Examining the Construct of Proficiency in a University's American Sign Language
(ASL) Program: A Mixed-Methods Study

By

Josée-Anna Tanner

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial
fulfilment of the requirements for the degree of

Master of Arts

in

Applied Linguistics and Discourse Studies

Carleton University
Ottawa, Ontario

© 2014
Josée-Anna Tanner

Abstract

American Sign Language (ASL) has become increasingly popular as a second language option at universities and colleges in North America. While a growing number of hearing, adult learners are enrolling in ASL classes, this has not been paralleled (yet) by an equal development in ASL research. There has been insufficient investigation into what constitutes ASL proficiency development and how proficiency can be validly and reliably assessed for this group of learners. This mixed-methods study explores the ASL program at a Canadian university. It investigates the construct of proficiency from three angles: instructors' understanding and definitions of ASL proficiency; how student proficiency is determined through current assessment practices and; student responses to assessment practices. Results of this study suggest that in this context ASL proficiency is not clearly defined. Without a clear construct definition, what current ASL assessments are actually measuring is unknown, and consequently may not be providing valid results.

Acknowledgements

Thank you to my supervisor, Janna Fox. Your guidance and support throughout this process were greatly appreciated, and your experience and insight were invaluable.

I am also indebted to those who agreed to volunteer their time and be participants in this study. Thank you for sharing your experiences with me.

A very special thank you to Jon Kidd, for helping me get here in the first place.

Thank you also to my family, especially my Mum, for cheering for me through this and all of my endeavours.

Thank you to Claire for all of the talks about work, life, writing and everything else.

And finally thank you to my partner Liam for helping me stay grounded. You have been wonderfully supportive throughout this process – from offering insightful feedback to making sure I stopped working when it was time to stop working. Thank you.

List of Abbreviations

ACTFL – American Council on the Teaching of Foreign Languages

ASL – American Sign Language

ASLTA – American Sign Language Teachers Association

CEFR – Common European Framework of Reference for Languages

CODA – Child of Deaf Adults. Hearing child with one (or more) Deaf parents.

EFA – Exploratory Factor Analysis

LSQ – Langues des signes québécoise

FL – Foreign language

IELTS – International English Language Testing System

L2 – Second language. Any language learned after the first or native language.

M2 – Second modality. Any language learned which relies on a different channel of articulation than the first or native language. Usually refers to hearing people learning a signed language.

MST – Maritime Sign Language

NCATE – National Council for Accreditation of Teacher Education

SLA – Second language acquisition

TOEFL – Test of English as a Foreign Language

TPP – Teacher preparation program

Glossary

Curriculum – The course of study within a program.

Deaf – (Capital D) People who self-identify as members of the social community and culture which revolves around deafness and the use of sign language.

deaf – (Lowercase d) The physical condition of being unable to hear.

Program – A group or selection of courses, identified by a university or other educational institution, as being part or the whole of a path of study.

Table of contents

Abstract.....	ii
Acknowledgements	iii
List of Abbreviations.....	iv
Glossary	v
List of Tables	viii
List of Appendices	ix
Chapter 1: Introduction	1
1.1 The Growth of American Sign Language	2
1.2 What is ASL?	7
1.3 ASL and Multimodality	11
Chapter 2: Literature Review	15
2.1 Cognitive Processing of ASL	15
2.2 Perspectives on Teaching and Learning ASL	21
2.2.1 Teaching ASL as a Second Language	21
2.2.2 Learning ASL as a Second Language	26
2.3 Defining Second Language Proficiency	32
2.4 Assessing Second Language Proficiency	39
2.5 Validity	46
Chapter 3: Method	53
3.1 Research Questions	53
3.2 Research Site and Design	54
3.3 Phase One Data Collection: Qualitative Data	55
3.3.1 Teacher-Participants	55
3.3.2 Teacher-Participant Demographics	56
3.3.4 Procedures	57
3.4 Phase Two Data Collection: Quantitative Data	58
3.4.1 Participants	58
3.4.2 Student Demographics	58
3.4.3 Instrument	60
3.4.4 Procedures	61
3.5 Phase One Analysis: Qualitative Data	63
3.6 Phase Two Analysis: Quantitative Data	65
3.7 Ethical Considerations	66
Chapter 4: Results and Discussion	68
4.1 Qualitative Results	68
4.1.2 Current Assessment Practices	69
4.1.2.1 ASL Assessment Format	69
4.1.2.2 Assessment Content	71
4.1.2.3 Test Type	72

4.1.2.4 Teacher-Participant Attitudes Toward Testing	73
4.1.3 Descriptions of Proficiency	74
4.1.4 Grading	79
4.1.5 Validity Concerns	81
4.1.5.1 Under-Representation of the Construct	82
4.1.5.2 Language Transfer	82
4.1.5.3 Method Effect	84
4.1.6 Teacher-Participant Perceptions of Students	86
4.1.7 Change	89
4.2 Summary of Qualitative Results	91
4.2.1 Addressing Research Question #1	92
4.2.2 Addressing Research Question #2	92
4.3 Quantitative Results	93
4.3.1 Why ASL?	94
4.3.2 Feedback	95
4.3.3 Test Format and Method Effect	95
4.3.4 Grading	98
4.3.5 Proficiency Development	99
4.3.6 The Next Level	101
4.3.7 Percentage of New Material by Level	102
4.3.8 Assessment Benefits	103
4.3.9 Productive and Receptive Testing	104
4.4 Proficiency Section Factor Analysis	105
4.4.1 Factor 1: ASL Discourse Fluency	105
4.4.2 Factor 2: Basic ASL Communication	105
4.4.3 Factor 3: Language Use Context	107
4.4.4 Factor 4: Comprehension	107
4.4.5 Factor Analysis Summary	108
4.5 Quantitative Results Summary	110
4.5.1 Addressing research question #3	111
Chapter 5: Conclusion	114
5.1 Limitations	114
5.2 Conclusions About ASL Proficiency Development	115
5.3 Conclusions About ASL Assessment Validity	117
5.4 Consequences of Current Assessment Practices	118
5.5 Future directions	120
References	122
Appendices	136

List of Tables

Table 1: Teacher-participants by level of experience and hearing status	57
Table 2: ASL enrolment by course/level code, fall 2013	59
Table 3: Cross-tabulation between instructor hearing status and student-reported amount of feedback.....	95
Table 4: Student response by ASL level to questionnaire item #32	101
Table 5: Student response by ASL level to questionnaire item #33	101
Table 6: Student response by ASL level to questionnaire item #34	101
Table 7: Cross-tabulation between ASL level and student-reported amount of new material	102
Table 8: Pattern Matrix for 4-Factor Solution (EFA)	106

List of Appendices

Appendix A: Ethics Certificate	137
Appendix B: Teacher-Participant Interview Questions	138
Appendix C: Student Questionnaire	139

Chapter 1: Introduction

I have been an ASL instructor for just two years. I am still a new teacher; I do not have the years of experience that others possess. I am also hearing and my first exposure to ASL was as a second language and a second form of communication. While I have only been involved in teaching ASL for a short time, much of that time has been spent planning, administering, and grading assessments. I arrived on the assessment scene with very little knowledge, but started learning fast. It was as I was teaching my second class that I began to wonder more about ASL assessment, and language testing in general. What was I really assessing? What did it mean for a student to be proficient in ASL at different levels, and what did that proficiency look like? What theoretical or pedagogical basis were my own assessments founded on? I became increasingly concerned that the ways in which I was using assessments, and the results I was seeing, were not in line with the outcomes I expected or hoped to achieve. At times I was not sure what my assessments were really measuring at all. There was a steady hum of questions about assessment always in the back of my mind. Before beginning this research, I gave a lot of thought to what I was most concerned with and most interested in knowing and narrowed it down to three research questions, which formed the basis of this study:

1. What constitutes ASL proficiency development for hearing learners acquiring the language in this context?
2. How is ASL proficiency determined through current assessment practice?
3. How well do students understand the construct of proficiency, and how do they respond to the current assessment practices?

1.1 The Growth of American Sign Language

The acceptance of American Sign Language as a second or foreign language, and consequently the teaching of it as a second language, has been sharply on the rise in Canada and the U.S. for at least the past two decades (Rosen, 2010; Jacobowitz, 2005; Wilcox & Wilcox, 1991; Peterson, 2009; Mckee & Mckee, 1992; Thoryk, 2010). Classes are offered at the elementary and high school level, and ASL has also become increasingly popular as a “foreign” language option at many universities and colleges. It is currently the fourth most commonly taught language course offered at colleges in the U.S. (Redden, 2007; Furman, Goldberg, & Lusin, 2010), following Spanish, French, and German. In the U.S., enrolment in ASL courses increased by 29.7% between 2002 and 2006; from 2006-2009 enrolment increased by an additional 16.4% (Furman, Goldberg, & Lusin, 2010). It is only lately that ASL has enjoyed this type of expansion. As recently as 1988 colleges and universities were still debating over whether or not ASL met the requirements of a foreign language. In 1984 a task force from the California State Universities and Colleges System rejected signed languages as foreign languages, grouping them together with computer languages, artificial languages, and dialects of English (Chapin, 1988). This change is visible in the Modern Language Association’s survey results, which record no college enrolment numbers for ASL until 1990 (Furman, Goldberg, & Lusin, 2010).

ASL was first recognized as a naturally occurring, rule-governed human language in 1965, following the work of William Stokoe, making it a much “younger” language in terms of the depth and breadth of research that has been devoted to its linguistic properties. It was not long after this recognition that ASL programs began to be

established, though they may not have been initially well-regarded. The first complete program in Deaf Studies (i.e. a program in which a student could enrol in a major) opened in California in 1975 (Ashton et.al, 2014).

While some students choose to take ASL as a pre-requisite for gaining entrance into an interpreting program, as a prelude to Deaf studies, or in preparation for a teacher-training program, there are also students who choose to take it out of personal interest – either as a one-time elective or as a language minor. How many of these students have a Deaf family member or friend is largely unknown, as are their individual motivations for enrolling in such a course and whether they plan to continue beyond a single class. One survey of 1115 new ASL students at university, conducted by Peterson (2009), provides some broad demographic information from which to begin. He found that students were overwhelmingly female (80%), monolingual English speakers between the ages of 18 and 24. His survey investigated incoming students' perceptions of ASL, finding that many students do not distinguish ASL from English, or believe that it is a manual form of English. Peterson's survey revealed that students' initial understanding or expectations of ASL are often at odds with factual information about the language. While it is not clear yet why ASL is so popular amongst hearing L2 learners, there could be some argument made for it being misconceived as a substantially 'easier' language to learn, which may play a role in driving up enrolment.

ASL is often identified as the natural language of Deaf and/or hard of hearing people in North America. This is a broad generalization, as it excludes Quebec, where *langue des signes québécoise* (LSQ) is primarily used, as well as east coast communities of Canada where maritime sign language (MST, a dialect of ASL), may be used. For the

purposes of this study, ASL was chosen as it is what is currently taught in the location being used as the study base, and also because it has one of the largest bodies of sign language research associated with it.

The recent upswing in the number of ASL programs across the country means that the language is being taught to an increasing number of hearing learners. In Canada, ASL is currently offered as a foreign language class at ten universities (Doré, 2014). Given the swift rise of ASL programs, both at colleges and universities, one pressing concern is the lack of research into the ways in which ASL is taught and learned. The problem, as pointed out by Quinto-Pozos (2009) is partially due to a lack of communication between ASL teachers and educators, researchers, and professionals from other language units. Additionally, there are very few resources (journal articles, books, papers) that investigate ASL pedagogy as it applies to hearing adult learners, and fewer still that have tried to apply theories of second language acquisition (SLA) to visual languages being taught to hearing students. Even more rare are attempts to develop new theories for ASL teaching or learning that are drawn out of visual language research. As Quinto-Pozos (2009) suggests

it appears that the teaching of ASL is informed, in large part, by the linguistic and cultural beliefs of its instructors and curricula developers – most of whom are members of the Deaf community... The teaching strategies that have been employed have allowed for regular instruction of ASL to many adult hearing students, though the efficacy of those strategies has generally not been examined empirically (p. 138).

There have been some calls recently (Thoryk, 2010) for improved instructional materials to support successful learners and teachers and to “guarantee continued respect for the validity of the language and the culture” (p. 116). It is no longer enough for an

instructor to simply be a proficient signer (Thoryk, 2010); according to Kanda and Fleischer (1988) they should also “be educated, demonstrating knowledge and application of educational and pedagogical principles along with formal study of the language being taught. Sign language teachers should be able to integrate second language teaching theory and methodology in their classrooms” (p. 193).

In 1988, with increasing recognition that acceptance of ASL was beginning to grow in the U.S., Gary Olsen, the then-executive direction of the National Association of the Deaf, cautioned that “A ground swell can be devastating if proper precautions are not taken. Without proper standardization of curriculum, instructor qualifications, and certification, this sudden spontaneity of acceptance of ASL can be devastating. ASL, as we know it today, could be drastically altered or lost” (Olsen, p. 100). Standardization is often considered to be a testing and assessment issue. In second language testing, the purpose of testing is to assess, in one way or another, the ability of a learner to use a language which is not their first language. Second language tests are also tools that provide information about a variety of issues which are intimately connected to language teaching. These concerns may be as immediate as measuring the proficiency of a particular learner in order to make decisions about that learner, such as giving them a grade at the end of a course. Language testing can “provide evidence on the results of learning and instruction, and hence feedback on the effectiveness of the teaching program itself” (Bachmann & Palmer, 1996, p. 8). Second language assessments have consequences beyond an individual learner’s grade, or beyond a single course; they can have serious implications for how a language is taught, and ultimately what the learners ‘get out of it’. Because of this, language assessments themselves have to be carefully

examined in order to determine whether the construct to be assessed is fully understood and to ensure that the validity of the test is grounded in its “usefulness” (Bachman & Palmer, 1996). Increased research and expectations of higher achievements in all aspects relating to language teaching, learning, and assessment are necessary if ASL, as a modern language, is to move out of its infancy and develop as a more mature field of study.

Since the primary concern in proficiency measurement is understanding the construct to be assessed (Messick, 1989a), the current study aims to contribute to this understanding of ASL assessments in several ways: 1) By documenting the construct of proficiency development for hearing learners of American Sign Language at the university level; and 2) By describing how the construct of proficiency is operationalized through current testing practice across four years of undergraduate study. Describing the current practices and relating them to existing second language assessment frameworks can provide a lens to make explicit what is currently implicit in ASL testing for the purpose of better understanding:

- What constitutes ASL proficiency development for hearing learners acquiring the language in this context;
- How proficiency is determined through the current assessment practice; and
- How well students understand the construct of proficiency and how they respond to the current assessment practice.

Second language assessments must be able to link arguments for test validity to test use (Bachman & Palmer, 1996), while at the same time allowing those engaged with ASL testing, including instructors, to provide fair and useful proficiency assessments to

an increasing number of students. Once the construct of ASL proficiency is better understood and articulated, tests can be designed that more usefully (Bachman & Palmer, 1996) represent the construct. This will also allow a learning trajectory to be elaborated. This does not currently exist for this group of learners. A clearer understanding of ASL proficiency and how it develops may also help teachers to determine if what they are teaching, and how they are teaching it, is supporting the development of student proficiency in the best way possible.

Having introduced the focus of this thesis, the next section will provide a brief introduction to, and description of, American Sign Language as a language. This is followed by a short discussion of multimodality in ASL.

1.2 What is ASL?

Before going into more depth investigating the research surrounding American Sign Language, it is beneficial to first have an understanding of what the language looks like: how it functions linguistically and semantically, and its importance as a language to the main community of users. American Sign Language (ASL) is one example of a signed language. Signed languages are also referred to as manual-visual languages or visual-gestural languages, meaning they are expressed through the hands, face, and body and received through the eyes. Signed languages exist all over the world and vary widely; while some may share common lexical items or certain grammatical forms, they are not necessarily mutually intelligible. Signed languages, including ASL, are naturally occurring human languages; while it is used as a primary language for many people who are deaf or hard of hearing within most of America and Canada, ASL developed independently from English.

ASL is not pictorial, nor is it a form of pantomime (Lane, Hoffmeister, & Bahan, 1996), however it is important to recognize that ASL does contain examples of iconicity, where the form of the signs “reflects some physical aspect of the entities or activities they represent” (Valli, Lucas, & Mulrooney, 2005, p. 5). Previously, linguists tended to emphasize the arbitrary nature of signed languages as they “had a definite sense that admitting the existence of iconicity in signs languages was admitting that sign languages were not “real” languages” (p. 5). However, the language remains more arbitrary than iconic. Semantically, the meanings that are attached to signs in ASL are unique to ASL. A similar-looking sign may have an altogether different meaning in another signed language, while the same signified object or concept may be expressed in a completely different way, visually, in different signed language. While English continues to have an impact on ASL (Peterson, 2009) and there may appear to be some overlap between the two, there is no one-to-one correspondence between an English word and a sign, and the structure of ASL also differs significantly from English.

ASL is a grammatical (rule-based) language, and those acquiring it as a first language from birth go through the same developmental stages as children learning a spoken language: with babbling beginning at 7-10 months of age, and progressing to word modification and rules for sentences between 22-36 months. In fact, “What seems to determine the progress of language acquisition is the complexity of the rules the child must master and not the mode of language, spoken or manual, in which they must be mastered” (Lane, Hoffmeister, & Bahan, 1996, p. 48). When discussing ASL phonology, Valli, Lucas, and Mulrooney (2005) include five sign parameters:

1. Movement. The way in which the hands and arms move in when producing a sign in three dimensional space. Movements may include specific directions (forward, down, left, etc.), taps, circles, arcs, and more.
2. Handshape. The shape(s) that the hand or hands take when producing the sign.
3. Location. The area within the signing space or on the body where the sign is produced.
4. Palm orientation. The direction in which the palm of the hand or hands face when producing a sign.
5. Nonmanual signals, which are also referred to as facial expressions.

At a basic phonological level, signs in ASL are composed of moves and holds (Lane, Hoffmeister, & Bahan, 1996). Moves and holds may be thought of as the outer boundaries of signs; all signs are made up of, at minimum, a move and a hold. Some signs will follow a hold, movement, hold pattern, meaning that the sign has a start position (a hold), a movement, and then an ending position (a final hold). Within those boundaries, “A hold segment has handshape, location, orientation and non-manual features, and likewise, a movement segment has handshape, location, orientation, and nonmanual features” (Valli, Lucas, & Mulrooney, 2005, p. 49). Together, these five parameters form the phonological building blocks of signs. Each part is significant; a difference in one segment may be responsible for signalling a contrast in meaning. The movement-hold model, developed by Liddell and Johnson (1989) further argued that signs are more than just simultaneous bundles of primes, and that in fact the recognition of particular sequencing of primes is necessary in order to provide “an adequate description of many phenomena in ASL” (p. 197).

The formal system of rules that governs ASL (its grammar) is not English-derived or dependent. Like any language, ASL can be modulated by grammar and syntax but, as Sacks (1989) writes, “There is an extraordinary richness of such devices in ASL, which serve to amplify the basic vocabulary hugely” (p. 84). By ‘basic vocabulary’ Sacks is not directing an insult towards ASL, but rather pointing out a fact: if one were to line up a dictionary of ASL and a dictionary of English, the English dictionary will always be thicker and contain more words than a sign dictionary will contain signs. Sacks (1989) points out that this does not limit the expressive properties of ASL, as the language has an “immense number of grammatical modulations” (p. 79) which allow a relatively small vocabulary to be highly flexible and adaptable, and which enables ASL to be used to discuss both concrete and abstract ideas.

