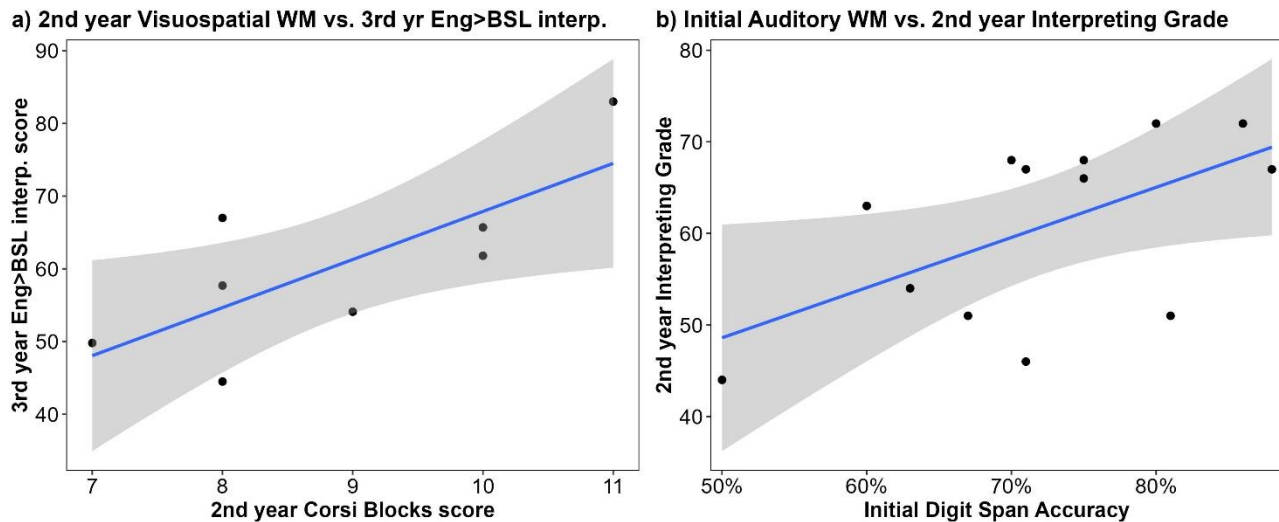
Figure 7: ~~Correl~~Associations between SLI student performance on WM tasks and SLI measures: a) 2nd-year visuospatial WM skill ~~comprehension~~ was positively correlated with ~~and~~ performance on the 3rd-year English-to-BSL interpreting task~~scores~~, however this relationship was marginally insignificant (n=8; r2=0.57, t=2.48, p=0.056); b) pre-degree ~~initial~~ auditory WM was correlated with ~~and~~ 2nd-year SLI grades, but this relationship was marginally insignificant (n=13; r2=0.24 t=1.96, p=0.076).

5.   *Mental rotation tasks and SLI performance*

There was no relationship between initial 2D-MR skill and second-year SLI grades (r^2=0.038). We found stronger correl~~positive associ~~ations between second-year SLI grades and 2D-MR skill in both second year (r^2=0.24, t=1.64, p=0.14) and in third-year (r^2=0.39, t=2.91, p=0.027),~~–~~ the latter was statistically significant. F~~Likewise, f~~or the English-to-BSL consecutive interpreting task, there were small positive correlations with second-year (r^2=0.2) and third-year 2D-MR (r^2=0.18), but these were not significant. However, for the third-year BSL-to-English interpreting task, the relationship with 2D-MR was stronger: second-year 2D-MR skill was a significant predictor (r^2=0.6, t=2.95, p=0.042; Fig. 8a) while ~~and there was a positive association with~~ third-year 2D-MR skill was a marginally insignificant predictor (r^2=0.31, t=2.11, p=0.088; Fig. 8b).

For 3D-MR, there was a moderate yet insignificant positive correl~~slight positive associ~~ation between initial rotation skill and second-year interpreting grades (r^2=0.24)~~-~~. This correlation ~~which~~ became stronger over time (vs. second-year 3D rotation: r^2=0.31; vs. third-year 3D rotation: r^2=0.54, t=2.85, p=0.029), the latter relationship being significant. Initial 3D-MR ~~was~~ also had a moderate but insignificant positive correlation ~~ly associated~~ with third-year English-to-BSL consecutive interpreting performance (r^2=0.27, t=1.32, p=0.24; Fig. 8c), although there was no correlation with ~~association for~~ later 3D-MR scores. Initial 3D-MR skill was also positively correlated with third-year consecutive interpreting from BSL to English and this relationship was marginally insignificant (r^2=0.48, t=2.004, p=0.101; Fig. 8d). However, this correlation ~~association~~ got weaker over time (vs. second-year 3D rotation: r^2=0.29; vs. third-year 3D rotation: r^2=0.16).
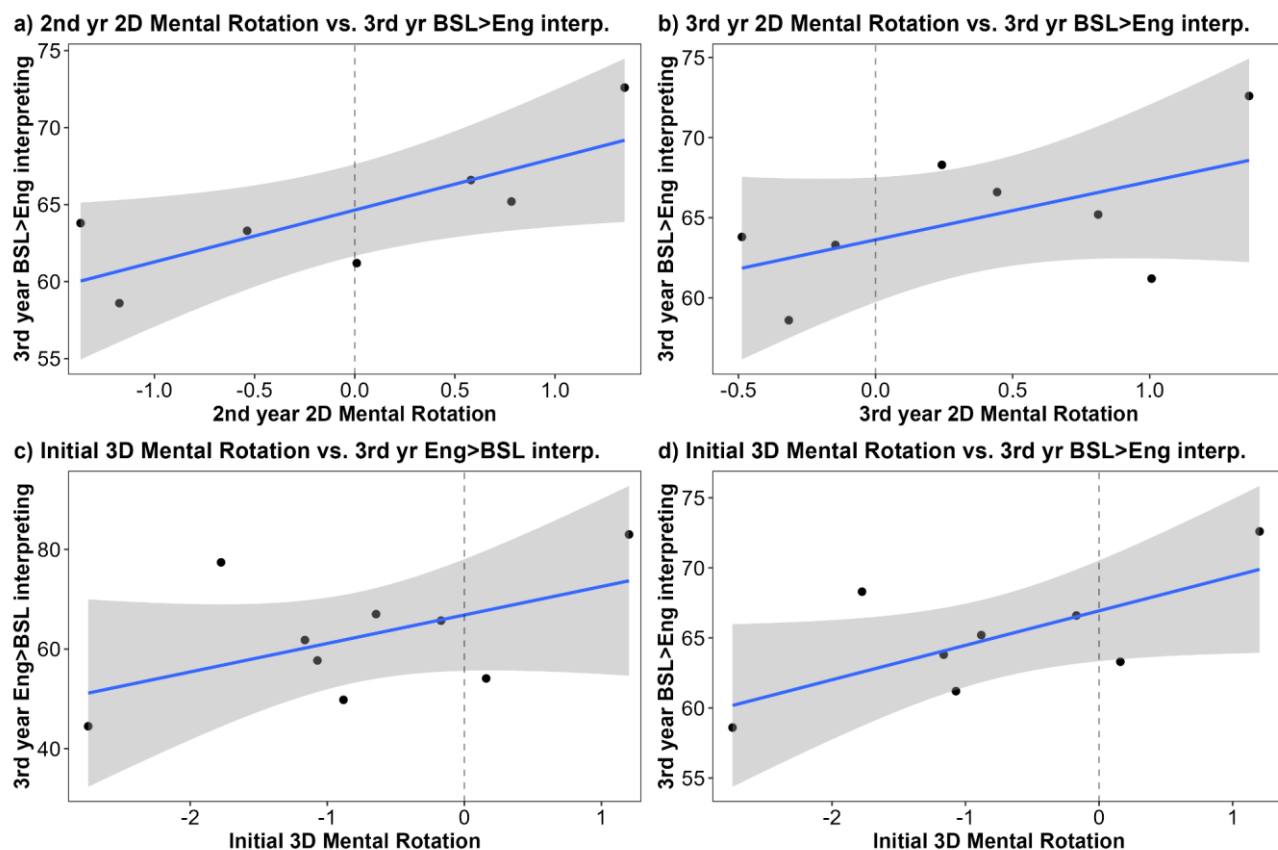
Figure 8: CorrelAssociations between SLI student performance on mental rotation tasks and SLI measures: a) 2nd-year 2D-MR was a significant predictor of and scores on the 3rd-year BSL-to-English interpreting task (n=7, r2=0.6, t=2.95, p=0.042); b) 3rd-year 2D-MR was correlated withand scores on the 3rd-year BSL-to-English interpreting task, but this relationship was marginally insignificant (n=8, r2=0.31, t=2.11, p=0.088); c) pre-degree 3D-MR was positively correlated with scores on thevs 3rd-year English-to-BSL interpreting task, but this relationship was not significant (n=9, r2=0.27, t=1.32, p=0.24); d) pre-degree 3D-MR was correlated with scores on theand 3rd-year BSL-to-English interpreting task, but this relationship was marginally insignificant (n=8, r2=0.48, t=2.004, p=0.101).

6. *Linguistic and other assessments and SLI performance*

There was no relationship between English sentence reading (Kirklees) at any time point and SLI grades, or with either of the consecutive SLI tasks. There was also no link between the pre-degree summarizing task and second-year SLI grades (r^2=0.002), nor with either interpreting task.
We also found no significant relationship with non-verbal reasoning (KBIT-2 Matrices) and any SLI measures., apart from one weak negative association between second-year non-verbal reasoning and second-year SLI grades (r^2= 0.12). There was no relationship between impulsivity (BIS) and SLI grades. There was a slight negative correlation between impulsivity and English-to-BSL consecutive interpreting, as well as between impulsivity and the consecutive interpreting task from BSL to English (r^2=-0.27), however these were not significant.

7. *Correlations among measures of BSL and SLI performance*

Second-year SLI grades were significantly correlated with both first-year BSL grades (r^2=0.68, t=7.21, p<.001) and second-year BSL grades (r^2=0.69, t=6.79, p<.001). The initial copy-sign task

was a marginally insignificant predictor of first-year BSL grades (r^{2}=0.46, t=2.25, p=0.054), but not second-year BSL grades (r^{2}=0.21, t=0.84, p=0.43). Interestingly, the initial copy-sign task had a stronger correlation with SLI grades in second year (r^{2}=0.36, t=2.08, p=0.076) than with BSL grades, but again this was marginally insignificant. Performance on the third-year BSL-SRT task was significantly correlated with first-year BSL grades (r^{2}=0.66, t=3.37, p=0.015) and strongly correlated with second-year BSL grades (r^{2}=0.53, t=2.25, p=0.074), which was a marginally insignificant predictor. The two consecutive SLI tasks (BSL-to-English and English-to-BSL) were also significantly correlated with each other (r^{2}=0.79, t=4.14, p=.009).

5. Discussion

1. Summary of key findings

In this exploratory study, we saw several significant relationships between cognitive and linguistic skills and SLI degree program outcomes. Firstly, we found that 2D-MR skill and initial self-rated BSL proficiency were significant predictors of selection for the SLI degree, whereas visuospatial WM predicted continuation on the course. Next, we examined the impact of the SLI degree program on cognitive and linguistic skills over time, by first collecting baseline measures of these skills and repeating testing at three further sessions. We found improvements in multimodal WM by the end of the second year of the degree, in visuospatial WM by the end of the first year, and in auditory WM by the final testing session in third year. While multimodal and visuospatial WM did not remain at these improved levels at subsequent sessions, both 2D- and 3D-MR skills were improved by the first year of the degree and consistently remained at these higher levels throughout our longitudinal study. In terms of cognitive and linguistic predictors of later BSL and SLI task performance, there were several significant and marginally insignificant results of note, despite our small sample size. First-year BSL grades were significantly predicted by pre-degree English comprehension; second-year BSL grades were strongly correlated with second-year 3D-MR (marginally insignificant); and third-year BSL-SRT scores were significantly predicted by second-year visuospatial WM. Second-year visuospatial WM was also strongly correlated with scores on the English-to-BSL interpreting task; and pre-degree auditory WM was strongly correlated with second-year SLI grades (both marginally insignificant). In terms of the 3rd-year BSL-to-English interpreting task, 3D-MR was a significant or marginally insignificant predictor at three different testing sessions (pre-degree, second-year and third-year). We now turn to consider some issues with data collection and attrition, and then move on to a detailed discussion of each of our three research questions and their related analyses in turn.

2. Issues with data collection and attrition

The various effects of the COVID-19 pandemic had considerable impact on our exploratory study, mostly in terms of student withdrawals from the degree program and the difficulties of re-testing students online at later sessions, making it more of a preliminary investigation. In particular, the earliest cohort (HW1), for whom the practical placement year fell during the first year of the pandemic, was affected heavily by withdrawals. Only a minority remained in the SLI degree program by final year (just 3 of the original 12, i.e. 25%; 1 student joined the year group below).

Some of those who withdrew continued on a BSL-only program while others switched to unrelated degree courses or left university altogether. Anecdotal evidence from program instructors and students themselves suggest that the main reason for the high withdrawal rate in this cohort was indeed the various impacts of the pandemic, which our statistical models hadve no way to account for without further demographic auxiliary variables. In particular, having to do remote interpreting placements seemed to be isolating and discouraging for students, and not the immersive experience it might have otherwise been. As a result, it is likely that capable and otherwise potentially successful students (particularly in this cohort) ended up withdrawing from their SLI program. However, it should be noted that the majority of students in the other two cohorts remained on their programs. We believe our data still offer valuable insights about SLI aptitude given the range of assessments tested across multiple sessions, given the present lack of longitudinal work with baseline measures in the literature.

3. Predictors of selection and remaining on SLI degree program

In our initial analyses, we saw some limited evidence of initial pre-degree assessments predicting whether students were either selected for, or continued on with, their SLI degree program. For the selection analysis, the best predictors were initial self-rated BSL proficiency and the 2D-MR task. It is fairly unsurprising that those students with some initial proficiency were more likely to be selected ahead of those with no knowledge of BSL and deaf culture. This result also highlights that it is difficult to truly test SLI aptitude in university interpreting students from a complete baseline of no BSL exposure at all, since many are unlikely to commit to a degree without some prior signing experience. Although timing and funding constraints meant we were not able to code enough of the copy-sign data to include it as a predictor in the selection model, the data we do have suggests that scores on the task correlated significantly with applicants' BSL proficiency self-ratings. While our copy-sign task featured real BSL signs, many of our participants would not yet have been familiar with all the signs used. In this sense, it is similar to the nonsense sign repetition tasks used by López Gómez et al. (2007) and Stone (2017), which also involved phonological encoding and perceptuo-motor skills and were found to be good predictors of later success in their cohorts. The 2D-MR task was also a good predictor of selection, and this is an interesting result given other strong correlations between initial 3D-MR performance and later outcome measures. However, our selection analysis was conducted on just two cohorts (HW0 and HW1) and on only five assessments, mostly due to time constraints at the interview stage (latter cohorts were tested during the first week of term, which meant we could test more assessments in a less stressful environment for participants). Future studies could aim to measure other skills assessments at the initial interview -stage which could not be included here due to practical constraints. In particular, 3D-MR and auditory WM would be good candidates for further skills to examine, since we found evidence that pre-degree performance in both of these domains in our other cohorts was strongly correlatedwas positively associated with third-year SLI performance and SLI grades respectively. Lastly, iIt should be reiterated noted that SLI program instructors did not have access to data from the initial assessment session when deciding whether to accept applicants to their degree programsstudents.