The most obvious, visible difference between ASL and a spoken language is that in ASL, lexical items and grammar are mapped onto the body of the signer and into the space surrounding her. The use of space, the face, and the body permeates every linguistic aspect of ASL, so much so that “much of what occurs linearly, sequentially, temporally, in speech, becomes simultaneous, concurrent, multileveled in Sign... what looks so simple is extraordinarily complex and consists of innumerable spatial patterns nested, three-dimensionally, in each other” (Sacks, 1989, p. 87). The visual-gestural modality of ASL is intimately connected to the form of the human body, and is shaped and bounded in part by available physical movements, articulatory points such as joints, and what is visible to the eye. Likewise, the grammar is so completely physical that there are rules not just for the shape and movement of signs, but for mouth shapes, eyebrow lifts, head tilts, eye-gaze, and body shifts. These can be obvious, such as a head shake to

indicate negation, and may be clear even to a novice or non-signer, while other aspects are more subtle, and may only become salient as experience with the language increases.

ASL is of both immeasurable linguistic and cultural importance to the Deaf community. As Lane, Hoffmeister, and Bahan (1996) write,

Signed language is the most important instrument for communicating in the DEAF-WORLD. The language competency of members varies, depending on whether signed language is their native language (true only of the children of members of the Deaf-World) or whether they learned the language as small children or later in life...But regardless of when they learn it, from the day Deaf Americans enter the Deaf-World, ASL becomes their primary language, and is, in itself, a big chunk of DEAF-WORLD Knowledge” (p. 6).

Cultural aspects associated with ASL should not be ignored when it comes to teaching ASL as a second language. Many ASL classes include cultural components which focus on the lives of Deaf people within North American or around the world, and which introduce students to the cultural differences between Deaf and hearing people.

1.3 ASL and Multimodality

According to Djonov and Zhao (2014), all human communication is multimodal. Multimodality “is based on the assumption that meaning is made through the many means (we call these modes) that a culture has shaped for that purpose” (Kress et al., 2005, p. 21). The process or act of meaning-making “involves selecting from different modes (e.g., written language, sound, gesture, visual design) and media (e.g., face-to-face, print, film) and combining these selections according to the logic of space (e.g., a film), time (e.g., a sound composition), or both” (Djonov & Zhao, 2014, p. 1). Djonov and Zhao define the domain of multimodality as “the ways different semiotic resources such as language, colour, dress, and gesture are concurrently deployed in communication

across a range of overlapping spheres such as politics, art, education, and everyday interactions” (p. 2).

The presentation of visual languages presents an interesting problem in multimodal studies and analysis. A visual language such as ASL is somewhat restricted in terms of modes and media. The media is necessarily going to be visual in some way, primarily face-to-face or through video – either pre-recorded or in real time. Since the primary users of the language are deaf the spoken mode only works one way – from deaf people who choose to speak, to hearing recipients. While there are vocalizations associated with particular signs, the sound of these (the aural) is far less important than the associated mouth movements (the visual). Since there is no written system for the language itself, any written communication is done in English or whatever the dominant written language of the social context is. There may be a tension between written English and ASL, as it can be difficult to reconcile semantic differences between the two. In other ways, the visual mode is much more expansive than it may be for spoken languages; what is considered a peripheral or additional mode for spoken language is part of the linguistics of ASL. Body movements, eye-gaze, facial expressions and use of space may be separate communicative modes when considering multimodal aspects of spoken communication, but for ASL they are a necessary part of the formal system of the language. The question becomes ‘how much of what is used during communication events in a visual language can be defined in terms of linguistic properties, and how much can be attributed to multimodality?’ Slobin (2008) argues that categorization of linguistic versus non-linguistic components of sign language are best set aside “until we have a better understanding of the range of gradient and body and facial components in

both signed and spoken languages” (p. 116). However, for the purposes of multimodal studies some categorization would need to occur, if only to examine how the modes and resources combine to make meaning.

Like any other language, ASL can be considered multimodal, although the understanding of the primacy of modes may require some alterations. As Jewitt (2007) describes, “communication can be understood as a product of people’s active engagement with the semiotic resources available to them at a particular moment in a specific sociocultural context” (p. 275). Meaning is made by choosing, combining, and manipulating these modes and resources, whatever they may be. With ASL, the function of the modes differs than what is presented in traditional multimodal studies; the visual is expanded and takes precedence, and aural modes are either incidental (vocalizations), or can only flow in one direction (speech). It may also be difficult at times to separate gesture from signs, though perhaps a working definition is that signs are part of the formalized system of the language, while gestures are socio-culturally specific, and are used and understood by both deaf and hearing people equally.

It may, perhaps, be especially important to acknowledge multimodal forms of communication in teaching ASL. Jewitt (2007) writes that “language is embedded in an ensemble of modes, and modes play a central part in meaning-making” (p. 280). In a language learning context, within a classroom, the teacher may engage with an ensemble of modes when presenting information to the class. Even an immersion setting, teacher to student (or student to teacher) communication is not achieved through ASL alone. Within the classroom, a teacher may present information using the language itself, but may also rely on gesture or acting, pointing, pausing/rate of signing and eye-contact. The teacher’s

signing may be followed, or preceded, by writing on the board. Signing may be done while concurrently showing video or still images. Some teachers (both Deaf and hearing) may use speech if they are comfortable doing so. In these respects ASL, and perhaps more particularly ASL instruction, can be considered multimodal. This is a fascinating area that requires future research, as it is not within the scope of this study.

In summary, ASL is a naturally occurring visual language, which includes semantic, syntactic, and phonological components (Battison, 1978). The phonological elements of the language are encoded through handshapes, locations, palm orientations, nonmanual signals, and movement. These combine to form lexical units which are referred to as signs. These signs are combined with, and governed by, a complete grammar to form utterances. This grammar is neither dependent on, nor derived from, English. This grammar is functionally expressed primarily through the use of space and facial expressions, but may also include other physical elements, such as eye-gaze and body shifting.

Now that the focus of the present study has been introduced and American Sign Language has been described, the next chapter outlines some of the relevant literature that can help inform a discussion of the research questions for this study. Five themes will be explored, in the following order: 1) What does empirical research say about the cognitive processing of ASL, and it is different for deaf and hearing people? 2) Perspectives on teaching and learning ASL as a second language. 3) Defining second language proficiency. 4) Measuring second language proficiency 5) Determining the validity of inferences that we draw from second language tests.

Chapter 2: Literature Review

The following chapter is an overview of literature that looks at American Sign Language and a few select areas of proficiency and language assessment. The body of knowledge that exists for ASL as an L2 is slowly expanding, but at the moment remains small. Empirical studies that investigate learning processes and assessment practices for hearing learners of ASL are scarce. In order to situate ASL within a larger body of L2 assessment literature, I have chosen to divide the literature review into five sections. I first present a brief review of studies that investigate the cognitive processing of ASL. I then outline some of the research that has been done into perspectives on teaching and learning ASL as a second language. Since in this study I am interested in how the proficiency of an L2/M2 learner is determined, I then include a review of literature on what it means to be proficient in a second language, how proficiency can be assessed, and finally ways in which the validity of second-language assessments can be determined.

2.1 Cognitive Processing of ASL

Mayberry and Fischer (1989) assert that “despite surface differences, sign language processing entails two levels of psycholinguistic processing similar to those of speech: form and meaning” (p. 742). On some level this appears to be true. Native signers, in recall tasks, make errors not on the meaning of signs, but on their phonological patterns (Bellugi, Klima, & Siple, 1975), while misconstructions in signs tend to be phonologically based (Newkirk, Klima, Petersen, & Bellugi, 1980). The focus, at least for native signers, appears to be on meaning rather than individual phonological handshapes, as shown by Siple, Fischer, and Bellugi (1977). In fact Klima and Bellugi (1979), in their study of sign language, argued that the underlying language form (universal principles) is

not affected by modality. It was their belief that signed languages encode meaning in the same way as spoken language. What they did not account for was whether the visual and spatial modality would lend itself more favourably to some linguistic encodings over others, or whether there would be different processing constraints or requirements for spoken versus signed languages.

Whereas spoken languages are phonologically sequential in nature, (with a possible acoustic representation mapped sequentially onto lexical items to allow for word recognition) a signed language is less dependent on serial presentation of phonemes. Emmorey and Corina (1990) found that even isolated single signs are more readily identifiable than spoken words, with an average time to identification of 240 milliseconds. They discussed two reasons for this: the first being that the visual nature of a signed language allows more lexical information to be accessed simultaneously, and the second that the cohort model for ASL lexicon retrieval has a much smaller pool of potential lexical candidates to pull from, leading to more rapid recognition and lexical access, with the movement parameter leading most directly to lexical identification.

An additional finding of their research, which looked at Deaf signers, is that there appears to be a significant delay in the performance of those who were exposed to ASL later compared to native signers, leading them to hypothesize that early language experience can affect processes in sign recognition. In an earlier study by Mayberry and Fischer (1989), similar results are reported. Native signers are able to process sign language automatically, while non-native signers are unable to do this. Non-native signers, they found, allocate more attention to phonological identification, while native signers are able to focus more on meaning. This held true even for highly experienced

non-native signers; they are consistently out-performed by native signers. While it is important to note that these findings apply to the late exposure of Deaf signers to an L1, not to ASL as a second language, they still raise interesting possibilities. If the same results hold true for ASL L2 learners, then there seems to be a distinct possibility that they may never achieve native signer fluency, at least at for sign comprehension. This is true for any language, and many people now argue that native fluency is not even necessary, or a realistic aim, in second language learning. However, it would be interesting to discover whether ultimate attainment of proficiency in an L2 is at all affected by modality.

In a later set of experiments, Emmorey (1993) compared hearing and deaf ASL signers' ability to generate visual images and detect mirror reversals in a mental rotation task. The hearing ASL signers performed similarly to deaf ASL signers on both tasks, and both performed better than non-signers, which indicates that the enhanced ability is a consequence of experience with a visual language rather than a product of auditory deprivation. However, the hearing ASL users included participants who had been born to Deaf parents, and had had exposure to ASL since birth. As a result, the hearing participants had been exposed to a signed language during a critical period of language development. At what stage of learning the same abilities would become available to hearing, adult learners of ASL is not known.

Poizner's (1983) study of movement in ASL found that both hearing non-signers and deaf signers perceive the same movements as salient psychophysical properties, but that experience with signing as a formal phonological system allows signers to categorize and prioritize movements differently than non-signers. This, Poizner argues, provides

evidence that experience with a visual-gestural language can affect movement perception, in much the same way that experience with a spoken language effects spoken sound perception. Again, the issue with such a study is that it compares native deaf signers to non-signers, and so provides little useful information on the arc of skill development that might be followed by hearing ASL students.

Building on studies like Poizner's (1983), Morford, Grieve-Smith, MacFarlane, Staley, and Waters (2008) examined the perception of handshape and place of articulation (location) in three groups of signers (Deaf native signers, Deaf non-native signers, and hearing non-native signers) in order to determine whether there is a difference in perception between the three groups. The researchers hypothesized for this study that

deaf non-native and hearing L2 signers would perform similarly to native signers on the identification task, just as signers and non-signers have in past studies...but that deaf non-native and hearing L2 signers would have poorer discrimination of the tested handshape contrasts than native signers...perceptually similar contrasts would show greater differences between groups" (p. 43).

Results were not as expected. It was shown that, irrespective of language background or experience, all participants "revealed discontinuities in their ability to discriminate between phonetic variants of handshape primes" (p. 50). In fact it was hearing L2 signers who displayed the "most uniform discrimination performance across the three perceptual regions that were compared" (p. 50). All of the participants had a similar number of years of experience with ASL, but varied by first age of exposure. Differences in discrimination performance – whether participants' discontinuities in perception occurred in regions on the periphery of a phoneme boundary, between peripheral and central stimulus pairs, or from the category centre – showed that "the earlier in life participants

were exposed to ASL, the more their perception skills are optimized to the linguistically relevant variation in the signed language” (p. 50). In fact, researchers “did not observe a heightening of sensitivity to important contrasts for either handshape or POA [place of articulation] in the native signers. All three groups tested were able to discriminate stimuli that straddled perceptual boundaries with a high rate of accuracy (p. 50). The results of this experiment show that visual language experience may only affect sensitivity to within-category variation, and only on handshape.

They explain the sensitivity of Deaf non-native and hearing L2 signers by postulating that “sensitivity to visual changes in handshape develops from birth in all humans, and is maintained to some degree across the lifespan” (p. 51). They then argue that it is possible that non-native signers (both Deaf and hearing) pay greater attention to phonemic detail in handshape, and this may at some point be “detrimental” to efficient processing of ASL (p. 52). However, based on these results alone, that is only speculation. More experiments would have to be designed and carried out in order to determine when, or if, the presumed “increased processing time” actually exists, and more importantly whether it is significant enough to pose a barrier to smooth face-to-face interaction.

In a paired-comparison discrimination task, Bochner, Christie, Hauser, and Searls (2011) examined ASL learners’ ability to differentiate between contrastive and non-contrastive differences in sign formation. This is something that often proves difficult for learners to do, particularly at the beginning stages of learning ASL. Results from 127 hearing learners at beginning and intermediate levels of proficiency were compared to results from 10 Deaf native signers. Results showed that movement contrasts are the most

difficult, while complex morphology and “Same” contrasts (where utterances in the presented pair had the same meaning) are moderately difficult. Handshape and palm orientation contrasts are relatively easy, and location contrasts are the easiest. The task was able to separate beginners from intermediate and native signers, but found no statistically significant difference in results between intermediate and native signers, perhaps because of the smaller sample size of these two groups. This study has important applications for teaching, learning and testing ASL as an L2. It indicates that “perceptual confusion in sign are not randomly distributed; rather, they are most likely to occur within the categories of movement and complex morphology” (p. 1318). The authors suggest that “the development of ASL curricula and delivery of ASL instruction should strive to facilitate learners’ ability to identify fundamental linguistic contrasts...Moreover, approaches to the assessment of ASL proficiency could benefit from the evaluation of learners’ ability to discriminate linguistic contrasts” (p. 1321). What it left out of the discussion is *how* the results could be developed into a practical system for proficiency assessment or in what order, and through what methods, learners can best acquire the ability to discriminate between linguistic contrasts.

It is unclear precisely what differences exist between Deaf and hearing signers in their perception and production of ASL. As the above studies indicate, there does appear to be some difference, particularly in processing time, between native signers and those who take up ASL or another signed language later in life, as well as differences between novice and experienced signers. In a parallel with spoken language SLA research, it also appears that experienced and novice signers allocate attention to different things. There are important implications for teaching and assessing ASL in being aware of these

differences, and there is ample potential for future in-depth studies that investigate the acquisition patterns of hearing learners.

2.2 Perspectives on Teaching and Learning ASL

2.2.1 Teaching ASL as a Second Language

As discussed in Chapter 1, ASL has become increasingly accepted and popular as a second or ‘foreign’ language at universities in Canada and the United States (Cooper, Reisman, & Watson, 2008; Cooper, Reisman & Watson, 2011; Miller, 2008; Quinto-Pozos, 2011; Jacobowitz, 2005). The growth in the number of classes offered, as well as increasing enrolment in those classes, has been paralleled by an increase in the number of available ASL curriculums and materials used by teachers (Rosen, 2010). Depending on the institution, and the guidelines that exist for the particular teaching context, teachers may use one specific commercially available curriculum or more than one; they may design their own materials, or may pull together material from several sources and combine them into their own, ever-evolving curriculum. Rosen (2010) suggests that the number of teachers who use different curricula is indicative that “they may not have an understanding of the theoretical, empirical, and pedagogical assumptions of the different curricula” (p. 349), arguing that some of the available curricula are based on older theories of pedagogy. He points out that other curricula endorse newer ideas, and that the inconsistencies amongst curricula “raise questions about the teachers’ understanding of the principles of and practices in second-language curriculum development and instructional strategies” (p. 349). It should also be noted that while much of current spoken second language pedagogy follows a content-based or task-based language

teaching approach, there are no content-based instruction or task-based language teaching curricula currently available for ASL.

Schornstein (1995) notes that in her experience as an instructor, following a curriculum which includes a textbook most often leads to students using the text as a vocabulary book. They draw signs from the book, but map them onto English syntax. With this method, there was little-to-no improvement in their communication skills. She then developed a new goal, which was primarily to have students “communicate as naturally as possible following ASL structure while continually expanding their range of vocabulary, which is not limited to what is in the book” (p. 400). For evaluation, Schornstein (1995) requires that students video any work they do outside of class, and that they share their experiences by signing, which is “a much better indication of their growing skills” (p. 405). She describes several elements which are included in her ongoing process of evaluation; students in her class take part in group discussions, formal individual presentations, and video responses to presentations. She refers often to the idea of fluency, and to the importance of providing students with opportunities to demonstrate their ASL skills in both formal and natural settings. However, she does not indicate what linguistic aspects are included in fluency or proficiency evaluations.

In terms of teacher training, Jacobowitz (2005) believes that the growing demand for ASL classes and ASL teachers means that the need for ASL teacher preparation programs (TPPs) has increased accordingly. She compares three ASL TPP sites in the U.S., examining them in terms of the program, their curriculum development, admission and enrolment requirements, and the courses. She found that the curricula of the three ASL TPPs support the four priorities of (American) national standards for teaching.

Additionally, each program used some form of ASL proficiency interview in order to assess the ASL level of students, and students could also demonstrate their ASL skills through letters of recommendation, ASL course work, and informal interviews with program heads.

Jacobowitz (2005) recommends increased cooperation between teacher preparation programs and the American Sign Language Teacher's Association (ASLTA) in order to create national ASL TPP program standards. The standards are similar, she suggests, to those set out by the National Council for Accreditation of Teacher Education (NCATE). Additionally TPPs and the ASLTA should work together to create a national ASL proficiency interview. There are three main potential problems with these suggestions. The first is: How accessible would these TPPs be to potential Deaf instructors of ASL? Jacobowitz recommends a standard be put in place in which teachers must possess, or be working towards, a terminal degree. There remain many problems in North America regarding deaf education (Cerney, 2007; Edwards, 2012; Lane, 1999), and the implementation of such standards in regards to ASL without first developing better access to education for deaf people themselves could be seen as further oppression of an already oppressed minority. Would these standards lessen or increase instances of what Lane, Hoffmeister, and Bahan (1999) refers to as audism? Second, Jacobowitz mainly addresses standards that are applicable to teacher preparation programs which focus on future teaching in levels K-12, and does not clearly address the different expectations placed on instructors working at the university level. Third, while the idea of creating a national standard for ASL proficiency interviews seems, on the surface, like a

positive development, how would such a standard be devised? What form of ASL would be used as the “standard”, and who would judge it?

Kanda and Fleischer (1988) take a broader view of what qualifies someone to teach ASL, and conclude that the necessary qualifications are

Linguistic and cultural competence undergirded by interaction with members of the Deaf community and accompanied by proper attitudinal characteristics are prerequisite. In addition, an ASL teacher should be educated, demonstrating knowledge and application of educational and pedagogical principles along with formal study of the language” (p. 193).

While acknowledging the importance of the above qualifications, Kanda and Fleischer (1988) also argue that “ASL teachers should recognize that they need first to be human beings; then to be teachers, and third, to be teachers of American Sign Language” (p. 1988). This results in a more holistic and humanistic view of ASL teaching, as opposed to one that focuses solely on certification.

Another standards-based approach to teaching ASL has been developed by Ashton et al. (2014). They outline standards for what should be taught throughout ASL in levels, beginning with kindergarten and continuing through to post-secondary. These were developed through cooperation between the ASLTA, the National Consortium of Interpreter Education Centres (NCIEC), and with financial support from the American Council of the Teaching of Foreign Languages (ACTFL). Within each level are five goals: communication, cultures, connections, comparisons, and communities; which “reflect the framework of communicative modes established by the ACTFL and incorporate the 5 C’s of foreign language instruction” (p. 10). Sample progress indicators included in the document represent a spiral approach to teaching ASL; concepts are revisited more broadly and deeply as students progress to higher levels of proficiency.

The creators of the standards note that the standards should not be taken as a curriculum, but as a guideline which “should be applied with flexibility in mind” (p. 10). The authors also recognize that students will enter into learning ASL at different points in time. While the standards build upon concepts that are introduced first to young learners (kindergarten) and are “used to demonstrate the typical progress of students in well-articulated ASL programs” (p. 10), the authors also recognize that students entering ASL studies for the first time at the secondary or post-secondary level “will need progress indicators more appropriate to novice learners” (p. 10) and that “a college-level course will require more extensive materials than those presented to an elementary or high school student, even though they may be working on the same linguistic or cultural topic” (p. 11).

The type of ASL learning trajectory proposed by Ashton et al. reflects an interest in creating a well-developed learning arc. According to Knigga (2012), long-term successful teaching (and learning) of ASL depends not only on teacher qualifications or curriculum, but also on having a clearly articulated ASL program. Drawing on both Lange’s (1982) and Pasesani and Barrette’s (2005) definitions of articulation, she determines that a strongly articulated ASL program would have both horizontal (across classes at the same level) and vertical (the hierarchal flow from level to level without overlap) continuity, where students achieve fluency in a reasonable amount of time, and in which the individual needs of the learner are considered and where there is “a cohesive relationship within instruction, content, assessment, and goals” (Knigga, 2012, p. 84).

Based on the review of current research undertaken for this thesis, there do not appear to be any longitudinal or cross-sectional studies looking at the effects and eventual

outcomes of standardization of ASL levels, teacher qualifications, or pedagogical approaches. What is currently known is that deaf and hearing instructors approach the teaching of ASL differently. In an examination of the perception of competencies important to teaching ASL, Newell (1995) found that deaf instructors rate educational foundations, such as knowledge of communication process theories, or issues affecting education in general, as significantly more important than hearing instructors. Conversely, hearing instructors put significantly more emphasis on the importance of competency in ASL structure. It is unclear at this time which approach to teaching has the most potential to provide a learning situation for hearing students that is linguistically strong, pedagogically sound, and provides the necessary level of interpersonal support to lead to successful learners. However, those two qualifiers - linguistically strong and pedagogically sound - are two key features that any language program should strive for.

2.2.2 Learning ASL as a Second Language

Just as teachers will face challenges in instructing their students in ASL, so will students face challenges in learning a visual language. These challenges may go beyond linguistic issues. Kemp (1998) distils the challenges of learning ASL into what he believes are the five key factors. Four of these - social dominance patterns and attitude, language shock, congruence, and motivation – deal with cultural, behavioural or attitudinal issues that may pose difficulties for students. Only one - properties of L1 grammar transferred into L2 grammar – considers linguistic difficulties. Kemp assumes that there are more psycho-social barriers to learning ASL than there are linguistic ones. ASL students do face challenges at a social and cultural level, some of which are unique to learning a visual language. Students, when entering an ASL class or program for the

first time, may assume that ASL is easy because it is sometimes believed to be a purely iconic language, or that it is “English on the hands”. While it is true that iconicity is more prevalent in signed than in spoken language (Taub, 2001) this does not necessarily mean that ASL is a completely iconic language. If this were true, “students would be able to understand signs without much instruction” (Jacobowitz, 2005, p. 77). Nevertheless, some students do take ASL for credit with the misconception that it will be an easy A, or at least easier than learning a spoken language (Jacobowitz, 2005; Miller, 2008; Quinto-Pozos, 2011) and are surprised when they find the class difficult.

Quinto-Pozos (2011) goes into greater detail describing more of the challenges that hearing, adult learners will encounter when learning ASL. Many of these include difficulties that arise from communicating in a new modality which requires them to “learn how to use their hands, arms, and bodies in new ways and having to acquire a manual phonology as an adult” (p. 144). As Pichler (2011) points out, the fact that many ASL L2 (M2) learners are adults, who would arguably have a fully developed motor system, will not preclude them from making errors in sign movements or handshapes, since “Most adults, faced with unfamiliar motor tasks, such as skiing or signing, execute them awkwardly until they have enough practice to maintain effective motor control in these new domains” (p. 101). This can be compared to learning a new spoken language, where speech and articulation rates increase and correlate with fluency (Wood, 2001). Lupton and Zelaznik (1990) found that the movement of signs produced by novice hearing ASL learners improves with practice over time. Muscles, they observe, have spring-like properties which generate a relationship between movement and velocity which “will show limit-cycle behaviour, in that the phase portrait will resemble a circle

or an ellipse” (p. 155). Specifically, they argue that “The emergence of limit-cycle behavior after several weeks of practice for the two-handed signs suggest that the development of coordination requires the organization of these processes more than it calls forth already existing processes” (p. 170). Over the course of a semester, sign movements become faster as well as more constrained in movement as students become familiar with the movements and are able to automatize them.

How L2 learners acquire the necessary skills in order to comprehend and control the communicative function of (signed) utterances in interaction with others appears to be relatively unexplored research territory. The pragmatics of sign language discourse has so far only been sparsely investigated (Metzger & Bahan, 2001). One study (Roush, 1999) explores how signers use both manual and nonmanual signs to “mitigate what might otherwise be construed as direct or even rude utterances” (Meztger & Bahan, 2001, p. 122). L2 learners of a visual language must be able to acquire not only the vocabulary and literal meanings (along with rudimentary grammar) necessary for basic communication, but at higher levels must also be able to process and produce the non-manual modifiers that mitigate, emphasize, clarify, or otherwise alter the intended meaning of an utterance. While this is necessary for learners taking up a second spoken language as well, what is distinct in ASL discourse as opposed to spoken language is that often these signals are produced non-manually, i.e. on the face or through subtle shifts in the body. How L2 learners of ASL can be supported in acquiring the pragmatic skill necessary for fluent communication is a significant issue since, as McDermid (2012) points out, there has been “ongoing concern in the field of [ASL] education and interpretation about the ability of hearing adults to gain mastery of ASL” (p. 15).

Learning to process language visually can be mentally and physically draining for hearing students, both at the beginning and at later stages of learning. As McKee and McKee (1992) note, one of the first tasks that hearing learners face “is learning to attend visually to linguistic information that is coded in a form for which they have no perceptual “schema”” (p. 131). They also investigate the perceptions of hearing students and ASL teachers (both Deaf and hearing) in order to identify the linguistic difficulties which are most often expressed by each group in relation to learning and teaching ASL. At the top of the list for both students and teachers are: coordination and articulation skills; grammar and syntax; and facial expression and nonmanual signals. While students “perceive ASL syntax and grammar to be one of their biggest weaknesses” (p. 135) they also report that “the expressive use of grammatical ASL facial expression and other nonmanual signals (NMS) is one of the most problematic aspects of learning ASL” (p. 138). In an exploration of students’ written comments relating to these problems the authors found that “difficulty is ascribed mainly to inhibitions about the use of face and body in communication, reflecting a clash between hearing and Deaf cultural norms for appropriate communicative behaviours” (p. 138).

There are other communicative norms for ASL that may pose problems for students. Students are often involved in ASL classes where they are expected to follow a ‘no-voice’ rule, which is “in line with general norms of Deaf culture and are often espoused by Deaf instructors” (Quinto-Pozos, 2011, p. 146). Proponents of this rule claim that it is motivated by both linguistic and cultural factors. Tevenal and Villanueva (2009) found that when ASL and spoken English are used together (what is referred to as SimCom), it produces unequal messages which drastically impacts the integrity of both

languages. Often signs become merely a shadow or manual code for the spoken language; visual grammar is dropped, and signs are produced following English syntax. Thus, when people attempt to sign and speak at the same time, what they produce is not ASL (or any *language*) at all. Not voicing is also believed to encourage students to make the most of their class time to become comfortable communicating visually, while at the same time introducing an important norm of Deaf culture; there are many Deaf people, whether they are teachers or not, who do not use their voices. In the classroom many teachers, like Schornstein (1995), state that they want students to learn respect for them as Deaf people, and respect for the Deaf community in general, by not talking in class (p. 400). However, students often find the “no-voice” rule difficult to maintain. Kemp (1998) rather harshly claims that the tendency among hearing students to speak or mouth English while signing provides them with a sense of “narcissistic gratification” (p. 258). This seems to suggest active rebellion or rudeness on the part of hearing students, but it is also possible that the behaviour stems from complete unfamiliarity with, and knowledge of, Deaf cultural norms and practices. It also stems from the process of learning a second language. Second language learners already have a prior language system in place, through which they develop inter-language knowledge. For ASL, cultural expectations may clash with language learning needs. Whether a complete ban on speaking is actually beneficial to students’ learning is questionable.

This is not to say that acknowledging cultural differences is not necessary. Deaf culture is not the same as hearing culture, and it is important that ASL learners are aware of that. Kemp (1998) holds the belief that “Deaf culture is different from any other culture. The problem with the perception of ASL, in my opinion, is that

Americans...probably think Deaf people are much the same as hearing people except that they don't hear" (p. 143). This attitude, that 'hearing people just don't understand' is not uncommon amongst some members of the Deaf community. In instructing students in the cultural components of ASL it is the job of the teacher to introduce them to Deaf culture in a way that is open and accessible. Personal and political stances towards hearing learners that are overtly negative may alienate students, and only serve to keep hearing and Deaf communities separate. Students may be surprised at first by the differences they encounter between Deaf and hearing people. Social interactions with Deaf people can be quite different from what many hearing people are accustomed to. There is much stronger and longer-maintained eye-contact; communication is more direct and up front, with fewer instances of hedging; and Deaf people are often more open in the amount of information, including what some hearing people might consider personal information, they share about themselves and others (Miller, 2008; Schornstein, 2005). This can lead to a learning situation which is potentially "totally foreign to most hearing people. This can cause a sense of discomfort, often described as culture shock, for individuals who are not used to these kinds of interactions" (Miller, 2008, p. 232). Learning about these differences, and learning how to communicate within Deaf culture, is essential when learning ASL. It may be necessary for teachers to find a balance between accommodating learners' linguistic needs and presenting or following Deaf communication norms.

Studying and learning a second language can be a valuable experience. Taking up a second language allows learners

to understand language as a structured system. Years of instruction in English grammar may never bring this point home to a student nearly so quickly as intense exposure...to an alternate system, one which allows you to put the same semantic messages into very differently shaped packages, not just with different

words but with different grammatical organization. ASL serves this purpose admirably (Chapin, 1988, p. 111).

As shown in the above discussion, there are challenges faced by both students and teachers of ASL. Teachers and students will approach these challenges from different backgrounds, experience levels, and with different beliefs. This is true of any language learning situation, but ASL faces some unique challenges. It presents linguistic information visually and spatially. It is a language that many students have no prior exposure to, and the modality difference between speaking and signing can produce new stressors on students. There may be resistance or mistrust of hearing students, or even hearing teachers, from the Deaf community. There is also a noticeable lack of resources which apply to teaching a signed language to hearing adult learners. Since ASL is still a relatively ‘young’ language there is sometimes not enough information available to teachers on best practices for teaching, and what information does exist does not always agree.

2.3 Defining Second Language Proficiency

Second language proficiency is the ability of a non-native speaker to use or function in a target language. However, there is little consistency between theorists, or even second language testing organizations, as to what definitively constitutes L2 proficiency. Rather it can be viewed as the conceptual framework for measuring a learner’s second language abilities. Beginning in the 1980’s, a common way of discussing second language proficiency was to describe it as a form of communicative competence, where L2 proficiency develops through interactions (with native speakers, with other language learners, or with texts) which are embedded in particular or meaningful contexts. Proficiency then addresses “broad task-universals, such as the

ability to manipulate language in order to give information of a personal nature, to solve some of the basic tasks of life in a given culture, the ability to narrate and describe in different time frames, and the ability to support in detail an opinion in an abstract topic” (Byrnes, 1987, p. 45).

Ellis and Barkhuizen (2005), in looking at learner language, divide it into three constructs: Accuracy, complexity, and fluency. By fluency, they do not mean a high level of language proficiency, but rather “the production of language in real time without undue pausing or hesitation. Fluency occurs when learners prioritize meaning over form in order to get a task done” (p. 139). These constructs can be measured individually, or taken together to provide a broader view of a learner’s proficiency level.

Canale and Swaine (1980), build on the work of Chomsky (1965), particularly his distinction between “*competence* (the speaker-hearer’s knowledge of his language) and *performance* (the actual use of language in concrete situations” (p. 4) as well as his (1980) distinction between grammatical competence and pragmatic competence. They propose a minimal framework of communicative competence which consists of grammatical, sociolinguistic, and strategic competences (p. 28), and which they assume

interacts (in as yet unspecified ways) with a theory of human action and with other systems of human knowledge (e.g. world knowledge). We assume further that communicative competence, or more precisely its interaction with other systems of knowledge, is observable indirectly in actual communicative performance (p. 29).

Thus, testing of language proficiency, via a communicative language program, must be “devoted not only to what the learner knows about the second language and about how to use it (competence) but also to what extent the learner is able to actually demonstrate this knowledge in a meaningful communicative situation (performance)” (p. 34). Canale and

Swaine do not, however, establish evaluation criteria or proficiency levels. They instead refer readers to the work of Carroll (1978), Munby (1978), and others.

Bachman (1990) describes language proficiency in terms of communicative competence, which he refers to as communicative language ability. His description is consistent with the work of Canale and Swain (1980), as well as that of Savignon (1983) “in that it recognizes that the ability to use language communicatively involves both knowledge of, or competence in the language, and the capacity for implementing, or using this competence” (p. 81). He proposes a framework for communicative language ability which consists of three components: language competence, strategic competence, and psychophysiological mechanisms. Language competence is “a set of specific knowledge components that are utilized in communication via language” (p. 84) and strategic competence is used to “characterize the mental capacity for implementing the components of language competence in contextualized communicative language use” (p. 84) while psychophysiological mechanisms refer to “the neurological and psychophysical processes involved in the actual execution of language as a physical phenomenon” (p. 84). Bachman further refines these categories. For communicative competence there are several different groupings of sub-skills clustered under the categories of: Organizational Competence (morphology, syntax, vocabulary, cohesion and organization); and Pragmatic Competence, which includes elements of sociolinguistic competence, and also abilities which are “related to the functions that are performed through language use” (p. 86). Strategic competence is divided into three components: Assessment, planning, and execution (p. 101) which exist in a constant relationship “so that communicative language use is not characterized simply by the production or interpretation of texts, but

by the relationship that obtains between a text and the context in which it occurs” (p. 102). Finally, by psychophysiological mechanisms Bachman means the neurological and physiological processes which are utilized in language execution, either receptive or productive.

Hulstijn (2011) argues that while Canale and Swain (1980), and Bachman and Palmer (1996), “were right in including other than purely linguistic competences in the construct of LP...these competences ought to be put in LP’s periphery, for the simple reason that although strategic and metalinguistic competences cannot exist without linguistic competences, linguistic competences may exist without strategic and metalinguistic competences” (p. 239). His definition of language proficiency is more specific, and separates core components of proficiency from peripheral ones as follows:

Language proficiency (LP) is the extent to which an individual possess the linguistic cognition necessary to function in a given communicative situation in a given modality...Linguistic cognition is the combination of the representation of linguistic information (knowledge of form-meaning mappings) and the ease with which linguistic information can be processed (skill). Form-meaning mappings pertain to both the literal and pragmatic meaning of forms (in decontextualized and socially-situated language use, respectively). Linguistic cognition in the phonetic-phonological, morphological, morphosyntactic, and lexical domains forms the center of LP (core components). LP may comprise peripheral components of a less-linguistic or non-linguistic nature, such as strategic or metacognitive abilities related to performing listening, speaking, reading, or writing tasks. (p. 242).

Verhoeven & Vermeer (1992) propose a model for L2 proficiency which they believe “can best be regarded as multidimensional” (p. 170). Proficiency, which they also term communicative competence, is observed indirectly as communicative behaviour. According to their study, it is “clear that some features of communicative behaviour are related to grammatical proficiency, while others are not” (p. 171). They find a correlation

between grammatical proficiency and fluency, as well as moderate correlations between grammatical proficiency and illocutionary and strategic proficiency, but no relationship between grammatical proficiency and sociolinguistic proficiency (p. 171). Based on their findings, they propose a model of communicative competence where second language proficiency is divided into five categories, some of which encompass smaller sub-categories: Grammatical Competence (Phonology, Lexicon, Syntax), Discourse Fluency, Sociolinguistic Competence (Cultural Rules, Expressiveness), Illocutionary Force, and Strategic Competence (Planning, Execution).

Despite the existence of various proficiency models, and testing of these models, there is still no singularly agreed-upon definition when it comes to second language proficiency. There is, as of yet, no general theory of SLA. In fact, “At best, SLA research has reliably and generally explained some few specific phenomena, for example certain selected syntactical or morphological constructions, and even these examples are arguable” (Klein, 1998, p. 529). Even if we assume that learners acquire a second language by following a universal sequence, questions still remain regarding the order of SLA. Is the order particular to the language being learned? How does the learner’s own approach to the language influence their proficiency development? What about factors that may stem from the learner’s L1? How does the learning context contribute to proficiency? On the latter point, Ellis (1985) writes that “It has become generally accepted that the human language faculty is a potent force in language acquisition. But internal processing and innate mechanisms are not the full story. It is also necessary to consider how the linguistic input contributes to the process of acquisition” (p. 73). As Byrnes (1987) points out “evidence from immersion projects and from research

comparing naturalistic with instructed learners seems to point increasingly to the crucial importance of formal instruction if rate of acquisition and ultimate level of attainment are considered” (p. 48).