In the continuation analysis, there was a significant relationship n association between pre-degree the visuospatial WM and remaining on the SLI degree, whereby higher visuospatial WM scores were associated with continuation. One interpretation of this result is thatWhile this is not a very strong effect, it could be that the cognitive demands of signing and interpreting in the new visuospatial modality are possibly a factor that (consciously or unconsciously) makes students more likely to withdraw from the program (see e.g. Grbić, 2009). Due to the high withdrawal rate in one specific cohort, it was not possible to include a mental rotation task in the selection this model, since different

cohorts did a different rotation task at the first session. However, it would be interesting to see whether other cognitive tasks involving visuospatial cognition are also associated with program drop-out, such as the copy-sign task. In terms of impulsivity, we did not replicate the result seen by Stone (2017) whereby of the BIS predicteding continuation on the SLI degree vs. switching to a Deaf Studies course. Due to testing session time constraints, wWe only introduced the BIS from session 3 onwards, where all testing was already online due to the pandemic, but ideally we would have also taken a baseline measure of risk-taking. It is quite likely that our participants' usual risk aversion strategies and levels of impulsiveness were inhibited or changed by the ongoing pandemic. The potential for impulsiveness to impact on learning behaviors (e.g. having the confidence to attend deaf social events to practice BSL, etc.) was greatly restricted due to pandemic-related safety measures. In future larger studies that follow students all the way through to graduation and beyond, it would be interesting to see whether specific cognitive and linguistic skills are associated with continuation or drop-out at particular points in time, for example, during placements, or once students start interpreting modules.

4. Changes in cognitive and linguistic skills during SLI program

In terms of changes in the cognitive and linguistic skills over the course of the degree, there was evidence of improvement in several domains. All three WM tasks (multimodal, visuospatial and auditory) showed some limited signs of improvement across testing sessions, although performance did not always remain at thise higher level. Visuospatial WM (Corsi Blocks) and multimodal WM (Dual N-Back) were improved by timepoints two and three respectively, but these changes did not hold to the final session, nor was there further improvement. These results are harder to interpret, but it could be that these regressions are related to reduced opportunities to practice BSL and SLI skills at certain points in the degree program. This is likely to have been the case at later sessions due to the effects of the COVID-19 pandemic. for multimodal and visuospatial WMs at the final testing session attained compared to the initial testing session. Alternatively, tThese regressions is may be explained by due to the varying demands of the SLI degree program tapping into different aspects of WM more frequently at different time-points during the course. Auditory WM attention (Digit Span) was only at a higher level at the final testing session, midway through the placement year, suggesting there may be potential for further improvement beyond our study, yet within the SLI degree. Further testing through to degree completion could be instructive here, as most of our predictions regarding WM weare related to SLI, and not BSL learning. In the first two years of the SLI program, the focus is mostly on acquiring BSL before beginning to interpret. Since the students do much more SLI in their placement year and in final year, we would expect greater potential for this to feed back into WM skills only towards the end of the degree, at later time points and beyond during professional practice. Visuospatial WM (Corsi Blocks) and multimodal WM (Dual N-Back) were improved by timepoints two and three respectively, but these changes did not hold to the final session, nor was there further improvement. These results are harder to interpret, but it could be that these regressions are related to reduced opportunities to practice BSL and SLI skills at certain points in the degree program. This is likely to have been the case at later sessions due to the effects of the COVID-19 pandemic.

As hypothesized, MR ability in SLI students improved over time, which fits with a wide body of literature of improved visuospatial cognition through sign language use (e.g. McKee, 1987; Talbot and Haude, 1993; Emmorey, 1998; Emmorey et al. 1998; Keehner and Gathercole, 2007; Kubicek and Quandt, 2021REFs). In particular, the improvements on the 3D-MR task between pre-degree and the final sessions three and four hadwere the strongest two-effect size across the entire longitudinal study. Results in the 2D-MR task were less clear-cut, with improved performance by the second session followed by a regression at later sessions in terms of speed-accuracy trade-off alone, but

faster reaction times by both the second-year and third-year session compared to the initial testing session. However, as suggested by Watkins, Abdlkarim and Thompson (2018), it is rotation around the vertical axis, as assessed by the 3D block rotation task, which should most closely resemble the plane in which sign language input must be rotated during language processing. The 2D-MR task used here involved simple circular rotation which does not map as clearly onto any visuospatial transformations during BSL use. In this sense, it is unsurprising that greater improvements over time were seen in the 3D-MR task and not the 2D-MR task, in tandem with L2 BSL learning during the degree program. The stronger improvement in the 3D-MR task involving rotation around the vertical axis, which is required during sign comprehension to understand someone signing from a different viewing angle, for example, ~~–~~provides some evidence that it is experience with rotation in this plane~~manner~~ through BSL practice~~experience~~ which is driving the improvement in 3D-MR, and not other types of rotation (e.g. signs that move from a palm-up to palm-down orientation).

For English reading comprehension, we did not see the same improvement over time as in Stone (2017), where Kirklees scores were significantly better by the end of third year versus the start of first year. Due to time constraints during testing, we could not repeat this task at our final third-year testing session. We assessed ~~included~~ it for the final time at the end of second year, which may be why we could not replicate this result. However, the Kirklees scores of our SLI students at the end of second year were just below the level of Stone's third-year students, who in turn scored significantly lower than the working BSL-English interpreters who were also tested. Given the lack of improvement in the present preliminary study and the conclusion of Stone (2017), it may be that experience working as an interpreter improves English skills (perhaps due to exposure to a wide range of vocabulary during interpreting) above and beyond any improvement seen during a university degree program. Age is generally a reliable predictor of vocabulary size (e.g. Ben-David et al., 2015), with better vocabulary scores as age increases. While it was not a significant covariate in the present study, our sample of university students unsurprisingly did not have an even spread of our ages (two-thirds of our participants were aged 23 or under at course begin).

5.   Predictors of BSL and SLI performance

There were several strong correlations~~positive associations~~ between the cognitive/linguistic skills assessed and measures of BSL and SLI performance, some of which were statistically significant.

Starting with BSL outcomes, first-year BSL grades were significantly predicted by English reading comprehension, a correlation that was particularly strong for those with lower English and lower BSL scores (Fig. 3a). This suggests that weaker English skills at the start of an SLI program may initially be a hindrance for language learning in another modality. However, SLI students who perform poorly in English initially seem to catch up by later testing sessions, by which point English comprehension was no longer predictive of BSL performance. Second-year BSL grades were correlated positively with 3D-MR skill in the same year, which fits with studies showing that MR improves gradually in line with sign language learning, and does not just improve once sign language fluency is reached (Kubicek and Quandt, 2021). Interestingly, however, there was also a promisingly strong correlation, albeit not significant, ~~relationship~~ between initial 3D-MR and third-year BSL-SRT scores (Fig. 5a). This suggests that at least some of the success in BSL performance may~~can~~ be predictable~~ed~~ from rotation scores at the outset of the degree, with an advantage for those who already start the SLI program with better MR skill (see e.g. Kartheiser et al., 2022, who conclude that adult L2 signers can apply pre-existing non-linguistic spatial skills to the sign language they are learning). Regardless of initial MR skills, however, the findings suggest that students who do not improve at MR as they advance in the program, do not perform as well as those who do. This may be

true not only for BSL comprehension, but production as well. Second-year visuospatial WM was also a significant predictor of SRT scores at the final testing point (Fig. 4a). Both mental rotation and visuospatial WM are likely implicated in this task, where increasingly complex grammatical constructions in BSL must be reproduced. Furthermore, tThe BSL-SRT, like the copy-sign task and nonsense sign repetition tasks which were predictive of sign language performance in previous studies (López Gómez et al., 2007; Stone, 2017), involves phonological encoding and perceptuo-motor skills (see also Martinez and Singleton, 2018, who found that visuospatial short-term memory was predictive of sign learning, concluding that perceptuo-motor processes play a big role in individual sign learning). However, we did not replicate the result seen by Stone (2017), whereby initial performance on the phonological encoding task (MLAT Number Learning) was predictive of early BSL module grades.