Clearly, second language proficiency is a complicated phenomenon. Why even be concerned with creating an operational definition of proficiency? Some theorists, such as Pienemann (1985) are not, and argue that it is an amorphous construct. As Long, Gor, and Jackson (2012) point out, several theorists have made similar arguments: proficiency lacks operationalization “and quite possibly, operationalizability” (p. 101). This is in opposition to learners’ current language-learning stage and subsequent progress which are “psycholinguistically meaningful and objectively measureable” (Long, Gor, & Jackson, 2012, p. 101). If this is the case, then second language proficiency has to be thought of in more general terms, perhaps best conceptualized as a set of threshold levels, descriptors, or bands that can be used to separate low, mid, and high proficiency learners by what they can be expected to be able to do with the L2 at each of these stages. There can be limitations to the reliability and validity of proficiency scales, particularly when a reductive scale is used; the quality of the information measured with this type of scale may lead to problems with the validity of the interpretations. However, one of the main reasons to use proficiency scales is to increase reliability. Global ratings of L2 abilities are used around the world, by employers from the public or private sector as well as by academic gatekeepers (Long, Gor, & Jackson, 2012). While there may be no theoretical basis for the distinctions drawn between levels of language proficiency Hulstijn (2012) maintains that “in educational practice it is quite common, perhaps even unavoidable, to speak of LP levels, such as beginning, intermediate and advanced levels, or to adopt the

level labels used by companies that produce language-proficiency tests” (p. 429). The acceptability of proficiency levels or proficiency frameworks “seems to rely primarily on the authority of the scholars involved in their definition, or on the political status of the bodies that control and promote them” (De Jong, 1990, p.72). However, this does not mean that ‘anything goes’. As North (1995) writes

If what proficiency scales and profile grids do – and there is no hard and fast line between the two – is to map behaviours (say this thing is *more difficult* than that thing, assess progress, report achievement), then there is at least an argument that they ought to take account of theories of scaling, and of experience with behavioural scaling (p. 446).

In the context of SLA within a university language program, what exactly will constitute proficiency for a particular language will be dependent on a number of factors, including (but not limited to) method of instruction, ultimate goal of the program or class, and whether an existing framework for language proficiency is already in place.

These ongoing issues regarding L2 proficiency, and L2 development exist in all second language learning contexts, regardless of the language. Second language research, even now, is “still grappling with the fundamental questions of how foreign and second language proficiency develops, and therefore what underlies such development” (Alderson, 2005, p. 1). Nevertheless, if second language students are going to be assessed, graded, and/or sorted based on their language ability, then proficiency, as a construct, needs to be defined in some way. Whether this is done in broader or more specific terms will depend on the goal of the language-learning context, program, or curriculum. It is not a simple matter, but it is crucial in order to “know what is required of the individual student, and more broadly, of the general student population to move from one proficiency level to another in a given language” (Long, Gor, & Jackson, 2012, p.

100). However, since it presents linguistic information using a visual modality, ASL L2 proficiency has to account not just for language proficiency generally, but more specifically how that proficiency develops manually and visually in learners who often have no prior experience with a visual language. This leads directly into additional considerations for how language proficiency, and specifically visual language proficiency, can be determined through testing, which will be discussed in the next section.

2.4 Assessing Second Language Proficiency

Examining how we assess second language proficiency, through an examination of the assessments themselves, is important because proficiency assessment “provides for an interactive, adaptive testing instrument that seeks to preserve crucial aspects of natural language use” (Byrnes, 1987, p. 48). Looking at proficiency tests can provide insight into at which point(s) learning can be enhanced, and to what extent the enhancement can be realized (p. 48). When describing the basis for L2 assessments, Bachman (1990) states that “if we are to develop and use language tests appropriately, for the purposes for which they are intended, we must base them on clear definitions of both the abilities we wish to measure and the means by which we observe and measure these abilities” (p. 81).

Bachman and Palmer (1996) divide language tests into three stages: design, operationalization, and administration. During the design stage the test is planned, taking into consideration the purpose of the test, tasks in the target language use domain, characteristics of the test takers, a definition of the construct to be measured, a plan to evaluate test usefulness, and the need to identify resources and plan for their allocation and management (Bachman & Palmer, 1996). Operationalization then involves

“developing test task specifications for the types of test tasks to be included in the test, and a blueprint that describes how test tasks will be organized to form actual tests” (Bachman & Palmer, 1996, p. 90), as well as a set of specifications that outline the scoring method.

As part of the test design stage, the construct that is to be investigated must be defined. If we desire to develop a test of proficiency in ASL, we must first decide on a theoretical or conceptual definition of this construct. This is an abstract sense of what the construct of proficiency means. Before a test can be developed to measure a student’s proficiency in a second language, there needs to be some measure (even if it is a general measure) of what it means to be proficient in that language, as different understandings of proficiency can lead to the development of different assessments. Language tests need not be based on a specific definition of proficiency, since no adequate one currently exists, but it is important to recognize, at the very least, “the need to base the development and use of language tests on a theory of language proficiency” (Bachman, 1990, p. 81).

Once that has been established, we can then move on to the task of operationalizing the construct. This is more concrete, and involves stating what discrete units are necessary to examine in order to obtain a measure of our construct (proficiency), and specifying what procedures we will use to measure it. The construct is functionally operationalized through the creation of test tasks. Bachman and Palmer (1996) write that “In language testing, our theoretical construct definitions can be derived from a theory of language ability, a syllabus specification, or both” (p. 89).

Zhang (2010) notes that proficiency classification in a second language is “one of the major, if not the only, reasons for which many language learners actually take language tests” (p. 119). Zhang writes in reference to large-scale standardized tests, such as the TEOFL or IELTS, and investigates the accuracy of proficiency classification under different measurement models, primarily classical test theory (CTT) and item response theory (IRT) where classification is based on a fixed cut-off score. He admits that “for tests with a limited number of items...proficiency classification is more challenging” (p. 138). His investigation looked only at tests for spoken (as opposed to signed) languages, which “measure general language competence by assessing skills in the key curricular areas such as listening, reading, grammar, vocabulary, and writing” (p. 135). In the context of a one-semester language class, a large, standardized test may not be desirable or practical. Determining proficiency through assessment is also going to be more of a challenge for ASL because the language has no written form. This gives those interested in developing ASL proficiency tests less to work with, and means that the weighting of items through which proficiency can be observed must be rebalanced to include only productive (signing) and receptive (comprehending what someone else signs) skills. There is no visual language equivalent for reading or writing. This has led to the development of ASL proficiency tests that often take the form of interviews. Proficiency interviews are considered direct or semi-direct, as the channel of stimulus and response both match the language ability channel (Bachman & Clark, 1987). Interviews are also examples of highly ‘functional’ language tests, since “the stimulus and the response involve the performance of a variety of language functions” (Bachman & Clark, 1987, p. 27).

The scale of measurement used for determining a score on a proficiency test should be criterion-referenced, according to Bachman and Clark (1987). This is because the alternative, norm-referenced scores, are “not suitable for use in either formative program evaluation or in comparative program evaluation, since it fails to reflect the degree of mastery of specific objectives and since both the score norms and then content may be inappropriate to either or both groups that are being compared” (p. 27). ASL is no different in terms of proficiency assessment than any other instructed foreign language; the measurement of what a student can do in ASL should be based, as Bachman and Clark (1987) suggest, on specific criteria. In a program of language learning, these criteria will be dependent on the level of instruction where goals for proficiency have been outlined by taking into consideration: general sequences of L2 acquisition; the mode of instruction; duration of course; and outcomes of a) the particular course, and b) the larger goals of the overall program of study. The main difference for ASL proficiency, then, will stem from the channel or mode of communication. A visual language like ASL has to account not just for the acquisition of lexical items and grammatical forms, but for the production and comprehension of these things as they exist through spatial use and physical movement as well.

Haug (2005) provides an overview of numerous sign language assessments, however, of the tests Haug discusses, almost all have been developed to assess the ASL skill of young, deaf signers. Some tests are appropriate for signers at a variety of ages, but most are intended to be used with children or young signers. In Haug’s review, several problems with these assessments are noted: The first is that “only a few of the tests described here report data on validity” (p. 63); the second is that “Few of the tests

discussed here are standardized” (p. 64); the third is that “only a small number of tests are available...Most of the other tests are not publicly available” (p. 65). For example, the American Sign Language Proficiency Assessment (ASL-PA) (Maller, Singleton, Supalla, and Wix, 1999) was developed to evaluate the ASL proficiency of deaf children between the ages of 6-12. While it is criterion referenced and the researchers claim that “The psychometric properties of the ASL-PA have been studied extensively” (Maller, Singleton, Supalla, & Wix, p. 264), it has only been used for research purposes, and the data presented in the study comes from only 80 participants, “thus, no large sample norms are available” (p. 264). Haug (2005) identifies several weaknesses with the test, including the limited age range, and also notes that the test “has been developed in the context of a research project and is not yet publicly available” (p. 67). Two of the authors of the ASL-PA, Supalla and Singleton, were also involved in a previous ASL assessment project for the *The test battery for American Sign Language morphology and syntax*. (Supalla et al., in press). In their review of available ASL assessment instruments, Maller, Singleton, Supalla and Wix (1999) describe this test as “One of the most comprehensive ASL tests available” (p. 250). The problem is that the publication has been ‘in press’ since 1995; to date it has not been published and is not publicly available, so it is difficult to know if the authors’ comments on the test are at all accurate.

A test developed for a particular target group must be properly adapted if it is possible to adapt it, before it can be used with a new group. If it is not possible, then a new test must be created. A deaf user of ASL is not the same as a hearing one; a hearing, adult learner cannot be expected to function, learn, or be optimally tested using the same methods that are applied to deaf learners or users of ASL. It is important to include these

considerations in the testing process because “the technical quality of tests is easily undermined if language constructs are not operationally defined as part of the process of test design and development” (Fairbairn & Fox, 2009, p. 11). So, while there are several tests that have been developed for deaf users of signed languages, few have been designed with hearing learners in mind.

When spoken language proficiency assessment is included, there is vastly more information and research available to draw from on. It is helpful to examine the construct of proficiency in spoken language research, as some facets of assessment are applicable (with some restructuring) to ASL. For example, Iwashita, Brown, McNamara and O’Hagan (2008) investigate spoken language test performances, which represent five tasks and five proficiency levels, in order to determine how distinct the proficiency levels actually are. They treat the performances as empirical data, and analyze them using measures of grammatical accuracy and complexity, vocabulary, pronunciation, and fluency in order to see which features have the strongest impact in determining overall learner fluency and the nature of proficiency at different levels of achievement. Researchers derived three broad conceptual categories of performance by going through the think-aloud comments made by raters as they listened and re-listened to speaker performances. These are Linguistic Resources, Phonology, and Fluency (p. 30). Each category has specific features identified within it. Linguistic Resources includes: grammatical accuracy, grammatical complexity, and vocabulary. Phonology includes: pronunciation, intonation, and rhythm. Fluency was regarded as a single feature, but was analyzed in several ways, such as filled pauses (‘ums’ and ‘ers’), unfilled pauses, repair,

total pausing time, speech rate, and mean length of run. This research suggests specific ways of categorizing aspects of proficiency that could potentially be very useful for ASL.

Hulstijn (2012) proposes a two-phase test for proficiency: the first being a discrete-point test, and the second an integrative skills test, where “the blessings of the communicative movement come to bear in the second phase of the exam” (p. 245). In his view “The validity of the integrative-skill testing in Phase 2 will increase, when, in Phase 1, it has first been ascertained that candidates have sufficient linguistic knowledge at their disposal” (p. 245). Hulstijn (2012) likens this two-phase format to a driving test, where “Just like a driving license should not be issued without the candidate having passed the practice test on the road, a L2P certificate should not be issued without an integrated skill test” (p. 245).

As the brief discussion above shows, what will be included in operationalizing proficiency, for the purposes of assessment, and how the assessment will be set up, can vary considerably. As Davies (1990) writes, “Assessing language produces complications regarding what is to be tested...and how the testing is to be done” (p. 9). Hulstijn (2012) adds that language proficiency assessment “not ‘only’ constitutes a methodological issue, but forces researchers to reflect on what LP actually is, the association of which with the linguistic feature under study they seek to establish” (p. 426).

What does all of this mean for ASL testing? In a study of the structure and content of sign language programs at higher education institutions in the U.S., Cooper, Reisman and Watson (2011) found that methods for assessing student progress tend to fall into six categories: written exams, where the student translates the teacher’s signs; written exams using fixed choices; live expressive presentations; videotaped presentations; out-of-class

assignments or papers; and in-class assignments or activities. Despite ranking third as the most-preferred method of assessment (after live signing and video-taped signing) the use of fixed-choice written exams is on the rise, having increased by 17 percent compared to 2004. However, this research was comprised of a broad survey investigating content of sign language programs, and so did not cover aspects relating to the validity and reliability of the assessment methods used. At the most basic level, before the administration of an L2 test, then, there must be at least two things in place: definitions of the abilities, and measurement criteria on which to evaluate them. Next we will consider how the validity of a test can be determined, what validity means, and why it is crucial for language tests.

2.5 Validity

Spolsky (1968) states that “The central problem of foreign-language testing, as of all testing, is validity” (p. 94). Validity looks at what a test is actually measuring, and whether it measures what it purports to measure (Lado, 1961; Cronbach, 1971). Weir (2005) defines validity as “the extent to which a test can be shown to produce data, i.e., test scores, which are an accurate representation of a candidate’s level of language knowledge or skills” (p. 12). Even if the construct that we wish to measure is known, it is still necessary to prove, and perhaps re-prove, that a test, and the way in which we use test scores, are valid. To this end, Messick’s (1989a) framework of validity remains “The most influential current theory of validity of relevance to language testing” (McNamara & Roever, 2006, p.12). His main concern is for the interpretation and use of tests scores, and poses the basic question for validity as follows: “To what degree – if at all – on the basis of evidence and rationales, should test score be interpreted and used in the manner

proposed?” (Messick, 1989b, 5). Messick’s stance on test validity also acknowledges that how and what we choose to measure is reflective of particular values, and that testers need to be aware of, and concerned for, the educational and social effects that tests can produce. He writes that “Hence, what is to be validated is not the test or observation device as such but the inferences derived from test scores or other indicators – inferences about score meaning or interpretation and about the implications for action that the interpretation entails” (p. 5).

Determining the validity of a second language test is not a necessarily a simple process. Test validity is multi-faceted, and must consider test content, criteria, and consequences (Messick, 1995), along with theory-based and scoring validity (Weir, 2005). No one facet alone “is sufficient to demonstrate the validity of a particular interpretation or use of test scores” (Bachman, 1990, p. 237). Validity is also not an all-or-nothing scenario; it is a relative concept (Weir, 2005) or a matter of degree (Messick, 1989b). Validity can be imagined as a spectrum, with several influencing factors, that will result in an assessment being either more or less valid.

Stevenson (1981) cautions against the dangers of allotting too much praise to a test based only on seemingly high face validity. Face validity “can be considered to be appearance of validity in the eyes of the metrically-naïve observer” (p. 41). Dependence on face validity as the sole validity measurement may occur in cases where the test used is a type of ‘direct test’, such as interview-style assessments. The test appears to be valid because it is thought to more directly access students’ language ability. This does not, by extension, make them automatically more valid (Bachman & Savignon, 1986, p. 382). While direct tests may be popular (Stevenson, 1981), the problem remains that

competence cannot be observed in a direct way; it is always manifested through an instance of language performance (Carroll, 1968). Believing that an interview-style test, or any particular test “is probably valid is not the point, at least not in the hard-nosed tradition represented by the spirit of validation” (Stevenson, 1981, p. 39). There must be an operationalized construct which the test can be shown to accurately and reliably measure.

Bachman and Palmer (1996) write that “If we want to use the scores from language tests to make inferences about individuals’ language ability, and possibly to make various types of decisions, we must be able to demonstrate how performance on that language test is related to language use in specific situations other than the language test itself” (p. 10). To this end, they stress that quality assurance in test development and test use is directly related the ‘usefulness’ of the test. Test usefulness is an amalgam of six different qualities: reliability, construct validity, authenticity, interactiveness, impact, and practicality. Bachman and Palmer (1996) state that “in order to be useful, any given language test must be developed with a specific purpose, a particular group of test-takers and a specific language use domain (i.e. situation or context in which the test taker will be using the language outside of the test itself) in mind” (p. 18). In subsequent work Bachman and Palmer (2010) describe how to connect scales of language ability to language performance in situations where language use is situated in discourse. Different aspects of language ability are rated separately, but all are tied on the rating scale to the construct that is being assessed. Bachman and Palmer provide a sample analytic scale (p. 345). This sample scale both defines the construct and provides performance criteria, and for each scale level there is a description which includes information about both range

and accuracy. Kline (1993) claims that “A test is said to be valid if it measures what it claims to measure” (p. 15). Test validity could include a number of types of validity (face validity, concurrent validity, predictive or content validity), but for language tests Bachman and Palmer are concerned with construct validity. Kline (1993) notes that the meaning of construct is synonymous with concept. He writes that

In the sciences constructs are frequently the object of investigation or study. A good example of a construct is the notion of species. It makes good sense to investigate this construct and to see how it may be defined such that different animals are properly classified. However there is no such thing as a species: it cannot be directly studied or observed. It is a category, constructed by the mind, which is useful in understanding the relationships of different types of living organisms. Members of a species are, however, open to observation and experiment (p. 23).

Cronbach and Meehl (1955), who first introduced the notion of construct validity, assert that “Construct validation is involved whenever a test is to be interpreted as a measure of some attribute or quality which is not “operationally defined” ” (p. 282), and “Construct validity must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured” (p. 282). For language tests, construct validity “pertains to the meaningfulness and appropriateness of the *interpretations* that we make on the basis of test scores” (Bachman & Palmer, 1996, p. 21). On a language test,

in order to justify a particular score interpretation, we need to provide evidence that the test score reflects the area(s) of language ability we want to measure, and very little else. In order to provide such evidence, we must define the construct we want to measure. For our purposes, we can consider a *construct* to be the specific definition of an ability that provides the basis for a given test or test task and for interpreting scores derived from this task. The term *construct validity* is therefore used to refer to the extent to which we can interpret a given test score as an indicator of the ability(ies), or construct(s), we want to measure (Bachman & Palmer, 1996, p. 21).

It should be briefly mentioned that test reliability, an additional critical test factor, is often considered alongside validity. Test reliability refers to consistency in measurement. A reliable test is stable over time, indicating a degree of test re-test reliability. A reliable test is also internally consistent; all parts of a test are measuring the same variable (Kline, 1993). Reliability will also mean that if there are two tests created to measure the same thing(s), then it should not make any difference which version the test-taker uses; she should obtain the same (or at least a very similar) score on either form. Reliability in testing can also pertain to the marking or scoring of tests. It should not matter which rater scores a test, as the scoring process and raters should both be reliable. These two qualities, test reliability and test validity, are both essential for test usefulness, and exist in a close relationship with each other. If a test is not reliable, then it is not valid. “Reliability is a necessary condition for construct validity, and hence usefulness. However, reliability is not a sufficient condition for either construct validity or usefulness” (Bachman & Palmer, 1996, p. 23). At the same time, test validity can be thought of as requisite to reliability; validity is required before reliability can be discussed in any meaningful way, since if the test in question is not valid, then any discussion of its reliability is moot.

Validity is not always as pressing a concern as it perhaps should be. Stevenson (1981) points out that situations where there is an acute need for measurement can create, “an air of need and ready acceptance that makes it hard to be objective when a measure appears that seems to fill the vacuum” (p. 40). This can create a “favourable climate of validation” (Stevenson, 1981, p. 40) where a test is accepted as valid on the basis that it fills a need and appears to work. The problem with this approach should now

be obvious: thinking a test is valid is not the same as it being proved to be valid. In cases where a particular test has been used for a substantial amount of time it is possible that there may be pressure to rationalize the validity of the test, otherwise all of the inferences which were based upon the results of that test are thrown into question.