In terms of SLI performance, we found that initial auditory WM was a promising indicator of second-year SLI grades (Fig. 7b), suggesting it may be a useful assessment to conduct at student intake. While a digit span task was also used by Stone (2017) as a control measure of general cognitive ability, we employ it here as a measure of WM. Using an auditory version of the task has allowed us to explore effects of different modalities in WM that previous SLI aptitude studies have not. There was also a significant correlation between SLI grades and 3D-MR in third year (r2=0.54, t=2.85, p=0.029). The third-year BSL-to-English interpreting task was also significantly correlated with second-year 2D-MR (Fig. 8a), and the correlation with pre-degree 3D-MR was also very high (r^2=0.48, t=2.004, p=0.101; Fig. 8d). The relationships here between MR skill and SLI outcomes were beyond what we had predicted: we had hypothesized that rotation would only be implicated directly in BSL tasks. However, it is plausible that rotation skills required during sign language comprehension and production (e.g. for syntactic and topographic uses of signing space) also come into play during measures of SLI itself. Furthermore, interpreting interactions also involve competent navigation of the spatial relations between interlocutors in physical space, where the viewing angle may be a factor. For example, in group scenarios, MR may be invoked to comprehend signing viewed from a non-frontal angle (Watkins, Abdlkarim and Thompson, 2018). Lastly, performance on the SLI task from English to BSL was significantly correlated with second-year visuospatial WM. It may be that WM in different modalities are implicated in different ways during SLI when working in different directions:, for example, visuospatial WM could be particularly important while planning and executing sign language production, as required by this task. It is possible that visuospatial WM is required constantly for production in the visuospatial modality during spoken to signed interpreting, whereas when interpreting in the opposite direction, visuospatial WM is only engaged to attend to the spatial relations in signing space, and is not required for the processing of specific signs.

Overall, the predictor assessments with the best and most consistent relationships to BSL and SLI performance were the MR tasks, in particular the 3D-MR task, and visuospatial WM. MR tasks were significant or marginally insignificant predictors of all the BSL/SLI outcome measures in the present study: associated with BSL grades, BSL sentence reproduction and, SLI grades and interpreting tasks in both directions, suggesting that MR this is an essential skill for SLI educators to pay attention to. Second-year visuospatial WM was a significant predictor of associated with final BSL sentence repetition and strongly correlated with English-to-BSL interpreting performance, which we interpret as WM in the visuospatial modality being necessary for the BSL chunking, planning and production required during both tasks. Most of the relationships between the cognitive/linguistic assessments and BSL/SLI performance emerged during re-testing at later sessions, although there were some notable exceptions which may point towards skills that have predictive value at the outset of a SLI degree program. Initial 3D-MR was strongly correlassociated with the same two final-year outcomes as second-year visuospatial WM (SRT and English-to-BSL interpreting), highlighting the key role of

cognition in the visuospatial modality for tasks involving planning and executing sign language production. We also saw that initial English vocabulary was important early on in L2 BSL learning, as well as and a promising relationship between initial auditory WM and second-year SLI grades, suggesting both of these tasks are worth assessing at intake. Although our data are only exploratory, we have some initial evidence which supports Marks' (2014) assertion that "[a] case can be made that there are cognitive skills that need to be present upon entry into [SLI] programs and others that can and need to be taught".

6. Implications and future directions

This exploratory study highlights multiple domains worth further attention for SLI educators and researchers. To our knowledge, this is the first longitudinal study of SLI students to take baseline cognitive and linguistic measures before the start of the training program and to relate them to performance on sign language and SLI tasks. We saw evidence that MR skill is implicated in not just sign language outcome measures learning but also in SLI performance, as well as links between visuospatial working memory and sign language production in particular. We see some evidence that good English skills are initially important for early BSL learning and a possible role of initial auditory WM in SLI, which should be investigated by future studies.

Since this is still a preliminary study, however, we do not advocate excluding SLI program applicants at intake based on performance in any of the assessments conducted here. As Robinson (2022) highlights, there are issues with (mostly hearing) SLI educators further restricting the pool of potential SLI students at intake, particularly when interpreter demand already exceeds supply. NeverthelessHowever, initial SLI aptitude testing could also help to instead highlight other related careers that do not involve SLI, which may be more appropriate for some candidates, before struggling with a lengthy degree program. Future studies testing (visuospatial) cognitive skills at SLI course intake should also consider the additional stresses of assessments at interview and issues such as stereotype threat, i.e. where performance is affected by awareness of a negative stereotype about one's social group. For example, Moè and Pazzaglia (2006) found stereotype threat to be relevant to MR and gender; a consideration given that most (BSL) interpreters are women (Napier et al., 2021), as were most of the SLI students in this study. Regardless of initial baseline performance, oOur preliminary results should make SLI training programmesprograms aware of skills that would be worth tracking in their students, as well as the possibility to offer additional targeted practice of skills that are likely to improvekey skills throughout SLI training. Since it is now fairly well- established that MR skill improves in tandem with sign language learning, it is also plausible that we could speed up this process by explicitly including rotation practice on assignments, tasks or games, whether in a sign language context or not. This could be implicit practice with comprehension of sign language from different viewing angles in group or dialogue situations, where MR is likely implicated (see e.g., Watkins and Thompson, 2019). Alternatively, this could be explicit training using practicing gamified versions of MR tasks, like the ones used here, to try to boost performance, which should in turn feed back into signing and interpreting performance (see e.g. Passig and Eden, 2001). Furthermore, SLI educators can use these results to diversify the teaching and learning experience to better support students' development in these areas.

The title of Stone (2017), "the trials and tribulations of a longitudinal study", bears repeating, because aptitude studies of SLI students are complex endeavors. No single longitudinal study can address all the design, methodology and data analysis issues. As Stone points out, we must be careful when interpreting the results of SLI aptitude studies, because most of the effect sizes are modestsmall, due to both the high levels of attrition and the small initial sample sizes. Even without

the impacts of the pandemic and the high drop-out rate in SLI degree programs, studies on the SLI student population in the UK are always likely to be small in scale, due to the population to sample from not being very large (around 60 new degree students per year in total). Despite the small *n* in this exploratory e present study, we hope that our sample is somewhat representative, having tested most of the students across two year--groups from two of the three UK universities where SLI programs are taught in degree form. However, issues of sample size and statistical power will likely continue unless an large scale, international, multi-center study is organized. In the present study, we had hoped to perform the first institution-level comparison in the UK, but participant recruitment and retention proved particularly difficult. Future A large-scale longitudinal studies y could also be facilitated by using online testing for many of the assessments, which we discovered worked well for students who had the means to take part remotely at later sessions. The tasks hosted on Pavlovia and GoReact were particularly successful. Online aptitude testing could facilitate access to a larger population of SLI students, hHowever, factors such as equity of access to technology and space to participate online should be considered carefully. Given the extra adaptations to SLI and sign language teaching programs which were required due to the pandemic (see e.g. Katz, 2021; Hornstra, 2021) and the increasing use of online teaching methods in L2 sign language learning, even pre-pandemic (Ackerman et al., 2018), we hope the present preliminary study demonstrates that online SLI aptitude testing is also feasible.

Further improvements to a future larger study could be gained by testing control groups of students on sign language-only degrees without the SLI components, and/or students on other degree programs, for example a spoken language interpreting course with a comparable 'placement' year abroad. This would allow us to tease apart the respective effects of learning BSL vs. learning BSL *and* SLI within a degree program context, as well as any potential effects of simply completing a university degree or interpreting degree. Furthermore, more explicit attempts could be made to retain participation of students who decide to withdraw from SLI programs. While we did not attempt to re-test students who had left their SLI course at later time points, continuing to include them as participants in the study would allow researchers to make stronger inferences about aptitude and the factors that can be attributed to changes in cognitive and linguistic skills over time. Since Stone (2017) also discovered several differences in skills between his student cohort and his group of experienced interpreters, another important avenue for future research weould be to follow SLI program graduates as they transition to the workplace, in order to pinpoint how long it takes graduates to perform at the level of experienced interpreters. This could help highlight gaps in SLI training curricula.

7. Conclusion

Overall, our exploratory study has revealed various some new insights about cognitive and linguistic aptitude for L2 sign acquisition and SLI. Crucially, we have tested a range of assessments before the beginning of an SLI training program and followed both their development over time and their impact on signing and interpreting outcomes. A numberSome of our preliminary results are consistent with previous findings which suggesting the importance of both phonological encoding and visuospatial WM in SLI student success. We also tested new domains such as mental rotation which, to our knowledge, has not been tracked in SLI students before. In particular, 3D mental rotation showed the biggest improvement over time and was strongly correlated associated with a range of BSL and SLI outcome measures, as well as there being some indication that pre-degree skill in this domain may be is associated with later signing and interpreting SLI performance. We also argue that visuospatial WM and 3D-MR are particularly important for tasks involving sign language production, with implications of modality for broader theories of cognition and language aptitude.

These preliminary results will hopefully inform SLI educators about relevant skills to identify and support during training programs, as well as provide a basis for longer-term studies of SLI aptitude through to professional proficiency.

Table 1: Demographics of the cohorts of SLI students who took part in the longitudinal study.

Figure 1: Example stills from the working memory tasks used in the study: a) Dual n-*back task (multimodal WM); b) Corsi blocks task (visuospatial WM); c) Digit span task (auditory WM).*

Figure 2: Example stills from the two mental rotation tasks used in the study: a) simple circular 2D shape rotation task showing an example 'same' pair rotated 90°; b) 3D-rendered block rotation task around vertical axis only, showing an example 'mirror' pair rotated 150°.

Table 2: All skills assessed in the longitudinal study and the test/re-test procedure by cohort.

Figure 3: Improvements in SLI student performance on four cognitive assessments over time. a) multimodal working memory; b) visuospatial WM; c) auditory WM; d) 3D mental rotationFigure 3: Changes in SLI student performance on four cognitive assessments over time: a) multimodal WM was significantly improved after two years of study compared to pre-degree (t=2.32, p=0.02, d=0.56); b) visuospatial WM was significantly improved after one year of study compared to pre-degree (t=2.01; p=0.044, d=0.29); c) auditory WM was significantly improved after three years of study compared to pre-degree (t=2.72, p=0.007, d=1.57); d) 3D mental rotation was significantly improved compared to pre-degree at all subsequent sessions (vs. first year: t=2.51, p=0.012, d=1.32; vs. second-year: t=4.6, p<0.001, d=0.91; vs. third-year: t=5.31, p<0.001, d=1.68).

Figure 4: Associations between SLI student performance on visuospatial WM and 3rd-year BSL-SRT scores: a) 2nd-year visuospatial WM; b) 3rd-year visuospatial WMFigure 4: Correlations between SLI student performance on working memory tasks and BSL measures: a) 2nd-year visuospatial WM was a significant predictor of 3rd-year BSL-SRT scores (n=7; r2=0.609, t=2.73, p=0.041); b) 3rd-year visuospatial WM was positively correlated with 3rd-year BSL-SRT scores, but this relationship was insignificant (n=9; r2=0.301, t=1.55, p=0.17).

Figure 5: Associations between SLI student performance on mental rotation tasks and BSL measures: a) pre-degree 3D MR and 3rd-year BSL-SRT scores; b) 2nd-year 3D MR and 2nd-year BSL gradesFigure 5: Correlations between SLI student performance on mental rotation tasks and BSL measures: a) pre-degree 3D-MR skill was positively correlated with 3rd-year BSL-SRT scores, but this relationship was not significant (n=9; r2=0.37; t=1.81, p=0.12); b) 2nd-year 3D-MR was positively correlated with 2nd-year BSL grades, but this relationship was marginally insignificant (n=12; r2=0.39, t=2.21, p=0.054).

Figure 6: Associations between initial SLI student performance on English reading comprehension and later BSL measures: a) pre-degree English comprehension and 1st-year BSL grades; b) pre-degree English comprehension and 3rd-year BSL-SRT scoreFigure 6: Correlations between initial SLI student performance on English reading comprehension and later BSL measures: a) pre-degree English comprehension was a significant predictor of 1st-year BSL grades (n=29; r2=0.19, t=2.90, p=0.008); b) pre-degree English comprehension was moderately positively correlated with 3rd-year BSL-SRT scores, but this relationship was not significant (n=9; r2=0.237, t=1.56, p=0.17).

Figure 7: Associations between SLI student performance on WM tasks and SLI measures: a) 2nd-year visuospatial WM comprehension and 3rd-year English-to-BSL interpreting scores; b) initial

auditory WM and 2nd-year SLI grades Figure 7: Correlations between SLI student performance on WM tasks and SLI measures: a) 2nd-year visuospatial WM skill was positively correlated with performance on the 3rd-year English-to-BSL interpreting task, however this relationship was marginally insignificant (n=8; r2=0.57, t=2.48, p=0.056); b) pre-degree auditory WM was correlated with 2nd-year SLI grades but this relationship was marginally insignificant (n=13; r2=0.24 t=1.96, p=0.076).

Figure 8: Associations between SLI student performance on mental rotation tasks and SLI measures: a) 2nd-year 2D-MR and 3rd-year BSL-to-English interpreting; b) 3rd-year 2D-MR and 3rd-year BSL-to-English interpreting; c) pre-degree 3D-MR vs 3rd-year English-to-BSL interpreting; d) pre-degree 3D-MR and 3rd-year BSL-to-English interpreting Figure 8: Correlations between SLI student performance on mental rotation tasks and SLI measures: a) 2nd-year 2D-MR was a significant predictor of scores on the 3rd-year BSL-to-English interpreting task (n=7, r2=0.6, t=2.95, p=0.042); b) 3rd-year 2D-MR was correlated with scores on the 3rd-year BSL-to-English interpreting task, but this relationship was marginally insignificant (n=8, r2=0.31, t=2.11, p=0.088); c) pre-degree 3D-MR was positively correlated with scores on the 3rd-year English-to-BSL interpreting task, but this relationship was not significant (n=9, r2=0.27, t=1.32, p=0.24); d) pre-degree 3D-MR was correlated with scores on the 3rd-year BSL-to-English interpreting task, but this relationship was marginally insignificant (n=8, r2=0.48, t=2.004, p=0.101).

6. Data Availability Statement

The datasets generated for this study can be found in the Open Science Framework (OSF) repository for this paper: https://osf.io/kjctg.

7. Conflict of Interest

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

8. Author Contributions

All authors developed the idea for this study. FW wrote the first draft of the manuscript and collected and analyzed the majority of the data, with data collection assistance from RT and SW. SW and CS facilitated initial data collection sessions and created and coded the final interpreting tasks. RT also coded some data and assisted with the analysis. All authors were responsible for revising the paper.

9. Funding

10. Acknowledgments

11.   References

Ackerman, J. M., Wolsey, J.-L. A., and Clark, M. D. (2018). Locations of L2/Ln Sign Language Pedagogy. *Creative Education, 09*(13), 2037. doi: 10.4236/ce.2018.913148

Baddeley, A. D., Hitch, G. J., and Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*(3), 438–456. doi: 10.1016/j.jml.2009.05.004

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29. doi: 10.1146/annurev-psych-120710-100422

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. doi: 10.18637/jss.v067.i01

Beal, J. S., Scheetz, N., Trussell, J., McAllister, A., and Listman, J. (2018). University American Sign Language Learners: Longitudinal Self- and Faculty Evaluation Ratings. *Journal of Interpretation*, *26*(1). https://digitalcommons.unf.edu/joi/vol26/iss1/5

Ben-David, B.M., Hadas, E., Goy, H., and Schneider, B. A. (2015). "Older is always better": Age-related differences in vocabulary scores across 16 years, *Psychology and Aging*, 30(4), 856–862. doi:10.1037/pag0000051.