The importance of validity cannot be overstated; it has direct consequence for language teaching as well as language testing, since “effective teaching depends on the availability of clear specifications” (Spolsky, 1968, p. 90). As Stevenson (1981) writes, “by examining the gap between best practice and best theory, and by demanding a better measurement for the width of the gap, we might learn something about the construct in the process...If we wish to go beyond faith in our measures, and beyond their face validity, we must also be willing to take a very critical look at where we stand, and how far validation must go” (p. 57). Creating a second language test requires a strong understanding of the construct which is to be tested, and a development process which allows the test to be checked and re-checked for validity and reliability. Testing and test development is necessarily an ongoing process which should contain a series of checks and balances. Opening up testing practices for investigation may not be comfortable, especially in situations where a particular style of assessment has been used, unquestioned, for some time, but it is necessary.

The above review of literature has looked at five themes: the cognitive processing of ASL in both deaf and hearing users; perspectives on teaching and learning ASL as a second language; how second language proficiency is defined in SLA theory; how second language proficiency may be assessed and; exploring how assessment validity can be

determined. Having concluded the literature review, the next chapter will present the method that was followed in conducting the present study.

Chapter 3: Method

3.1 Research Questions

This study is framed by several key issues surrounding ASL proficiency and second language testing. The previous review of literature presented these issues in five sections: the cognitive processing of ASL; perspectives on learning and teaching ASL; definitions of second language proficiency; measuring second language proficiency; and determining the validity of inferences that we draw from second language assessments. Each section identified issues that arise when one tries to investigate what constitutes ASL proficiency for hearing, adult learners, and how proficiency in a second language is determined. This study seeks to investigate ASL proficiency and proficiency assessment within a particular learning context, while keeping these concerns in mind. My research questions, which were introduced in Chapter one, are as follows:

- 1) What constitutes ASL proficiency for hearing learners acquiring the language in this context?
- 2) How is proficiency determined through current testing practices?
- 3) How do students respond to current testing practices?

It is important to first acknowledge that my stance in writing is linked to a “perspective on social phenomenon emphasizing the role of human action and agency in social life” (Kelle, 2004, p. 204). All of the participants in this study are program stakeholders, either as students or as teachers. Their actions over time have shaped, and will continue to shape, how the program evolves. They will not always agree with each other, and their actions may at times even be at odds. I respect all of my participants, but I also acknowledge that they are human; in being human they will say and do things that

are outstanding, but also things that may reflect less flatteringly on themselves or on the program within which they are teachers or learners. It is vitally important to openly discuss both the positive and the negative findings of this study. As Alderson (2009) affirms,

it is impossible to understand language education, and particularly its processes and products, if one only looks at the claimed successes, the theories advocated, the sanitised accounts of behaviour. Much that happens within language education goes unreported, is unacknowledged in the literature, and at best is considered periphery to academic study, if not downright irrelevant. (p. 235).

This is an important stance to keep in mind when reading the results and the discussion sections that follow.

3.2 Research Site and Design

Since no two language learning contexts are the same and “the guidelines that we can find in the literature rarely provide the exact answers to one’s specific research questions” (Dörnyei, 2007, p. 16), the current study was designed using a mixed-methods approach. This was done in an effort to increase strengths while eliminating weaknesses (ibid, p. 42) that may be present when using only a quantitative or only a qualitative approach. A mixed-methods approach was also the only way to really answer the research questions as I have framed them.

This research was carried out at a large Canadian university. ASL is offered as a minor at this university, and so the program has four years, or four levels (equal to four credits worth) of courses. It was important for this study to engage in multi-level analysis in order to gain a richer understanding of both how instructors approach proficiency and assessment, as well as how the learners are responding to the program and the current assessment methods. Using a mixed methods approach was also important for

maintaining a level of external validity in a study that is very context dependant. As the program is so large (and continuing to grow and expand) it was important not only to get a sense of how instructors operate within the program, but to make visible a snapshot of the students as well. Following approval by the ethics review board (appendix A), data collection began. Qualitative and quantitative data were collected concurrently, but were not merged until after analysis had been completed on both sections separately, in a convergent parallel research design. By mixing qualitative and quantitative datasets, “the researcher provides a better understanding of the problem than if either dataset had been used alone” (Creswell & Plano Clark, 2007, p. 7). For the sake of clarifying the methods used in this study, procedures for qualitative and quantitative data collection and analysis are described separately. Total data collection for both phases took place over a two-month period during the 5th-12th week of classes in the fall semester of 2013.

3.3 Phase One Data Collection: Qualitative Data

Qualitative research typically collects data in a natural setting and is followed by analysis that is inductive as well as deductive in order to establish patterns, and is summed up in writing that “includes the voices of the participants, the reflexivity of the researcher, a complex description and interpretation of the problem, and its contribution to the literature or a call for change” (Cresswell, 2013, p. 44).

3.3.1 Teacher-Participants

Participants for the qualitative portion of this study consisted of both full-time and contract instructors in the university’s American Sign Language program. Letters of invitation were hand-delivered individually to all instructors in the program. I decided to hand-deliver the invitations rather than relying on email because of the expectation and

importance that face-to-face communication has in the Deaf community. As three of the seven instructors in the program at the time were Deaf it was necessary to provide the information in written English, and to also to give a signed explanation of the research and to answer any questions in ASL. For consistency of recruitment practices I decided to hand-deliver and explain the research in person to each instructor, regardless of their status as Deaf or hearing. All instructors were informed that agreeing to participate would entail two parts: A one-on-one semi-structured interview with the researcher, and classroom observation. Since classroom observation time was not able to be scheduled with each instructor, this data was not used in analysis.

3.3.2 Teacher-Participant Demographics

As shown in the table on the following page, this study included five teacher-participants. Two instructors have had less than five years' experience teaching ASL and were considered novice, while three have more than five years' experience and are considered 'experienced'. The participants ranged in age from mid-twenties to over fifty, though age was not considered as a factor in this study. Three of the teachers are hearing, and use English as a primary language. Two instructors are Deaf, and use a signed language as a primary language. Teacher gender is not noted, as it was also not a factor which was considered in this study.

Three of the participants are former students of the program. Their insights are valuable, because they have experienced the program, and the learning process for ASL, two ways: first as students, and now as instructors. In Chapter Four the information from this table is summarized after each quote, so that each comment is attributable *not* to an individual participant, but to level of experience (N or E), and Deaf or hearing status (H

Table 1: Teacher-participants by level of experience and hearing status

Participant #	Novice or Experienced (N or E)	Deaf/Hearing (D or H)
1	N	H
2	E	D
3	E	D
4	E	H
5	N	H

or D). Including both native and non-native users of ASL provides an important point for comparing attitudes regarding teaching and assessing a visual language. As Medgyes (1999) proposes, these two groups (native and non-native speakers) represent “different species” (p. 25) of language teachers. While the native speaker has long been promoted as the linguistic ideal to which students aspire, non-native speakers are now being recognized as important contributors to language use (Kramsch, 1997; Llurda, 2004). Non-native speakers have experience switching from their native language to the target language, and this “enhances their understanding of the demands of the learning situation” (Llurda, 2004, p. 318). For investigating ASL proficiency it is important to keep the perspectives of both of these groups in mind, and to think about what each can offer. Do native and non-native signers describe and assess proficiency differently? Do they identify the same features of ASL as necessary for clear and comprehensible signing and sign language development?

3.3.4 Procedures

Ten key questions related to ASL proficiency and assessment practices were developed (appendix B). These served as the basis for semi-structured interviews with teacher-participants. Following ethics board review and approval for this research, interviews with Deaf and hearing instructors began. Interviews were conducted in the preferred language of the instructor. If the language was ASL, then the interview was

video-recorded. Participants interviewed in English were given a choice of being video or audio-recorded. Each participant was asked a series of questions related to ASL proficiency, assessment, teaching, and learning and it applied to the level that he/she taught most often. The set of questions was the same for each interview, with variations arising depending on participant response. Interview participants had the option of choosing to not answer any questions, or of ending the interview at any time without consequence. Following the interviews course outlines from all instructors were requested in order to add to the data pool. These were used only to determine if teachers included course outcomes in their outlines.

3.4 Phase Two Data Collection: Quantitative Data

3.4.1 Participants

The participants for this section of the study were all students (N= 391) enrolled in the university's ASL program during the fall semester of 2013. Participants were drawn from all four levels of ASL instruction. The ASL program currently includes several hundred students. Using a wide-scale survey was the most pragmatic way to gather information from such a large pool of participants.

3.4.2 Student Demographics

Since the survey was not completed by all students, it was more accurate to examine enrolment numbers using information provided by the university's modern languages department. The table on the following page shows the enrolment numbers by course code for fall 2013.

Both ASL 1010 and ASL 1110 are first-year courses. ASL 1110 is an intensive course – it is comprised of all of first-year (usually spread out over both the fall and

Table 2: ASL enrolment by course/level code, fall 2013

Course code:	ASL 1010	ASL 1110	ASL 2010	ASL 3010	ASL 4010
Number of students enrolled:	459	28	75	50	21

winter semesters) compacted into the fall semester. For the purposes of this study both ASL 1010 and ASL 1110 are simply considered ‘first-year’, though in tables I have chosen to display them separately. Courses at the first-year level (1010 and 1110) made up 76.9% of the total enrolment for the fall of 2013. After first year, student numbers drop significantly: 11.8% of enrolment was in second-year, 7.8% in third-year, and 3.1% in fourth-year. This represents an 84.5% decline between first and second year. It is not unusual for language classes to have their highest enrolment numbers at the introductory level(s). Enrolment numbers in the university – for languages other than ASL – reveals that this trend is usual. However, as noted by Furman, Goldberg, and Lusin (2010), attrition rates between introductory and advanced ASL classes are often higher than those found in other language classes, averaging at a ratio of 4:1 for four year institutions in the U.S. If we apply the term ‘introductory’ to first and second-year courses, and ‘advanced’ to third and fourth-year courses in this context, the resulting ratio is 7.9:1, almost double what is reported by Furman, Goldberg, and Lusin (2010).

In total, four first-year classes (out of a total of 12) did not participate. 56% of all first-year, 85% of second-year, 64% of third-year, and 85% of fourth-year students completed the questionnaire, meaning that the study is representative along demographic lines. Population size for fall 2013 was 633, with a total of 391 responses indicating a 3.07% margin of error and a 95% confidence level. The total number of surveys completed was sufficiently large to proceed with quantitative analysis of results.

Participants were 67.9 percent female, and 31.1 percent male. The greater number of female students enrolled in ASL courses is in keeping with results found by Peterson (2009), though the female/male split is somewhat more even. The four most common majors were engineering (19.4%), psychology (14.3%), criminology (6.2%) and communications (6.2%). The next highest were linguistics (5.7%) and journalism (4.9%), with all other majors making up 4.7% or less (respectively) of the total enrolment by major.

82.9 percent of respondents said that they had no prior experience with ASL before entering the program. Of those who responded that they had had prior experience, written-in information indicated that most had only had limited exposure to the language, such as learning a few isolated signs in elementary school, or having learned the manual alphabet.

3.4.3 Instrument

The instrument used for gathering quantitative data was a questionnaire (appendix C) that consisted of two parts. The first section was made up of 43 questions. The first ten questions collected general demographic information. The remaining 33 questions elicited participants' opinions using a 6-point Likert scale (1=strong disagreement/negative response to 6=strongly agree/positive response). A 1- 6 scale was purposely selected in order to require that participants to favour one side (agreement/disagreement) over the other. The only question that did not fit this pattern was question #21, to which participants responded by ticking one of three available boxes.

The second section of the survey was made up of 26 ‘can do’ statements. Participants again responded using a 6-point Likert scale (1=strong disagreement/unable to perform the associated task to 6=strongly agree/can easily perform the associated task) that asked students to determine how easily they can fulfill an action in ASL. The scenarios used in the proficiency self-assessment were developed in part by examining levels of proficiency as described in the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) (Council of Europe, 2001) which I modified in order to accommodate the manual and visual requirements of ASL.

I piloted the first section of the survey in two first-year ASL classes that I was teaching at the time. These classes did not participate in the final survey. Their feedback and comments aided in clarifying and rewording sections of the survey. The proficiency section of the questionnaire was added after the pilot was complete. Due to time constraints this section was not piloted before administration.

3.4.4 Procedures

Surveys were distributed during the 9th-11th week of class to ensure that all students would have experienced at least two ASL assessments that term. I decided to use paper surveys and to distribute them during class time in order to attain the highest response rate possible. Arrangements were made with each instructor beforehand, and permission to distribute and have students complete the surveys during class time was obtained. For each class visited, I introduced myself and provided an explanation of the focus of my research and a description of how the surveys fit into my research design. This was done first in ASL, and then in English. It was clearly stated that completing the surveys was voluntary and would have no bearing on the students’ grades. Surveys and

consent forms were then distributed to the students. Two classes (aside from the two that I was the instructor for) were not surveyed, as the class instructor felt it would take too much time away from the class. In order to maintain continuity of participant recruitment, students from these classes were not invited to participate.

Participation was voluntary. Those who chose to participate were given time to fill the questionnaires out immediately. Survey completion took approximately 20-30 minutes. I remained in the classroom during the time that students filled out the surveys. This was a conscious decision, made so that if any questions arose I was on hand to clarify them. Upon completion of the surveys, consent forms and surveys were collected in separate piles so that the names of students who participated could not be attributed to any particular survey.

Immediately following each classroom visit, the surveys were deposited into colour-coded envelopes according to the class level and instructor. I did this so that it would be clear, later, which surveys had come from which classes. This also meant that while hearing status of instructor was not included as a survey question, this information was still available to me.

Since the surveys were not distributed to every class, precise numbers for total student enrolment, and enrolment by level were not known. In order to rectify this, and to provide a clearer picture of the robustness of the program and levels, enrolment numbers for the past two years were requisitioned from the language department. These figures are used only for determining enrolment and numbers of students at each level.

3.5 Phase One Analysis: Qualitative Data

Interviews which were conducted in ASL required an interpretation from ASL into written English. I use the term interpretation rather than translation, because there is no direct ‘fit’ between ASL and English. In some cases this meant I needed to add a script for un-signed, but pertinent information (eye-rolls, shrugs, facial expressions) which alter the tone or meaning of the manually produced signs. As discussed by Hole (2007), transcribing and transforming signed discourse into written English can be complex. For this study the process involved moving between the recorded visual-language interview and the written transcript several times in order to develop a transcript that was as faithful to the meaning of the interview responses as possible. Interviews which were conducted in spoken English were transcribed into written English. “Ahs” and “ums” (whether voiced or signed) were transcribed, but length of time for pauses was not tracked. The five interviews amounted to a combined duration of four hours and thirty minutes, and were transcribed over fifty-seven typed, single-spaced pages.

Once written transcripts had been produced from all interviews, these transcripts were then coded using a modified constructivist grounded theory approach to coding (Charmaz, 2006; Charmaz, 2008). Initially the codes were heavily informed by the research questions. As coding occurred concurrently with collection, codes and categories were modified and reorganized several times as new information clarified and modified the overall picture. This approach to coding was used so that salient codes would emerge over time, leading to a more well-rounded investigation of the interview responses and draw out important information that was not directly addressed by the interview or research questions. Preliminary codes, chosen through a process of memoing and note

keeping, were developed as data collection occurred. Once all interviews had been completed, these codes were used to organize the initial full coding of the data. The preliminary codes included: grades/grading; teacher perspectives on students; perspectives on ASL testing; current testing practice; and descriptions of proficiency. Each interview was examined individually, line by line, for evidence of data that fit those categories. New categories and codes that emerged were also recorded. Following this, a fresh “file” was created for each participant, and the interviews were re-coded, with special attention paid to information that was repeated by several participants. Verbatim extracts from what seemed to be the crux each person’s comments were also recorded. These files were then compared and contrasted, and the results of this comparison resulted in restructured, final categories which are more representative of the core variables. The evolving and emerging theory was informed by my own experiences as a teacher in the program as well as de-briefing sessions held periodically with colleagues. I did not approach the process of coding with an empty head, but as Dey (2010) recommends, I strove to keep an open mind about the “cogency and the relevance” (p. 176) of any ideas I found myself turning to when coding. The constant process of comparing and contrasting data while developing a theoretical frame results in what Isabella (1990) describes as “a reconceptualization, often based on a creative leap, that should account for and encompass all nuances in the data” (p. 12). Since there is a level of subjectivity involved in coding qualitative data, a second coder was used, with agreement of codes reaching 85%.

3.6 Phase Two Analysis: Quantitative Data

All questionnaire responses were entered and analyzed using the SPSS statistical package (v.21). While inputting the data, a total of 4 questionnaires were excluded because 1) participants had responded to less than three quarters of the questions 2) the participant had clearly not read the question (e.g. selected the same response throughout the whole questionnaire, or responded in a zig-zag pattern). All responses where the participants had circled an area *between* the points of the Likert scale were inputted at the lower response (e.g. if a respondent had circled the empty space between 2 and 3 then the response was entered as 2.) A reliability analysis of all questionnaire items (excluding demographic information) showed strong internal consistency (Cronbach's $\alpha = .924$).

Descriptive statistics were used to provide meaningful summaries of the data from the first section of the questionnaire. This examination commonly included frequency distributions, and cross tabulation to examine differences across groups. Since responses were selected from a Likert scale that ranged from 1 to 6, this was used to gauge general agreement or disagreement, i.e. responses of 4 and up were considered to be agreement/positive responses, while those that were 3 or lower were interpreted as disagreement/negative response.

For the proficiency section of the questionnaire an exploratory factor analysis (EFA) was used to investigate the data in order to determine what factors accounted for the variance in students' self-reported reported ASL proficiency. This approach meant that no constructs were imposed a priori. Rather, results were allowed to emerge from the data. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy for this dataset was

.941 which is ‘marvellous’. Bartlett’s test of sphericity indicated the correlations in the data were appropriate for factor analysis ($p < .001$).

Factor selection was guided by eigenvalues of 1 or more and whether they were located at an ‘elbow’ on a scree plot. I selected principal axis factoring as the extraction method. This approach “treats each X variable as a measurement that, to some extent, may provide information about the same small set of factors or latent variables” (Warner, 2013, p. 860). I then used an oblique rotation method (Promax), which assumes that the factors are correlated (Gorsuch, 1983). Analysis of the proficiency questions converged in five iterations into four factors that account for 56.775% of the variance amongst student responses. Factor one is represented by nine items; factor two by seven items; factor three by six items and; factor four by four items. Factor four is the weakest, as it is only represented by four items. However, as this is an exploratory factor analysis I decided it was acceptable to keep it included.

3.7 Ethical Considerations

I do not consider myself a cultural insider, or a member of the Deaf community. However, my position in this study (as an instructor myself) granted me access to teacher-participants. I was very fortunate that I was able to recruit several interview participants for this study, and I recognize that this is most probably due to the fact that they know me and have worked with me for several years. As such there was already, before this project began, a level of familiarity and in some cases trust between myself and my participants. I take this trust seriously, and have done as much as I can to ensure that the personal and possibly sensitive nature of the information I have been given is treated respectfully. The research site is small and the number of participants was limited.

I have done my best to present the data in a way that tries to maintain confidentiality, but also feel it important to be up front in acknowledging that there may be some details in the information that would allow participants to be recognized by those who are familiar with them, or with the site of the study.

My participants also engaged in honest expression which created a level of conflict for me because I was not at arms' length. I have tried to be careful to protect my sources and not violate that trust. In some instances this meant I was forced to filter the data when making decisions about which parts to present. This was a learning opportunity for me: being at arms' length from my teacher-participants would have allowed me to report on all data without concern for protecting my participants. However, I would have missed out on a lot of key insights.