Berteau-Pavy, D., Raber, J., and Piper, B. J. (2011). Contributions of age, but not sex, to mental rotation performance in a community sample. *PEBL Technical Report Series*. http://sites.google.com/site/pebltechnicalreports/home/pebl-technical-report-2011-02

Bontempo, K., and Napier, J. (2009). Getting it right from the start: Program admission testing of signed language interpreters. In C. V. Angelelli and H. E. Jacobson (Eds.), *Testing and assessment in translation and interpreting studies* (pp. 247–295). John Benjamins Publishing Company. doi: 10.1075/ata.xiv.13bon

Bontempo, K., Napier, J., Hayes, L., and Brashear, V. (2014). Does personality matter? An international study of sign language interpreter disposition. *Translation and Interpreting*, *6*(1), 23–46. doi: ti.106201.2014.a02

Carroll, J. B., and Sapon, S. M. (1959). Modern Language Aptitude Test. *Psychological Corporation.*

Carroll, J. B., Sapon, S. M., Reed, D. J., and Stansfield, C. W. (2010). *Modern Language Aptitude Test: Manual.* Second Language Testing Foundation.

Carver, C. S., and White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology, 67*(2), 319–333. doi: 10.1037/0022-3514.67.2.319

CFA. (2012). *Interpreting: National Occupational Standards.* Retrieved from https://dpsionline.co.uk/wp-content/uploads/2019/03/National-Occupational-Standards-Interpreting-2006.pdf Last accessed: 12th April 2022

Chmiel, A. (2018). In search of the working memory advantage in conference interpreting – Training, experience and task effects. *International Journal of Bilingualism*, *22*(3), 371–384. doi: 10.1177/1367006916681082

Christoffels, I.K., de Groot, A.M.B. and Kroll, J.F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency, *Journal of Memory and Language*, 54(3), 324–345. doi: 10.1016/j.jml.2005.12.004.

Cokely, D. (1986). College Level Sign Language Programs: A Resource List. *Sign Language Studies*, *50*, 78–92. JSTOR.

Cormier, K., Adam, R., Rowley, K., Woll, B., and Atkinson, J. (2012). The BSL Sentence Reproduction Test: Exploring age-of-acquisition effects in British deaf adults. *34th Annual Meeting of the German Association of Linguistics, Frankfurt am Main*.

Corsi, P. M. (1972). *Human memory and the medial temporal region of the brain*. [Doctoral Thesis]. McGill University, Montreal.

Croschere, J., Dupey, L., Hilliard, M., Koehn, H., and Mayra, K. (2012). The effects of time of day and practice on cognitive abilities: Forward and backward Corsi block test and digit span. *PEBL Technical Report Series*.

Department for Work and Pensions (DWP). (2017). *Market review of British Sign Language and communications provision for people who are deaf or have hearing loss.* https://www.youtube.com/watch?v=7MguokqjMco

Dong, Y., Liu, Y., and Cai, R. (2018). How Does Consecutive Interpreting Training Influence Working Memory: A Longitudinal Study of Potential Links Between the Two. *Frontiers in Psychology*, *9*. doi: 10.3389/fpsyg.2018.00875

Ebi, K. (2019). How climate change could make our food less nutritious. *TED2019, Vancouver, BC*. https://www.ted.com/talks/kristie_ebi_how_climate_change_could_make_our_food_less_nutritious

Emmorey, K. (1998). The impact of sign language use on visuospatial cognition. In *Psychological Perspectives on Deafness 2*, 19–52.

Emmorey, K. (2002). *Language, Cognition, and the Brain: Insights From Sign Language Research*. Psychology Press.

Emmorey, K., Klima, E., and Hickok, G. (1998). Mental rotation within linguistic and non-linguistic domains in users of American Sign Language. *Cognition*, *68*(3), 221–246. doi: 10.1016/s0010-0277(98)00054-7

Forbes, S. (2020). PupillometryR: A Unified Pipeline for Pupillometry Data. R package version 0.0.3. https://CRAN.R-project.org/package=PupillometryR

Frishberg, N. (1986). *Interpreting: An Introduction.* RID Publications: Silver Spring, Maryland.

Ganis, G., and Kievit, R. (2015). A New Set of Three-Dimensional Shapes for Investigating Mental Rotation Processes: Validation Data and Stimulus Set. *Journal of Open Psychology Data, 3*(1). doi: 10.5334/jopd.ai

Gile, D. (1997). Conference interpreting as a cognitive management problem. *Applied Psychology*. Sage, London, 196-214.

GoReact. (2022). Video + Feedback for Sign & Foreign Languages. Video Assessment For Skill Development & Feedback | GoReact. https://get.goreact.com/sign-foreign-languages/

Grbić, N. (2009). Sign language interpreter training in Austria: An integrated approach. In J. Napier (Ed.), *International Perspectives on Sign Language Interpreter Education* (pp. 3–14). Gallaudet University Press.

Hall, M. L., and Bavelier, D. (2011). Short-term memory stages in sign vs. speech: The source of the serial span discrepancy. *Cognition*, *120*(1), 54–66. doi: 10.1016/j.cognition.2011.02.014

Haug, T., Nussbaumer, D., and Stocker, H. (2019a). *Gebärdensprachdolmetschen: Die Bedeutung der Kognition beim Dolmetschen*. Forschungskolloquium, Zürich. https://www.researchgate.net/publication/332963868_Gebardensprachdolmetschen_Die_Bedeutung_der_Kognition_beim_Dolmetschen

Haug, T., Nussbaumer, D., and Stocker, H. (2019b). Die Entwicklung von Instrumenten zur Überprüfung von kognitiven Fähigkeiten, gebärdensprachlicher Kompetenz und Dolmetschleistung von Gebärdensprachdolmetscherinnen. *Das Zeichen*, *111*, 130–143.

Hebbali, A. (2020). blorr: Tools for Developing Binary Logistic Regression Models. R package version 0.3.0. https://CRAN.R-project.org/package=blorr

Hedderly, R. (1996). Vernon-Warden Reading Test (Revised). *Dyslexia Review, 7*(2).

Hornstra, K. (2021). Teaching through a screen: The challenges of virtual American Sign Language. *The News Record.* https://www.newsrecord.org/news/teaching-through-a-screen-the-challenges-of-virtual-american-sign-language/article_78db1128-a17a-11eb-9453-2fbafbfc1f7a.html

Huhtinen, K. (2014). *Admittance tests 2002−2008 of the Degree Programme in Sign Language Interpretation at the Diaconia University of Applied Sciences: Student selection, admittance test results and their connections to study success.* [MA Thesis]. Diaconia University of Applied Sciences. https://theseus.fi/bitstream/handle/10024/83538/Huhtinen_Kati.pdf

Humphrey, J. (1994). Effective screening procedures for entering students in interpreter education. In E. A. Winston (Ed.), *Mapping our course: A collaborative venture. Proceedings of the Tenth National Convention of the Conference of Interpreter Trainers* (pp. 144–173). https://islandora.wrlc.org/islandora/object/cit%3A9/datastream/PDF/download

Hunt, D. I. J., and Nicodemus, B. (2014). Gatekeeping in ASL-English interpreter education programs: Assessing the suitability of students for professional practice. In D. I. J. Hunt and S. Hafer (Eds.), *Biennial Conference – Our Roots: The Essence of Our Future Conference Proceedings* (pp. 44–60).

Jaeggi, S. M., Buschkuehl, M., Jonides, J., and Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences of the United States of America, 105*(19), 6829–6833. doi: 10.1073/pnas.0801268105

Kartheiser, G., Kurz, K. B., Emmorey, K., Brozdowski, C., and Hauser, P. C. (2022, June 23). *Nonlinguistic Spatial Cognitive Skills Facilitate Second Language Learning in the Visual Modality.* 4th International Conference on Sign Language Acquisition, Boston University Wheelock College of Education & Human Development [online].