Having provided information on the design of the study, the analytical approaches and the procedures followed, in the next chapter I present the results of the study. An interpretation and discussion of the findings is included alongside the results, rather than as a separate chapter in order to address and contextualize the results as they are presented.

Chapter 4: Results and Discussion

Based on the quantitative and qualitative analyses described in Chapter 3, this chapter presents the results of the study, which are guided by the research questions. In proceeding with this research I was interested in discovering three things: 1) What constitutes ASL proficiency for hearing learners acquiring the language in this context? 2) How is proficiency determined through current testing practices? And finally, 3) How well do students understand the construct of proficiency, and how do they respond to current testing practices? I begin with a presentation and discussion of the results of the qualitative analysis. These results are all drawn from the teacher-participant interviews and examination of course outlines, and they address the first two research questions. This is followed by the presentation and discussion of the results of the quantitative analysis of student surveys.

4.1 Qualitative Results

Although the research questions guided initial coding (because I was looking for specific responses in relation to those questions), I also allowed other codes to emerge that capture the dynamism in the data. These qualitative results are reported as themes (Saldaña, 2013). As such they do not exemplify every possible code and category, but are representative of the comments and opinions, observations, and remarks that the teacher-participants discussed most often.

The six primary themes arising from analysis of interview data are:

1. Descriptions of current assessment practices
2. Descriptions of ASL proficiency
3. Grading of ASL assessments

4. Validity concerns
5. Perception of students
6. Program change

These are presented in the order that the teacher-participants reported as having the greatest influence on their decisions, actions, and opinions. For each theme, direct quotes from the interviews are included so that participants' own voices and responses are incorporated. As described in Chapter 3, each quote is followed by two letters indicating whether the teacher-participant is experienced (E) or novice (N), and whether they are Deaf (D) or hearing (H).

It should be stated at the outset that this study is not capable of supplying conclusive answers to questions of best methods or best practices. Neither a set of best practices for teaching, nor a best method for assessment has been established by either second language theorists or teachers. The aim here is to identify and describe the current approaches and attitudes to proficiency and proficiency assessment held by ASL teachers in this context, and to compare and contrast this with student responses. This will, it is hoped, provide a sense of how effective current assessment practices are.

4.1.2 Current Assessment Practices

4.1.2.1 ASL Assessment Format

All of the teacher-participants reported that they follow the same test format, which I will now briefly describe. Each ASL test is made up of two sections: the first is referred to by teachers as the 'required section'. This is an introduction where students give key information about themselves, such as their name, major and minor, teacher's name, the current date, etc. This is used by all of the teacher-participants and includes the

same information, with only one or two very slight variations. This is followed by a ‘free’ section, in which teachers may ask that students do any number of things, for example – respond to specific prompts, produce a pre-practiced monologue, show how they would use specific ASL structures, or discuss a given topic. This format does not change with the level of the course.

So the tests are broken down into two parts, so there’s the required elements and then there’s the other part... [the introduction is] the same one that continues into third year and fourth year as well. (EH)

So there’s a required section and a free section. (NH)

I tell them right off, you know, there’s five marks for the required, it’s called the self intro... (NH)

While there is no official program document (that I am aware of) stating that teachers must pattern their tests in this way, this procedure is used by the most senior teacher, and has become a norm of the testing culture of the program. In writing on the concept of culture, Trompenaars and Hampden-Turner (1998) describe the norms of a culture as those things that define a group’s sense of right and wrong and provide a sense of how one should behave. The value that this test format provides, or what aspects of proficiency it accounts for, was not discussed by teacher-participants. It appears that the format is accepted as a given and acts as an informal level of social control.

Current tests are (with the exception of one test in third-year) all expressive in nature. ‘Expressive’ is a term that is used frequently within the ASL program. It refers to the productive aspect of ASL (signing) rather than the receptive or comprehension aspect (understanding what someone else signs). When teacher-participants refer to expressive tests, they mean those assessments where students are evaluated solely on their ability to

sign. There is no formal (graded) assessment of students' ASL comprehension abilities, and teacher-participants reported little-to-no concern for this lack of receptive language testing. In terms of test time, participants also report that their tests are short, with total testing times (excluding set up) that range from three to seven minutes, occasionally hitting the ten-minute mark. Buck (2009) writes that "all other things being equal, the longer the test the more accurately it will measure the construct, and the more thoroughly it can cover a broad range of content" (p. 170). He notes that this must be balanced by "the practical constraints of how much time test-takers can be reasonably expected to spend taking the test" (p. 170). In some instances, there may be pressure to condense language assessments into shorter and shorter time spans. But when a test time is too short, an adequate sample for a reliable test cannot be produced (Buck, 2009).

4.1.2.2 Assessment Content

Current testing practices within the program are at once restrained (by the format) but very open (test content). What results is a context where teachers have a lot of freedom, and limited guidance. As a result instructors teach, and assess, different things:

There are some topics that other teachers focus on that I don't generally bring up in class, so I wouldn't test on that. I only test on material that I cover in class and sometimes it's just not possible to cover things that other teachers are covering because there's only a set amount of time. (NH)

[What I assess] depends on the goal... There are different ways of teaching that have to show on the test. For example, gesture, that started to be included in tests three years ago. Now there's hand configuration being included on tests - that's new. It changes all the time, so it depends. (ED)

No, there aren't always specific topics. The program goal for each semester is really the culture lecture. All of the vocab and concepts are connected to the lecture. (ED)

Some teachers handle this by surveying colleagues about what they plan to teach and/or test and then planning lessons that are similar. This strategy is reported more by the novice teachers, and may provide structure in the face of uncertainty. Meeting with other instructors is also a way of procuring information and advice in a program where there is no curriculum to follow.

How free or open the ‘free section’ is depends on the instructor. Some of the teacher-participants report that they only supply students with a topic, and it is up to the student to use the topic as a catalyst for language production. Others ask that their students to show them specific language structures. Teachers are granted this freedom because the ASL program permits each teacher to develop his or her own curriculum to use within the classroom; teachers are able to choose assessment material and content that they feel best fits their class. There are loose and flexible global proficiency expectations for each level, but there are no firm competency or linguistic objectives. This flexibility means that when selecting topics for language production, teachers are not necessarily selecting the same topic(s). From a consistency standpoint this is potentially problematic; different topics require not only different vocabulary, but also different skill sets and different linguistic structures. For levels with more than one instructor, it means that what is assessed in one class is not necessarily assessed in others.

4.1.2.3 Test Type

ASL assessments are a combination of test types. Teachers are interested in both predicting how well a student will be able to use ASL for communicative purposes (proficiency), and also whether they have mastered something that has been taught and practiced in class (achievement). Depending on the instructor’s approach to assessment,

the assessment may include aspects of both test-types. The test content may be drawn from what has been taught in class, but the assessments also reference forward, to the application of language achievement in relation to the sort of tasks that students may later be required to perform in ASL. In some cases, students are also asked to demonstrate knowledge of Deaf culture. It is not clear in these latter assessments if language ability is in any way scored separately from cultural knowledge. Including items other than those designed to highlight language runs counter to the idea that language tests should test language and little (if anything) else. According to Weir (2005) the 1960s brought about a shift in language testing, “towards a view that language might be divorced from testing literary or cultural knowledge...which focuses on language as against an assortment of language, literature and culture” (p. 17-18). Introducing elements other than those which are necessary for determining language ability into a language test may compromise the validity of test results, since it introduces variables outside of the domain of how students use the language. This alone does not mean that the assessments are bad or invalid, but it is one point to consider on the sliding scale concept (Weir, 2005) of validity. An additional validity concern is that without any assessment of receptive language skills, it is impossible to claim that the current assessment practices actually provide a full picture of students’ language proficiency or their skill progression.

4.1.2.4 Teacher-Participant Attitudes Towards Testing

Even though teachers in this context have free reign over the content and grading of their assessments, attitudes toward testing were predominately negative. Tests were considered a necessary evil; teacher-participants reported feeling that education is grade centered, and so they must administer tests because it is required by the university.

Therefore, tests are important for their assistance in allowing instructors to apply a grade to student work and to give out marks, but appear to have little redeeming qualities of their own:

...the tests themselves aren't really important, no. But we *must* have tests, so. If I could throw them out I'd be happy. (ED)

It's important. It's important because the university says it's important. Yeah because it's - everything is grade centered. (EH)

Well, obviously we need a way to grade them. (NH)

...it's the only way we have of giving them their grade, so it's like priority number one. (NH)

There is a shifting of responsibility away from the teachers and onto the institution; assessments are used to “satisfy the accountability demands of an external authority” (Shepard, 2000, p. 4). Testing is viewed as a requirement rather than a potentially useful tool, and the responsibility for the fact that tests exist at all is put onto the larger educational institution rather than the teachers or even the program. Even what teachers perceive as the main benefit of testing, feedback to students, is a tangential result. Tests themselves are not held in high esteem.

4.1.3 Descriptions of Proficiency

The expectations for students' signing abilities do appear to increase with each level; teacher-participants report that new structures and new foci are added as the levels increase. First year includes a small vocabulary of signs and only very basic grammar. Second-year incorporates use of space and ASL syntax. By third year students are expected to be able to combine grammar and structure, and are introduced to new aspects of ASL discourse, such as methods of correct turn-taking and debate. By fourth-year

students should be able to discuss almost anything, though they will still face some limitations. They should also have developed a “style” to their signing. By the third and fourth-year level skills particular to visual languages become more apparent in teacher-participant descriptions of proficiency. At these levels, high-proficiency students are described as rhythmic, musical, visually pleasing, comfortable, smooth, and easy to watch. These descriptors emphasize the visual nature of the language, and highlight select areas where what is included in visual language proficiency differs from that of spoken languages. This is a very basic outline, as reported by the teacher-participants, of the key language features that students encounter at each level. The specific details which teachers examine when determining a student’s proficiency, and how they apply grades, varies widely.

Across all levels and instructors, the most common indicators of proficiency identified by teacher-participants were: facial expression, connectivity, good hold-movement-hold, and low visual noise. Beyond these commonalities, teachers described very different proficiency indicators, even within the same level. For first year teachers, the most common recurring factor for high proficiency was facial expression or animation. This was discussed by all three instructors who teach at the first-year level. Hold-movement-hold (Liddell & Johnson, 1989) is brought up by two instructors. These two particular aspects of ASL – facial expression and movement – are the two that research has shown hearing learners struggle with the most (McKee & McKee, 1992; Bochner, Christie, Hauser, & Searls, 2011). This suggests that perhaps proficiency expectations and student acquisition patterns may not be well-matched, at least at the introductory level. All other proficiency indicators that teacher-participants discussed

were different, or were rated at different levels of importance. This indicates that the construct of proficiency is not consistently defined throughout the program, even at the same level, and is articulated differently by each instructor:

I think the biggest difference between those three groups, low, mid and high range students, you're going to see it in their face. Right away. A student who has good animation, it's rare that they're going to be in that low group. (NH)

However, for another instructor, proficiency is about what is *not* shown as much as what is shown and includes the ability to transfer language between contexts:

I look at their free section kind of as, whether I understand it, but also whether they've taken something that we've done in class and brought it into their video correctly. So if we've just talked a bunch about time, or time topic comment, and they're putting time at the end of their sentence then they haven't quite got what I taught in the lecture. (NH)

For most teacher-participants, judging a student's proficiency comes down to relying on an internal framework. This was connected to each teacher's sense of 'understanding'. Understanding, and variations of it, was discussed by every teacher in relation to how they determine proficiency, and by direct extension to how they grade. Instructors often referred to whether they themselves understood their students:

...I'm mostly looking to see if it's understandable. (NH)

...I do look at overall whether it was understandable. (NH)

In the free section...they can use gesture, they can do what they want. The goal for that section is to understand what they're talking about. I don't care about mistakes. (ED)

...again it comes back to the nature of the error do I still – how easy is it for me to still understand. (NH)

Sometimes the hold-move-hold is messy or sloppy and I can't understand it...
...when it's clear, I understand it. (ED)

...it's rare that I go below a four, to be honest. Because I understand most of them. (NH)

If I can't understand it, it's wrong. (ED)

As seen here, 'understandable' and 'understand' is used as a measure when determining proficiency, but that measure is based on an internal vision, particular to each instructor, of what proficient signing looks like. The instructor, and her understanding, is heavily relied upon as the model for proficiency. At a basic level, this concept of understanding is consistent with goals of communicative language assessment, which (in productive language assessment) prioritizes appropriate language use rather than grammatically correct sentences (Kathleen & Kitao, 1996). However, it is perhaps more accurate to describe current ASL assessments as having communicative elements rather than being communicative language tests: though the students sign face-to-face with a partner, the non-signing partner is merely a sounding board rather than a true interlocutor. The language students produce for an assessment is often rehearsed rather than spontaneous and students are rarely required to produce appropriate responses in real time.

In evaluating student performance on assessments, some teachers find that vocabulary errors, or incomplete signs are important, for others it might be precise movements, while another will ignore those and focus on paralinguistic features such as hand dropping. These differences were most strongly pronounced between teachers at the same level, but there were key cross-level variations as well. While in first-year there is a heavy reliance on gesture in lieu of signs:

I think that's the main goal of this first term proficiency. Giving them a bit of confidence, showing them that they don't have to stick to, they don't have to know specific vocabulary items, they can still act things out. (NH)

This was not so in second-year:

...it's not that I don't like gestural communication, but there shouldn't be a heavy reliance on it. It should be more specific and less sort of like, it puts the onus less on the listener to fill in the blanks. (EH)

I would love it if all the first year teachers would introduce the idea that ASL structurally does not resemble English. Unfortunately there are some instructors who don't want to burden the students with that so they just let them play with the language however it comes out. (EH)

This suggests that after using ASL one way for (in most cases) a year, students may be surprised to find out that it is actually much more difficult, and much less iconic, than what they were initially led to believe. This may create pressure on instructors teaching at subsequent levels, as they must be the ones to get students to discard a lot of what they have been using in favour of more correct structures.

Inter-level differences in proficiency expectations can lead to tension between levels, and may make it more difficult for students to transition successfully from one class to the next. The skills students learn and are accustomed to in one level may not be accepted in the next. While first-year teachers focus on getting students to use the language as much as possible, and will encourage them to supplement their signing with gesture, there is little focus on structure. This becomes problematic in second year and later, when correct structure is expected. Despite there being a general proficiency increase between levels, by the end of fourth year it seems that students' understanding of the linguistics of ASL remains low:

[By fourth year students have] sort of very layman skills. I would love to have more emphasis on the real linguistics, but again, most students are thoroughly uninterested in that. (EH)

Apart from the four most common proficiency indicators listed earlier (facial expression, connectivity, good hold-movement-hold, and low visual noise) each teacher is working from a different construct of proficiency, and this directly influences how they apply grades to students' work.

4.1.4 Grading

As Stedman (1984) writes, “most tests are used for the purpose of assigning grades to students. The primary rationale for this practice is presented as a means of motivating students, maintaining standards, and controlling quality” (p. 2). With each instructor deciding individually what it means to be proficient in ASL, instructors also grade their assessments differently. There is no consensus, and no rules, that are applied across the board when it comes to grading. It is left up to the individual classroom teacher:

We're the instructor, it's our course, we do have some wiggle room as to some ability to – to grade the way that we feel is appropriate. (NH)

Right, how do I grade. So for that test it's can they analyse and can they express themselves, how do they critique someone else? It means I focus on facial expressions and grammar that connects to their signing. Are both aspects included and do they mesh properly? ...Plus, sometimes I watch for classifiers because classifiers help students stay away from English. Grading will also sometimes include things that I see, generally things like, is their signing smooth, does it flow well, is it clear, is there facial expression - all of that. (ED)

Yeah. So I follow the five point, well ten point scale, so five points for the required section and five points for the free section. When I'm looking at the required section I'm looking to see if they have every part of the required section, if the signs are produced correctly. Since we go over them so often they should be produced correctly. I'm, I'm looking to see whether anything's missing, how their fingerspelling is because that tends to degrade over time...But the required section since it's practised so much really, it's easy to grade because you can tell right away if something is off. (NH)

So the really common errors that were popping up I made sure that I had a consistent amount that I removed, amount of points that I check off. That works well for the self intro, for the free elements everyone's doing something different so again it's subjective, it's hard, I mean I don't really have an answer. I look at a few videos, I try to make sure, I think of them overall as well, so this student got a 7.5. This one, you know, definitely should not be more than that because this one student, overall, is stronger than the other. So sometimes I have to fiddle with the marks... (NH)

So that one [the introduction] I'm nitpicky about the signs, very very nitpicky about the signs. In the second part the "free elements" [air quotes], not really free at all. That one, it depends on the test. Today, on the comparison test, I was looking for a specific structure, I wanted them to use spatial locations to set up their two items and then to make a comment on it that I had provided. Within that the sign production I let it slide. So I told them what they had done wrong, but that didn't cost them anything. (EH)

Without clear definitions of the construct to be measured by a test, making decisions about students' test performances is difficult. It results in grading that differs depending on whose definition of proficiency is used, whose grading process is followed, and what their decision making style (Baker, 2012) is like. No participants reported creating or adhering to common threshold levels or proficiency bands when determining grades.

While all participants described how they graded, and what (to them) separates a high proficiency from a low proficiency student, no teacher-participant connected how they grade, and what determines a student's grade, to any particular number or range of numbers. There is no consensus or description of how the assessments are quantified. What, for example, does it mean when a student receives a grade of 70% on an assessment? What scale is the grade tied to? It is also unclear in some instances what assessment grades are meant to signify. Do they represent a mastery of material, a level of proficiency, or a combination of both? Are they a predictor of future language ability outside of the classroom? There is a gap where definitions of proficiency, scoring, and

meaning or use of the score should meet. The validity of the ASL assessment practices are brought into question, because “the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on tests scores” (Messick, 1989b, p. 5) is unchecked. Without clear definitions of the construct to be examined, rules for scoring, and a consistent application of score interpretations across classes, grades become meaningless. What teachers are engaging in is not testing, but teacher-made assessment activities which are not generalizable beyond the classroom.

4.1.5 Validity Concerns

Throughout the interviews, teachers expressed concerns about proficiency and testing stemming from a number of areas. Coding revealed a striking lack of faith in their tests as accurate tools for measuring student proficiency. Teachers questioned several facets of their tests’ validity, particularly content/context, and scoring. Teachers also indicated a high level of concern for a possible test method effect that they believe interferes with correctly measuring proficiency. While teachers may know what they are looking for when grading they set up their tests in an ad-hoc fashion. There were no reports from teacher-participants of a priori validation. There was also no attempt to fully describe and operationalize the construct to be measured before administering a test, save for one instructor. This instructor described deciding beforehand what specific linguistic features to test, and then building a test to try to measure those features. All other instructors use a much looser test design strategy; they attempt to look at overall communication ability by presenting students with a topic and asking them to produce a language performance based on that topic.

4.1.5.1 Under-Representation of the Construct

Whether the tests are more or less structured did not make a difference in teachers' feelings regarding test validity. No teacher felt that they got credible test results, and several teacher-participants expressed concern that the tests they are using do not provide an accurate read on students' proficiency. This may be because the current tests under-represent or fail to fully describe the construct to be measured:

I think that they can learn a lot throughout the course and that's not necessarily reflected in all their videos. I think they might have good skills that don't show up in the testing process. (NH)

In some cases validity is in question because the test performance compared to class performance is vastly different. One teacher-participant commented:

...today was test four and almost everyone got nine or ten out of ten. So I'm looking back and I'm, 'oh well, that's pretty high, but I know that you're not that good in class', so, right. (EH)

If the content is too narrow, then students may very well have other language skills which are not visible on the test because they are not included as part of the focus. At the same time, when students know the content of the assessment ahead of time, their focus when practicing may be only be on what they know they will be assessed on, leading to assessments that end up measuring how much a student is able to successfully memorize and retain in the short term.

4.1.5.2 Language Transfer

Doubts about the success of assessments in truly accessing student proficiency are also expressed through teacher-participants discussion of language transfer. Where students are able to sign in the classroom, or on an assessment, these skills may not generalize to other language use contexts:

...what they're pulling out for a test, even though it's rehearsed and stuff, is that – I mean are they showing me what they've learned if they're able to put it together and rehearse it then they have it, but if then they can't do it spontaneously then, is that really a true reflection of their skill? I don't know. I know it would fall apart if they met a Deaf person and had to communicate with them... (EH)

They're fine in the classroom, but as soon as they meet a Deaf person they're like....eeuuugh, what do I do? (NH)

But when it comes to just free activities in class...they just play and they sign English. So what, I don't know. Am I seeing [on a test], is it like formulaic and they understand the formulas I've provided or can they transfer that into real life language? I don't know. Because they don't show me in class that they understand and that they can do it. Or they understand, they can do it, maybe they're just too lazy to do it? (EH)

The ability to transfer knowledge and use it in new situations indicates understanding (Shepard, 2000). When teachers question the validity of their assessments by demonstrating doubt in their students' ability to transfer their language skills to new contexts, it signals that the assessments, and perhaps even the instruction, are neither flexible nor generalizable. It is possible that the assessments are capturing only a very small range of what it means to become proficient in a language without building bridges to how those skills are applicable in situations beyond the classroom. Teacher-participants reported concern for how well students' abilities (both expressive and receptive) would function out in 'real life'. When discussing comprehension outside of the classroom, teacher-participants are concerned about two things: whether other ASL users will be able to understand the language that students produce, and whether students will be able to comprehend what other people sign:

Of course not. (ED) [In response to a question which asked whether the instructor's ability to understand the student meant that the student's signing would be comprehensible to other people as well]

I know what I've taught them, I know the vocabulary that *they* know, so there's a limited amount of signs that they should be using, but...to a Deaf person on the street that they meet, is this still going to be understandable? (NH)

So it takes a lot of exposure and a lot of practice to keep your receptive skills up to par. So yeah I think ... students have really mediocre receptive skills. I'm happy if they can understand me and if they can understand the other instructors. (EH)

The fourth year students here... their understanding of each other in conversation is weak... They don't understand each other. They don't mix enough. (ED)

From these results it seems that there are two approaches to proficiency occurring simultaneously: what Bachman (1990) refers to as the 'interactional/ability' approach compared to the 'real-life' approach. In teaching and subsequently developing assessments, teacher-participants are working from an ability stance where language proficiency is defined through component abilities. However, when they reflect on their students' language performances they compare them to expected real-life scenarios, where proficiency is a characteristic of language use in a particular domain, and find them lacking.

4.1.5.3 Method Effect

Several teachers commented that they feel their students perform better in class than on tests, leading to a general belief amongst a majority of the teachers that students 'just don't do well with testing'. Often, this is blamed on a test method effect. While teachers unanimously agree that video testing is necessary for a visual language, they also believe that it contributes significantly to test anxiety:

In some ways I think it [the testing] hurts them, because they're so nervous. Some of them do really well in class and I'm surprised by their video, like it's not as good quality as what I know they're capable of. (NH)

...in the class they're very involved and they're always willing to start practising and sign with other people and then they get into the test, and the test makes them nervous, and they just don't produce something that's as good as they could do in class. (NH)

They freeze up in the lab. They start out too formally. In the class, it's casual. (ED)

I have a couple of students it's clear they're soooo nervous. They are weak students - I've seen it in class so I know what to expect but the video test makes it even worse for them. You can see how nervous they are. (NH)

The latter comment is, perhaps, telling – the fact that with a video test you can “see how nervous they are”. It is possible that teachers are more susceptible to the perception that students are anxious because the tests are video-recorded. Sweat, shaking hands, and rapid breathing are all made visible by the testing method, much more so than would ever be evident on a written test. This does not necessarily mean that video assessments are any *more* problematic or anxiety-laden for students, but only that their test anxiety is more conspicuous. As a result, some teachers tend to be more forgiving when grading a video test, as a way to balance what they see as a detriment to students' ability to perform well.

The current assessment practices in the ASL program under study permit instructors to see only a slice of language use. Ideally assessments are developed using principles that draw out a performance slice that is “representative of the abilities we want to measure” (Bachman, 1990, p. 34). The above examples show that instructor-participants express concern that their assessments may under-represent the construct of proficiency, and that the test format may present a significant impact on student performance. They also express worry that both what they teach and what they assess may not transfer to contexts outside of the classroom. These validity concerns indicate

that a cohesive relationship between instruction, assessment, and goals (Knigga, 2012), is not currently well-developed within the program.

4.1.6 Teacher-Participant Perceptions of Students

As described above, teacher-participants relate that they perceive student anxiety as a significant factor in assessment performance. This belief leads them to offer detailed explanations of what will be assessed, in the hope that this will lessen students' anxiety. Teacher-participants also describe being exhaustive in their feedback to students. This increasing transparency seems, on the surface, to satisfy fairness principles (Shepard, 2000), since students are made clearly aware of the criteria on which they are being evaluated:

I'm doing mini modules and then testing immediately and I'm testing in a very very structured way. So, here are three sentences that I've written in English and I want you to sign them, based on what we've just learned. (EH)

So I tell them right off, you know, there's five marks for the required, it's called the self intro, and this is what you have to say. And I literally write out the sentences in ASL gloss for them, so there's no excuses, um, for missed elements or what's the order, like, we've practiced it a *lot*. (NH)

...it has to be something we've done before. I don't want to throw them for a loop I don't want to give them anything new that's going to surprise them. It needs to be stuff we've done repeatedly in class, and that they've had a chance to get feedback in already. (NH)

So they know exactly, they know what they're being tested on, specifically all the details that I'll be looking at, all the ASL elements I'll be looking at like structure and verbs and this and this and this... These tests are really low stakes in a way. There are five of them. They're worth ten percent each, but only four, the best four count. (EH)

I do give very detailed notes. I pretty much write down every mistake they make. (NH)

...on the first test they leave and the free section is short, tiny. I inform them by bawling them out - c'mon! More free! I add comments to the test like "not enough free, not enough free". Their mark goes from 5 to 4 because it's too short. That's ten percent, come on, give me more! That improves fast. On the most recent test almost all of them were over 5 minutes. (ED)

I explain the required versus the free section and say that they're five marks for each to the students before their first test. And before every, every subsequent test we have a review of the required section so I at least write down on the board, generally second and third tests we go over it again so that they have then seen me producing the signs over and over again. (NH)

The danger in this approach is that it can encourage a move from assessment *of* learning to assessment *as* learning (Torrence, 2007). The problem is not with assessment as learning as a pedagogical choice - assessments *can* function as instances of learning. The potential issue occurs when the inferences drawn from this type of assessment are simultaneously used as evidence that a student has demonstrated or obtained a particular level of proficiency in ASL. The concern, as expressed by Torrence (2007) is that "‘criteria compliance’ comes to replace ‘learning’" (p. 282). Torrence (2007) further argues that

such transparency encourages instrumentalism. The clearer the task of how to achieve a grade or award becomes, and the more detailed the assistance given...the more likely the candidates are to succeed. But transparency of objectives coupled with extensive use of coaching and practice to help learners meet them is in danger of removing the challenge of learning and reducing the quality and validity of outcomes achieved (p. 282).

Along with the impression that students experience, and are significantly impacted, by test anxiety, teacher-participants reported that students do not put in the necessary time and effort into practicing in order to improve their language abilities:

I find that students in first year don't feel that they should practice as much as they should. They're learning a language, so, it's really very important. And when

I was learning the language I practiced a lot, constantly, and I still practice quite a bit. But they kind of see, a lot of them I think, see it as maybe something that they don't need to practice, so they only practice during class time. (NH)

I think they benefit from being forced to practice for the tests. If there were no tests then a lot of them probably wouldn't be reviewing or practising the material that's covered in class. (NH)

Different classes have different characteristics. Some classes are weak at interacting, really weak. Some classes have no grammar. What do you do, what happened there? Some are lazy. (ED)

In the above examples teacher-participants associate student performance, and by extension proficiency, with practice. Whether a student is actually practicing outside of class, on his or her own time, can never be known for certain. Yet teacher-participants attribute low performance to laziness and an inferred lack of practice. However, rather than allow students to take responsibility for their own practice habits and accept the consequences of choosing to not practice, teachers discussed alternative ways of providing practice opportunities to students:

...this year there were four journals that were meant to encourage practice and expression practice. The focus was on practice - come on, practice! (ED)

I would love that, if we could have a lot of TA's hired, specifically to run tutorial groups, and have students practice, you know, be forced to practice outside of class. (EH)

In setting up increased opportunities for student practice, student attributes of 'lazy' or 'fail to practice' are shifted back (at least in part) under the control of the instructor.

Teacher-participants in this study hold two main perceptions about students, and these can be expressed as a cause-and-effect relationship. 1) Students are anxious about video tests, and this is why they may not perform well. 2) Students do not put enough time and effort into practice, and this is why they may not perform well. Whether or not

this cause and effect relationship actually exists is less important than whether teachers believe it exists. If students are thought to be lazy in addition to being incapable of handling the pressure of assessments, then it makes sense that teachers are trying to do what they can (in some cases through over-explaining and over-supporting), in order to seek improvement in student ASL proficiency. This may not be actually helping student language development in the long run. However, without the support of program-wide outcomes and grading rubrics, it is also possible that no teacher wants to be the one giving out the lowest grades. The result is that teacher-participants seek strategies for improving student success, even when an argument could be made that the responsibility should rest (at least in part) with their students.

4.1.7 Change

Teacher-participants discuss the impact that change has on their assessment and teaching in several ways. Some describe practices that they have tried and then stopped, while others discuss change that they expect to occur as they evolve as instructors. Still others discuss changes coming from the program level that they feel pressure to comply with:

The answers that I'm giving now [about what is taught and assessed], I'm very well aware that I might give you totally different answers next term, or next year if I teach this course again, one or two more times, I'll learn and grow and all that jazz and like, I'll have more perspective. (NH)

...when I started here...I needed to teach classifiers, but now I no longer do because they already have those skills. So I can put that aside and it means I can give them tougher topics (ED)

...the test requirements change every year. It tends to get harder and include more and more. (ED)

...the program is changing really fast, you know that already. What I say right now, next year may be different. Everything that happens in first year puts pressure on second, third and fourth to change also. Change happens really fast here. (ED)

Without a program-wide curriculum, the program is susceptible to indiscriminate change. While some change over time is expected in any program, constant change makes it difficult to track what is being affected, and how. Continuous change also results in ‘moving targets’ for both students and teachers. When change happens without a plan or a goal, and when no work is done to investigate the effects of one change before another occurs, the consequences of the changes become muddled and unclear. In a surprising result, only one teacher-participant expressed concern about maintaining continuity within the program:

I think it’s important to work with your colleagues and also to be somewhat on the same page because my, the students that I have now are going to move on. Some of them are going to move on and do the next level so I want them to be graded in a way that’s somewhat similar and you know the content, the topics and the proficiency expectations should be somewhat similar because when they move on and get another teacher they should be on the same page. (NH)

What this participant is describing is a form of standard setting. Continuous change results in shifting goals and makes it impossible to set standards. Standards are necessary in assessment because “It is important to know what any performance on a test actually means, which implies that we need to be able to interpret a test score” (Alderson, 2005, p. 61). In this context each instructor sets his or her own standards and so the meaning (or fairness) of the marks depends entirely on a) the difficulty of the test being taken and b) the interpretation of the results by the instructor.

4.2 Summary of Qualitative Results

Spolsky (1978) describes the job of language testers and test creators as being fundamentally different than in other subject areas because language testers are less likely to be trained in measurement. They are, instead, subject matter (in this case language) specialists. So while modern Western education is highly concerned with formal, objective testing, at the moment the current testing practice in the ASL program reflects what Spolsky (1978) refers to as the “pre-scientific period” (p. v). The program and assessment practices are less concerned with statistics, reliability, and demonstrable validity, and more dependent “on the judgement of an experienced teacher, who can tell after a few minutes' conversation, or after reading a student's essay, what mark to give” (p. v). The qualitative results show that each teacher-participant knows what he or she is looking for when grading; each instructor has a mental map for what it means to be proficient in ASL at the level they teach. A potential problem, as highlighted by the results of interview analysis, is that each teacher is looking at something different. This means that the construct of proficiency is not consistent between instructors, if it is defined at all. Based on teacher-participant responses, results for research question #1 are inconclusive. Teacher-participants in the program all define proficiency differently. Without a common construct definition a description of development cannot occur. This is a key finding, as it suggests a critical area in which ASL assessment practices can be strengthened.

4.2.1 Addressing Research Question #1: What constitutes proficiency development for hearing learners acquiring the language?

Teacher-participants' definitions of proficiency exist alongside divergent teaching and assessment practices. As a whole, courses from first year to fourth year seem to present an increasing level of general second language ability. However, while there is a very loose structure at each level, how the levels build on each other is unclear. Why or how each course contains what it does is unknown. A theoretical basis for why certain things are taught in a particular level is absent. In terms of defining proficiency, this means that while each instructor was able to articulate what proficiency looks like to him or her, the responses were not consistent. Comments from teacher-participants also indicate that the assessment, course content, vocabulary, and expectations change from year to year. Examination of ASL course outlines revealed that only one outline contained projected course outcomes. This, coupled with the fact that the program has no program-wide curriculum, explains why no strong trajectory for student proficiency development became clear: there currently is none. However, responses from teacher-participants did reveal four common indicators of proficiency: facial expression, connectivity, good hold-movement-hold, and low visual noise. The fact that these commonalities exist suggests that there may be an emergent construct for ASL proficiency that warrants further investigation and refinement.

4.2.2 Addressing Research Question #2: How is proficiency determined through current testing practices?

Taken all together, the qualitative results form a matrix of how proficiency is determined through the current assessment practice. The short answer to research

question #2 is that ASL proficiency is determined through expressive language assessments that use communicative elements to capture a demonstration of students' ASL expressive ability. The longer (and much more complex) answer is that instructors use these assessments in order to capture instances of language performance, but subsequent marking of the assessments involves no rubric, and is instead reliant upon each participant's own definition of proficiency. How the teacher-participants engage with proficiency assessments is influenced by their perspectives on students, as well as their feelings towards testing in general. These assessment practices may change from term to term, either as a result of changes implemented by the instructor or through a trickle-down effect where change in one level forces changes in another. The rate of change makes it difficult to discuss with any certainty the effects, or washback, of assessment practices. Without assessment practices that remain consistent over time, there is no way to gauge test reliability.

Now that the qualitative results have been presented and discussed, and research questions one and two addressed, Chapter 4 continues with the presentation and discussion of quantitative results.

4.3 Quantitative Results

A total of 395 questionnaires were collected from student-participants. After discarding 4 unusable surveys (those where students had responded to less than three quarters of the questions, or where the participant had clearly not read the questions and instead responded using a pattern such as always circling 4's) there were a total of 391 surveys remaining. These were entered into a database using the SPSS statistical package (v. 21). The surveys were completed by student-participants in-class in order to obtain as

many results as possible. Students who completed the survey did so on a voluntary basis, and were not compensated in any way. Due to the necessity of securing time and permission through the instructor, two classes were not surveyed, as the instructor indicated that it was not desirable/there was not enough time. I also chose not to survey my own two classes, and so an additional two classes were removed from qualifying.

4.3.1 Why ASL?

As reported in Chapter 3, students majoring in engineering made up 19.4% of the total enrolment by major. The attraction of engineering students to ASL is difficult to explain, though it is possible that they see it as an easy A. A cross tabulation between major and reason for taking the course showed that the response 'I wanted something to help keep my CGPA up' was chosen more often by engineers than by students in any other major (22% of engineering majors chose this response, compared to 12% of psychology majors, the next highest enrolment by major). When looking at why students were taking the course, the most common response was 'Personal interest'. This was followed by (in descending order of choice): It fits my timetable; It was suggested by a friend; I want to minor in a language; It's required for the ASL minor; I need a language credit; I wanted something to keep my CGPA up; Other; It was suggested by an academic advisor. Of those who chose 'Other' as a response, some wrote in that they have a deaf family member, that they plan to use ASL in their future career or studies, or that they simply wanted to fill an open elective. Some indicated that they simply enjoyed the class, with one respondent specifying that their 'passion developed after 1st year'.

4.3.2 Feedback

The majority of students (91.8%, N=391) responded that feedback on tests is important to them (indicated by a Likert scale response of 4 or above). 56.8% selected 6 as their response, indicating the highest level of agreement with the statement ‘Feedback on my tests is important to me’. However, when asked about the amount of feedback they are receiving, only 62.4% responded with a Likert scale choice of 4 or above, which indicates that some students may not be receiving as much feedback as they would like.

As seen in table 3 below, the reported level of feedback differs depending on whether the instructor is Deaf or hearing, with hearing instructors providing significantly ($p<0.001$) more feedback than their Deaf colleagues. This surprising finding suggests that there may be substantive differences between ASL classes, depending on whether an instructor is Deaf or hearing. Although beyond the scope of this thesis, this is a finding worth of further investigation.

Table 3: Cross-tabulation between instructor hearing status and student-reported amount of feedback

	Survey question #25: How much feedback on your tests are you given?						
	1 (None)	2	3	4	5	6 (A lot)	Total
Deaf (signed L1)	26	44	50	41	33	15	209
Hearing (spoken L1)	0	12	15	38	62	55	182
Total	26	56	65	79	95	70	391

4.3.3 Test Format and Method Effect

In the instructor interviews, several teacher-participants discussed what they believe to be a method effect. They reported that video testing, while pragmatic and necessary for testing a visual language, may also contribute to performance anxiety and interfere with accurate proficiency assessment. A slight majority of all students (63.2%,

n=245/388) feel more nervous during a video test than during a written one. Broken down by level, this translates into 66% of first year students, 59.6% of second-year students, 54.5% of third-year, and 47.3% of fourth-year students, indicating a slight downward trend in reported test-nerves as ASL level increases. However, the majority of students 73.4% (n=283/386) responded that they feel just as prepared for a video test as for a written test, and for most (66.2%, n=255/385) not being able to look over their answers at the end of a test is not a concern. Additionally, 70.3% (n=388) of respondents indicated that they are comfortable being filmed, and at the end of a test, 75.5% (n=295/391) feel that they have communicated clearly. If there is a method effect, then according to student responses it is unclear how exactly it is affecting performance. Student-participant responses also reveal that there appear to be few-to-no issues with the video testing system itself, with 89% (n=343/385) of student-participants indicating that it is easy to use. While students may be more nervous during a video test, additional responses to survey questions indicate that for most it is not having an effect on their ability to communicate clearly. As discussed in the qualitative results, it is possible that student anxiety is simply more visible during an ASL assessment, leading instructors to assume that if they can see evidence of nervousness, this must mean that the students are performing badly.