Katz, S. (2021, February 26). The COVID Zoom Boom Is Reshaping Sign Language. *Scientific American.* https://www.scientificamerican.com/article/the-covid-zoom-boom-is-reshaping-sign-language/

Kaufman, A. S., and Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test, Second Edition*. Pearson, Inc.

Kazak, A. E. (2018). Editorial: Journal Article Reporting Standards. *American Psychologist, 73*(1), 1-2. doi: 10.1037/amp0000263

Keehner, M., and Gathercole, S. E. (2007). Cognitive adaptations arising from nonnative experience of sign language in hearing adults. *Memory & Cognition*, *35*(4), 752–761. doi: 10.3758/BF03193312

Kessels, R.P.C., van Zandvoort, M.J.E., Postma, A., Kappelle, L.J., de Haan, E.H.F., 2000. The Corsi Block-Tapping Task: Standardization and Normative Data. *Applied Neuropsychology* 7, 252–258. doi: 10.1207/S15324826AN0704_8

Kubicek, E., and Quandt, L. C. (2021). A Positive Relationship Between Sign Language Comprehension and Mental Rotation Abilities. *The Journal of Deaf Studies and Deaf Education*, *26*(1), 1–12. doi: 10.1093/deafed/enaa030

Lara Burgos, P. and de los Santos Rodríguez, E. (2000). *Técnicas de interpretación de lengua de signos.* Madrid: CNSE.

Leitner, K. (2012). *Studienabbrüche im Studiengang ÖGS - Ursachen und Konsequenzen* [MA Thesis, Karl-Franzens-Universität Graz]. https://resolver.obvsg.at/urn:nbn:at:at-ubg:1-35340

Lemon, J. (2006). Plotrix: a package in the red light district of R. *R-News* 6(4): 8-12.

Liesefeld, H. R., Fu, X., and Zimmer, H. D. (2015). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1140–1151. doi: 10.1037/xlm0000081

Liesefeld, H. R., and Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*(1), 40–60. doi: 10.3758/s13428-018-1076-x

Lin, Y., Lv, Q., and Liang, J. (2018). Predicting Fluency With Language Proficiency, Working Memory, and Directionality in Simultaneous Interpreting. *Frontiers in Psychology*, *9*. doi: 10.3389/fpsyg.2018.01543

Liu, M., Schallert, D. L., and Carroll, P. J. (2004). Working memory and expertise in simultaneous interpreting. *Interpreting*, *6*(1), 19–42. doi: 10.1075/intp.6.1.04liu

López Gómez, M. J., Molina, T. B., Benítez, P. P., and Santiago de Torres, J. (2007). Predicting proficiency in signed language interpreting: A preliminary study. *Interpreting, 9*(1), 71–93. doi: 10.1075/intp.9.1.05lop

Lüdecke, D. (2021). sjPlot: Data Visualization for Statistics in Social Science. R package version 2.8.10. https://CRAN.R-project.org/package=sjPlot

Macnamara, B. (2009). Interpreter Cognitive Aptitudes. *Journal of Interpretation*, 2008–2009, 9–32.

Macnamara, B. N., and Conway, A. R. A. (2016). Working Memory Capacity as a Predictor of Simultaneous Language Interpreting Performance. *Journal of Applied Research in Memory and Cognition, 5*(4), 434–444. doi: 10.1016/j.jarmac.2015.12.001

Macnamara, B. N., Moore, A. B., Kegl, J. A., and Conway, A. R. A. (2011). Domain-general cognitive abilities and simultaneous interpreting skill. *Interpreting*, *13*(1), 121–142. doi: 10.1075/intp.13.1.08mac

Madley-Dowd, P., Hughes, R., Tilling, K., and Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, *110*, 63–73. doi: 10.1016/j.jclinepi.2019.02.016

Marks, J. (2014). Cognitive Measures in Interpreter Education: An Examination of Admissions Screening in ASL-English Interpreter Education Programs. In D. I. J. Hunt and S. Hafer (Eds.), *Our Roots: The Essence of Our Future Conference Proceedings*. https://www.unco.edu/cebs/asl-interpreting/pdf/osep-project/2014-cit-proceedings.pdf

Martinez, D., and Singleton, J. L. (2018). Predicting sign learning in hearing adults: The role of perceptual-motor (and phonological?) processes. *Applied Psycholinguistics*, *39*(5), 905–931. doi: 10.1017/S0142716418000048

Martinez, D., and Singleton, J. L. (2019). Individual differences in lexical learning across two language modalities: Sign learning, word learning, and their relationship in hearing non-signing adults. *Acta Psychologica*, *198*, 102892. doi: 10.1016/j.actpsy.2019.102892

Mayberry, R. I., and Fischer, S. D. (1989). Looking through phonological shape to lexical meaning: The bottleneck of non-native sign language processing. *Memory & Cognition*, *17*(6), 740–754. doi: 10.3758/bf03202635

McKee, D. (1987). *An analysis of specialized cognitive functions in deaf and hearing signers.* Unpublished doctoral dissertation. University of Pittsburgh, Pittsburgh, PA.

Moè, A., and Pazzaglia, F. (2006). Following the instructions!: Effects of gender beliefs in mental rotation. Learning and Individual Differences, 16(4), 369–377. doi: 10.1016/j.lindif.2007.01.002

Monikowski, C., and Peterson, R. (2005). Service Learning in Interpreting Education: Living and Learning. In M. Marschark, R. Peterson, and E. A. Winston (Eds.), *Sign Language Interpreting and Interpreter Education: Direction for Research and Practice* (pp. 188–207). Oxford University Press. https://www.oxfordscholarship.com/view/10.1093/acprof/9780195176940.001.0001/acprof-9780195176940-chapter-8

Morales, J., Padilla, F., Gómez-Ariza, C. J., and Bajo, M. T. (2015). Simultaneous interpretation selectively influences working memory and attentional networks. *Acta Psychologica*, *155*, 82–91. doi: 10.1016/j.actpsy.2014.12.004

Moser-Mercer, B. (1985). Screening Potential Interpreters. *Meta: Journal des traducteurs*, *30*(1), 97. doi: 10.7202/003631ar

Mueller, S. T. (2014). *PEBL: The Psychology Experiment Building Language (Version 0.14)* [Software]. Available from http://pebl.sourceforge.net

Mueller, S. T., and Piper, B. J. (2014). The Psychology Experiment Building Language (PEBL) and PEBL Test Battery. *Journal of Neuroscience Methods, 222,* 250–259. doi: 10.1016/j.jneumeth.2013.10.024

Müller, K. (2020). here: A Simpler Way to Find Your Files. R package version 1.0.1. https://CRAN.R-project.org/package=here

Nakano, S. (2021). A Review of Studies Related to the Aptitude in Sign Language Acquisition by Hearing Adults: Cognitive Factors. *Bulletin of the Faculty of Co-Education, Gunma University, Humanities and Social Sciences*, *70*, 165–174.

Napier, J., Skinner, R., Adam, R., Stone, C., Pratt, S., & Obasi, C. (2021, August). A demographic snapshot of the profession: The 2021 Census of sign language translators & interpreters in the UK. *Association of Sign Language Interpreters UK & Heriot-Watt University.*

Nicodemus, B., and Emmorey, K. (2013). Direction asymmetries in spoken and signed language interpreting. *Bilingualism: Language and Cognition*, *16*(3), 624–636. doi: 10.1017/S1366728912000521

Ortega, G., and Morgan, G. (2015). Input processing at first exposure to a sign language. *Second Language Research*, *31*(4), 443–463. doi: 10.1177/0267658315576822

Padden, C. A. (2000). Simultaneous Interpreting across Modalities. *Interpreting*, *5*(2), 169–185. doi: 10.1075/intp.5.2.07pad

Padilla, F., Bajo, M. T., and Macizo, P. (2005). Articulatory suppression in language interpretation: Working memory capacity, dual tasking and word knowledge. *Bilingualism: Language and Cognition*, *8*(3), 207–219. doi: 10.1017/S1366728905002269

Passig, D., and Eden, S. (2001). Virtual Reality as a Tool for Improving Spatial Rotation among Deaf and Hard-of-Hearing Children. *CyberPsychology & Behavior*, *4*(6), 681–686. doi: 10.1089/109493101753376623

Patton, J. H., Stanford, M. S., and Barratt, E. S. (1995). Factor Structure of the Barratt Impulsiveness Scale. *Journal of Clinical Psychology, 51*(6), 768–774. doi: 10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Pedersen, T. L. (2020). patchwork: The Composer of Plots. R package version 1.1.1. https://CRAN.R-project.org/package=patchwork

Peirce, J., Hirst, R., and MacAskill, M. (2022). *Building Experiments in PsychoPy (2nd ed.).* SAGE Publishing.

Peterson, R. (1999). The Perceptions of Deafness and Language Learning of Incoming ASL Students. *Proceedings of the 16th National Convention of the Registry of Interpreters for the Deaf*, 185–237.

Pimsleur, P. (1966). *The Pimsleur Language Aptitude Battery.* New York: Harcourt, Brace, Jovanovic.

Pivac, L. (2014). Learner Autonomy in New Zealand Sign Language Interpreting Students. In D. McKee, R. S. Rosen, and R. L. McKee (Eds.), *Teaching and Learning Signed Languages: International Perspectives and Practices*, 197–221. Palgrave Macmillan UK. doi: 10.1057/9781137312495_10

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Robinson, O. (2022, February 22). Abled Arrogance, Not Linguistic Carework. *Medium*. https://notanangrydeafperson.medium.com/abled-arrogance-not-linguistic-carework-77434eae2cee

Robinson, O., and Henner, J. (2018). Authentic Voices, Authentic Encounters: Cripping the University Through American Sign Language. *Disability Studies Quarterly*, *38*(4). doi: 10.18061/dsq.v38i4.6111

Russo, M. (2011). Aptitude testing over the years. *Interpreting, 13*(1), 5–30. doi: 10.1075/intp.13.1.02rus

Russo, M. (2014). Testing aptitude for interpreting: The predictive value of oral paraphrasing, with synonyms and coherence as assessment parameters. *Interpreting*, *16*(1), 1–18. doi: 10.1075/intp.16.1.01rus

Russo, M., and Pippa, S. (2004). Aptitude to Interpreting: Preliminary Results of a Testing Methodology Based on Paraphrase. *Meta : Journal Des Traducteurs / Meta: Translators' Journal*, *49*(2), 409–432. doi: 10.7202/009367ar

Shaw, S. (2011). Cognitive and motivational contributors to aptitude: A study of spoken and signed language interpreting students. *Interpreting*, *13*(1), 70–84. doi: 10.1075/intp.13.1.05sha

Shepard, S., and Metzler, D. (1988). Mental Rotation: Effects of Dimensionality of Objects and Type of Task. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(1), 3-11. doi: 10.1037/0096-1523.14.1.3

Speekenbrink, M. (2021). sdamr: Statistics: Data Analysis and Modelling. R package version 0.1.0. https://CRAN.R-project.org/package=sdamr

Stone, C. (2017). Sign language interpreter aptitude: The trials and tribulations of a longitudinal study. *Translation & Interpreting, 9*(1), 16. doi: 10.12807/ti.109201.2017.a06

Strobach, T., Becker, M., Schubert, T., and Kühn, S. (2015). Better dual-task processing in simultaneous interpreters. *Frontiers in Psychology*, *6*. doi: 10.3389/fpsyg.2015.01590

Talbot, K. F., and Haude, R. H. (1993). The relation between sign language skill and spatial visualization ability: Mental rotation of three-dimensional objects. *Perceptual and Motor Skills*, *77*, 1387–1391. doi: 10.2466/pms.1993.77.3f.1387

Timarová, Š., and Ungoed-Thomas, H. (2008). Admission Testing for Interpreting Courses. *The Interpreter and Translator Trainer*, *2*(1), 29–46. doi: 10.1080/1750399X.2008.10798765

Torchiano, M. (2020). effsize: Efficient Effect Size Computation. R package version 0.8.1. doi: 10.5281/zenodo.1480624

Tzou, Y.-Z., Eslami, Z. R., Chen, H.-C., and Vaid, J. (2012). Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: Evidence from Mandarin–English speakers. *International Journal of Bilingualism*, *16*(2), 213–227. doi: 10.1177/1367006911403197

Volk, C. (2014, October 8). Sign Language Interpreter Education: Time for a National Call to Action. *StreetLeverage*. https://streetleverage.com/2014/10/sign-language-interpreter-education-time-for-a-national-call-to-action/

Wang, J. (2016). The relationship between working memory capacity and simultaneous interpreting performance: A mixed methods study on professional Auslan/English interpreters. *Interpreting*, *18*(1), 1–33. doi: 10.1075/intp.18.1.01wan

Wang, J., and Napier, J. (2013). Signed language working memory capacity of signed language interpreters and deaf signers. *Journal of Deaf Studies and Deaf Education*, *18*(2), 271–286. doi: 10.1093/deafed/ens068

Warden, V. B. (1956). *Construction and Standardisation of a Reading Test.* [Unpublished MA thesis]. University of London.

Watkins, F., Abdlkarim, D., and Thompson, R. L. (2018). *L1 and L2 sign recognition: The role of visual angle*. 3rd International Conference on Sign Language Acquisition, Koç Üniversitesi, İstanbul. https://osf.io/p7hkt/

Watkins, F., and Thompson, R. L. (2019). *Angle-diverse input and sign vocabulary learning in novice L2 British Sign Language learners*. Theoretical Issues in Sign Language Research 13, Universität Hamburg. https://osf.io/9vfzq/

Webb, S. (2017). *Job Demands, Job Resources, Wellbeing and Student Outcomes: A Study of Sign Language Interpreter Educators' Perceptions* [PhD Thesis, Heriot-Watt University]. https://www.ros.hw.ac.uk/bitstream/handle/10399/3442/WebbSL_1117_soss.pdf

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., ... and Yutani, H. (2017). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686, doi: 10.21105/joss.01686

Wickham H., and Seidel, D. (2020). scales: Scale Functions for Visualization. R package version 1.1.1. https://CRAN.R-project.org/package=scales

Williams, J., Darcy, I., and Newman, S. (2015). Second Language Working Memory Deficits and Plasticity in Hearing Bimodal Learners of Sign Language. *Psychology of Language and Communication*, *19*(2). doi: 10.1515/plc-2015-0008

Williams, J. T., Darcy, I., and Newman, S. D. (2017). The beneficial role of L1 spoken language skills on initial L2 sign language learning: Cognitive and linguistic predictors of M2L2 acquisition. *Studies in Second Language Acquisition*, *39*(04), 833–850. doi: 10.1017/S0272263116000322

Williams, J., and Newman, S. (2016). Phonological substitution errors in L2 ASL sentence processing by hearing M2L2 learners. *Second Language Research*, *32*(3), 347–366. doi: 10.1177/0267658315626211

Wilson, M., and Emmorey, K. (2003). The Effect of Irrelevant Visual Input on Working Memory for Sign Language. *The Journal of Deaf Studies and Deaf Education*, *8*(2), 97–103. doi: 10.1093/deafed/eng010

Witter-Merithew, A., and Johnson, L. (2005). *Toward Competent Practice: Conversations with Stakeholders*. Registry of Interpreters for the Deaf Conference 2005.

Yorke, M., Bridges, P., and Woolf, H. (2000). Mark Distributions and Marking Practices in UK Higher Education: Some Challenging Issues. *Active Learning in Higher Education, 1*(1), 7–27. doi: 10.1177/1469787400001001002

Yudes, C., Macizo, P., and Bajo, T. (2011). The Influence of Expertise in Simultaneous Interpreting on Non-Verbal Executive Processes. *Frontiers in Psychology*, *2*. doi: 10.3389/fpsyg.2011.00309