In a related result, 75.3% (n=294/391) of students feel they sign better in class than they do on tests. This is also reported by several of the teacher-participants, and is an interesting phenomenon. Whether this is objectively “true” would need to be investigated further using different methods than those employed in this research. However, there are several things which might be contributing to this perception, particularly instances of

immediate feedback, repair, and support which occur in the classroom. Within the classroom, students' language ability is supported by the instructor and by peers. Either the instructor or peers may supply signs that a student has forgotten, and cue them when they are on the right (or wrong) track. This feedback allows a student to engage in immediate repair or reiteration. Within the classroom, students are working within the Zone of Proximal Development (Vygotsky, 1978), which represents "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). In this study both teachers and students report that students perform better in class. I suggest that a possible reason for this perception is because of the high volume of feedback and available language support from teachers, and also between students themselves, in the classroom. Nodding, confirmation of understanding, (or not understanding), a slower rate of signing from a sympathetic and skilled instructor, and having the teacher or peers fill in incomplete or forgotten signs, all allow a student to do more with their current language level than they would be capable of on their own. All of these factors and supports are important for language learning, but this does not mean that what is seen in the classroom is necessarily an indicator of higher proficiency. All that can be said is that language use in the classroom shows what a student can do with support from a sympathetic and highly fluent instructor who is familiar with their signs and knows what they have been taught. Language skills as observed through video assessment, where a student is producing an unsupported language performance, are possibly more indicative of the proficiency level of a student at the time of assessment. Consideration also needs to be given to class size;

with approximately 40 students in each first-year class¹, it is questionable how accurate a teacher's assessment of an individual student's in-class skill actually is.

4.3.4 Grading

The vast majority of students (93.9%, n=367/391) reported that the tests are directly related to what they are learning in class. This shows agreement with teacher-participant reports that what instructors decide to assess depends on what they have been teaching in class. However, any relationship between in-class learning and grading was not so clear.

74.9% (n=292/390) of students reported that their test grades accurately reflect how much they feel they are learning in their ASL course. This means that 25.1% feel that their test grades do not accurately reflect how much they are learning. The precise relationship between grades and learning within this context is unknown, and the survey question (#25, see appendix for full questionnaire) was not accurate enough to determine what the relationship actually looks like. Do students (those who feel that their grade is reflective of what they are learning) believe that their grade is *equal* to how much they have learned? Do those with lower marks actually feel that they are learning more than what they believe their grades show? Or are those who are satisfied with their grade more likely to report that the grade matches how much they believe they are learning? A very close number of respondents (78.6%, n=307/391) reported that their test grades accurately reflect how they feel they are doing in the course. This shows some indication that students are indeed relating their mark on assessments to their overall performance in the class. This is not unexpected, especially considering reports from teacher-participants

¹ This was the average size of a first-year class during the time this research was conducted. The class size may or may not be the same at the present time.

² I had originally intended to include student-participant interviews in this study, but due to time constraints

which revealed that student assessment is a high priority because it is the only way of providing students with a grade. What comes first: students feeling they are doing well in the course and accepting a grade as being a reflection of that feeling, or seeing a grade and then determining how they feel based on that number? As in the qualitative results, the problem goes back to the meaning of grades, and to the inferences that can be drawn from them – in this case the assumption that grades are an indication of how proficient a student is in ASL. In this context, what a student's grade actually indicates is not known; the assessment process is not currently structured in such a way that a clear connection between the two can be validated.

Almost half (49%, $n=190/388$) of students reported being surprised by their grades and 66.8% ($n=259/388$) felt positively about their tests results. On an additional survey question 77.1% ($n=299/388$) of students reported that they received grades which were higher than expected ($M=4.1$, $SD=1.1$). The inferred relationship then could be said to be that students are more likely to be surprised by their grade because it is higher than expected rather than low. A possible extension of this, which would require further research, is the question: do students who receive higher than expected grades believe that they are learning more than they actually are?

4.3.5 Proficiency Development

Three survey questions attempted to assess students' linguistic understanding of ASL, albeit in a very elementary way. The questions asked students whether they were able to explain or describe three different aspects of ASL: phonology, syntax, and facial grammar. 76.8% of all students surveyed reported that they are able to explain the function of key phonological features of ASL. 72% indicated that they can explain the

basic aspects of ASL structure and syntax to some extent, and 66.5% say they are able to describe the function of ASL grammatical features.

The tables on the following page show the breakdown of responses to these questions by ASL level.

If the responses are examined by level, 29.6% of students in 1010 and 38.5% of students in 1110 (both first-year classes) responded with a Likert scale choice of 5 out of 6, for question #32, and a further 9.3% and 23.1% respectively responded with 6 out of 6, indicating a high level of agreement. Results are similar for questions #33 and #34: first-year students consistently self-report a high level of ASL linguistic understanding. In regard to self-assessment questionnaires, rating scales, and check-lists Oscarson (1989) argues that they are capable of yielding valid results, “although the individual ability for realistic appreciation varies a great deal. Training in using the materials seems to be needed in many cases.” (p. 8). Teacher-participant responses indicated that these types of language abilities are not introduced to students until the second year or later, so it is suspect that so many first-year students responded with such high level of agreement to the statements.

91.7% (n=357/389) of student-participants indicated that they know what aspects of ASL they are being tested on. This is not surprising when considered in light of the interview results which demonstrated that teacher-participants often make it exceptionally clear what will be covered on ASL tests. 87.2% (n=338/388) also responded that they know how to improve their signing proficiency. Given the teacher-participant responses as well as my own experiences, I would argue that the self-reporting of first-year students is beyond what their objective proficiency or understanding is likely

Table 4: Student response by ASL level to questionnaire item #32 (N=388)

	#32. I can explain the function of the key phonological features of ASL (movement, location, hand configuration, etc.)					
	1=Not at all	2	3	4	5	6=Yes, easily
ASL level: 1010	2 (0.8%)	20 (8.1%)	44 (17.8%)	85 (34.4%)	73 (29.6%)	23 (9.3%)
1110	0 (0.0%)	1 (3.8%)	4 (15.4%)	5 (19.2%)	10 (38.5%)	6 (23.1%)
2010	0 (0.0%)	1 (1.6%)	10 (15.9%)	21 (33.3%)	23 (36.5%)	8 (12.7%)
3010	0 (0.0%)	1 (3.0%)	4 (12.1%)	10 (30.3%)	9 (27.3%)	9 (27.3%)
4010	0 (0.0%)	2 (10.5%)	1 (5.3%)	3 (15.8%)	7 (36.8%)	6 (31.6%)

Table 5: Student response by ASL level to questionnaire item #33 (N=386)

	#33. I can explain the basic aspects of ASL structure/syntax					
	1=Not at all	2	3	4	5	6=Yes, easily
ASL level: 1010	5 (2.0%)	31 (12.7%)	49 (20.0%)	80 (32.7%)	61 (24.9%)	19 (7.8%)
1110	0 (0.0%)	0 (0.0%)	6 (23.1%)	6 (23.1%)	8 (30.8%)	6 (23.1%)
2010	0 (0.0%)	1 (1.6%)	10 (15.9%)	20 (31.7%)	26 (41.3%)	6 (9.5%)
3010	1 (3.0%)	0 (0.0%)	4 (12.1%)	2 (6.1%)	16 (48.5%)	10 (30.3%)
4010	0 (0.0%)	1 (5.3%)	0 (0.0%)	1 (5.3%)	11 (57.9%)	6 (31.6%)

Table 6: Student response by ASL level to questionnaire item #34 (N=387)

	#34. I can describe the function of key ASL grammatical features as they appear in questions and sentence topicalization					
	1=Not at all	2	3	4	5	6=Yes, easily
ASL level: 1010	9 (3.7%)	29 (11.8%)	61 (24.8%)	94 (38.2%)	45 (18.3%)	8 (3.3%)
1110	0 (0.0%)	3 (11.5%)	6 (23.1%)	8 (30.8%)	6 (23.1%)	3 (11.5%)
2010	0 (0.0%)	3 (4.8%)	13 (20.6%)	22 (34.9%)	21 (33.3%)	4 (6.4%)
3010	1 (3.0%)	2 (6.1%)	2 (6.1%)	9 (27.3%)	12 (36.4%)	7 (21.2%)
4010	0 (0.0%)	1 (5.3%)	0 (0.0%)	5 (26.4%)	10 (52.6%)	3 (15.8%)

to be. Repeating this type of inquiry and extending it with additional questions would be likely to yield clearer and more accurate results.

4.3.6 The Next Level

60.9% (n=237/389) of student-participants reported that they knew what would be expected of them at the next level, should they choose to continue. This is a difficult response to interpret, since without course outcomes for each level, and different teachers creating different course outlines, it is impossible for students to know what exactly the next level will entail. However, the question itself is vague; ‘know what will be expected of me’ is not clearly defined, and could be interpreted as referring to a number of variables, such as signing ability, assessment, class work, or even general class routines and behaviour.

4.3.7 Percentage of New Material by Level

As seen in table 7 below, there is a steady downward trend in the reported amount of new material as the course level increases. Initially this may seem to fit with general language learning trends which indicate learners make the most progress in language learning from beginner to intermediate stages of learning. However, the question referred to course material, not amount learned. The responses can be taken to mean that there is less new material presented to ASL students as the course level increases. A chi-square test showed that the relationship between these variables was significant $\chi^2(24, N = 391) = 246.17, p < 0.001$.

Table 7: Cross-tabulation between amount of new material and ASL level (N=391)

	#7: How much of the course material is new to you?						
	All	Almost all	Most	Half	Some	A small amount	None
ASL level: 1010	194 (77.9%)	41 (16.5%)	8 (3.2%)	2 (0.8%)	3 (1.2%)	1 (0.4%)	0 (0.0%)
1110	9 (34.6%)	15 (57.7%)	2 (7.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
2010	17 (26.6%)	21 (32.8%)	9 (14.1%)	9 (14.1%)	8 (12.5%)	0 (0.0%)	0 (0.0%)
3010	12 (36.4%)	3 (9.1%)	3 (9.1%)	1 (3.0%)	10 (30.6%)	4 (12.1%)	0 (0.0%)
4010	3 (15.8%)	3 (15.8%)	3 (15.8%)	0 (0.0%)	4 (21.1%)	4 (21.1%)	2 (10.5%)

This may suggest an uneven distribution of new material over the four levels, with increasing redundancy in third year and above, as less and less new material is offered. It is also possible that the program arc is too heavily front loaded, with an abundance of new material being introduced in first and second year, followed by a drastic reduction from then on. However, it is also possible that at more advanced levels a smaller amount of material is covered, or perhaps even re-visited, but at greater depth. While the content of each year may seem different from the teacher-participants' perspectives, student responses indicate high overlap. For teacher-participants who

indicated that what they teach changes every year, and that it tends to get harder and harder, there is a danger that this ‘progress’ may actually hurt students in the long run. Language learners learn at different rates. If teachers or curriculum planners increase the difficulty of their expectations for student achievement at the same level every year, then it’s feasible to think that some students, who may develop very good skills over time, would drop the language because they feel they are not able to keep up. This could also be an explanation of why there is so much overlap – it may be that teachers at levels other than first year realize that students are not as skilled in certain areas as they need to be, and so they re-visit and re-teach them. To students this can appear as ‘old content’ when in fact what teachers are doing is trying to increase student proficiency in something that they may have encountered previously, but not deeply enough. The question then becomes, if first year includes, as one teacher-participant put it, ‘more and more’ how effectively are students really learning? If some students can handle more language input, does that mean the majority of students can?

4.3.8 Assessment Benefits

As discussed earlier, the majority of students surveyed feel that their test and quiz grades accurately reflect how they feel they are doing in the course, and that their grades accurately reflect how much they feel they are learning in the course. This is a key finding, as it shows that students (as would be expected) equate grades with a measure of success. However, as the qualitative results showed, the current ASL assessment practices do not connect grading, final marks, and proficiency in any systematized way. Also, while 77% (n=301/391) of students surveyed reported that ASL assessments are beneficial, nearly half (46.6%, n=182/390) felt that their signing would improve at the

same rate, even if there were no tests. If current assessments are not aiding students in improving their signing then it is unclear what students consider to be the benefit of assessments. Brookhart (2003) argues that assessment is part of instruction, and that “Good classroom assessments are not only measurements of learning, but genuine episodes of learning themselves” (p. 7). In her view, the problem is not the integration of teaching and assessment, but that current measurement theory is not designed to apply to this combined process. For formative assessment, where student learning is being monitored and the educational process developed or improved, classroom assessments make sense. However, if the assessments are not helping in a formative sense, and are not valid indicators of proficiency, then the question is ‘what *are* they good for?’

4.3.9 Productive and Receptive Testing

During the interviews, instructor-participants reported that they engage in productive language testing only. With the exception of one test at the third-year level, testing involves having students use the language in a wholly expressive capacity. Students are not formally assessed on their ability to understand what someone else signs. Compared to a spoken language, this can be considered equivalent to testing speaking skills, but not listening. Survey results showed that 2.3% (n=9/387) of students believe they are assessed on their ability to understand what someone else signs (receptive), 47.2% believe that they are assessed on their ability to sign (expressive), and 49.9% believe that they are assessed on an equal balance of both. At each ASL level, responses were split almost 50-50 between those who felt they were assessed on their expressive skills, and those who felt they were assessed on both expressive and receptive skills. As

described by the instructors themselves, this is simply not true, but it is unclear why students think it is.

4.4 Proficiency Section Factor Analysis

An exploratory factor analysis (EFA) relies on the basic assumption that “within a collection of observed variables, there exists a set of underlying factors, smaller in number than the observed variables, that can explain the interrelationships among those variables” (Pett, Lackey, & Sullivan, 2003, p.4). Four EFA factors emerged from the procedure outlined in Chapter 3. Clustering of proficiency questionnaire items under these factors were examined for similarities in wording, meaning, and rank. Table 8 on the next page shows the pattern matrix that was produced by the factor analysis. (See appendix C for the full questionnaire.)

4.4.1 Factor 1: ASL Discourse Fluency

Factor 1 was represented by nine items, and included those proficiency questions from 18-26. These can be explained as high-level items that represent fluency and increasing ease in participating in ASL discourse. The statements on the questionnaire required high-level ASL skills, such as the ability to create properly structured and coherent ASL discourse on almost any topic; maintaining grammatical control of complex signing; and following complex and extended signing.

4.4.2 Factor 2: Basic ASL Communication

This factor included seven items, represented by questionnaire items from 1-7. This factor can be explained as basic communication skills, since all of the statements ask students to gauge their ability to function using low-level ASL abilities. These questions included many skills that students are accustomed to using in the classroom, and which

Table 8: Pattern Matrix for 4-Factor Solution (EFA)

<i>Item #</i>	<i>Statement/Question</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>
24	Create properly structured and coherent discourse in ASL on almost any topic, using a wide range of connectors and varying organizational patterns	.984			
25	Maintain consistent grammatical control of complex signing, even when my attention is otherwise engaged (thinking ahead, making note of others' reactions, etc.)	.957			
23	Understand a native signer who is signing to me quickly, using as many colloquialisms as they would when communicating with another native signer	.856			
22	Express myself in ASL on a wide range of topics fluently and spontaneously, using clear descriptions and themes	.823			
26	Take part in a formal job interview in ASL where I am required to give accurate information about my interests, qualifications, and experience	.789			
21	Develop well-structured points and conclusions as part of an argument/discussion in ASL, without having to obviously search for expressions	.730			
20	Understand fingerspelling at full speed, as long as I know what the context is.	.570			.406
19	Follow complex and extended signing, as long as the topic is familiar to me (for example, a signed lecture or speech on a topic that I am familiar with)	.454			
18	Appropriately interject in a signed conversation in order to give my opinion on a topic that I am somewhat familiar with.	.388		.378	
4	Make myself understood using short, simple signed sentences and gestures when I am describing things I am very familiar with		.787		
5	Ask and answer questions in ASL about personal details in a simple way		.783		
3	Create a short monologue describing myself in ASL		.735		
2	Understand simple statements or questions in ASL ("Hello", "How are you?", "What's your name?", "It's cold today", etc.)		.706		
1	Recognize familiar and very basic signs, such as those concerning myself, my family, and my immediate surroundings		.621		
7	Use a series of signed sentences to describe everyday situations/places (i.e. daily routines, my home, my educational background, etc.)		.552		
6	Understand high-frequency signed phrases that are used often in the classroom		.403		
10	Buy clothes in a store where all the staff only sign.			.898	
13	Give a brief but accurate signed description of a photograph or painting while looking at it.			.871	
16	Understand the main points of a news broadcast in ASL			.715	
15	Give directions in sign to someone on the street with only slight hesitations and rephrasing			.654	
17	Order a complete meal in a restaurant in ASL			.602	
11	Describe my favourite hobby in ASL in some length, using appropriate vocabulary			.601	
12	Understand fingerspelling when it is slow and deliberate, and when the word or name being spelled is familiar to me				.798
8	Understand a native signer who is signing to me slowly and carefully (i.e. someone who is deliberately adapting their signing to suit you)				.555
9	Understand most short, simple, personal ASL narratives about basic subjects (shopping, travel, work, etc.)				.374
14	Repeat back part of what someone has signed to confirm mutual understanding				.346
<i>Eigenvalues</i>		10.55	3.51	1.36	1.08
<i>Percentage of variance explained</i>		39.04	11.85	3.67	2.21

Extraction Method Principal Axis Factoring
Rotation Method Promax with Kaiser Normalization

they are frequently exposed to. This includes asking and answering personal details in a simple way; understanding simple statements or questions; and using signed sentences to describe everyday situations or places.

4.4.3 Factor 3: Language use Context

The six items that represent factor 3 all involve using ASL in a context outside of anything currently experienced or practiced in the classroom. Three of the questions (10, 15, and 17) involve using ASL in an interactive context with other (possibly native) signers – buying clothes, giving directions, and ordering a meal. Question #11 asked student-participants about their ability to describe a favourite hobby in some length. Question #13 asked whether they could accurately describe a photograph or painting while looking at it. Question #16 asked about whether a student could understand a news broadcast in ASL. All six of these questions require students to imagine themselves in language-use situations that are not currently covered in class. In each of these cases, students would need to transfer their language skills to contexts that they may not have experience with; the questions ask students to gauge their ability to use ASL in a real life situation. These questions also all necessitate using the language in a more meticulous and focused way; questions include qualifiers such as ‘using appropriate vocabulary’, ‘accurate’, and ‘with only slight hesitations’. As a group, all of the questions that cluster under this factor require that students be accurate in their language use or understanding.

4.4.4 Factor 4: Comprehension

Factor four is only represented by four items, making it the weakest factor. However, all four items can be classified as ‘comprehension skills’. All of the

statements that grouped under this factor ask students to gauge their ASL comprehension; the word “understand” or “understanding” appears in each statement. This factor also co-loads onto question 20 (high proficiency) which examines comprehension as well. Despite the fact that this is the weakest factor in terms of the amount of variance it explains, this could be an artefact of the number of items (only 4) associated with it on the questionnaire. Regardless, this is an important finding as it indicates that comprehension is a separate factor in ASL proficiency.

4.4.5 Factor Analysis Summary

The results of the factor analysis show that what accounts for variance in student-participants self-assessment of proficiency are four factors: ASL discourse fluency; basic ASL communication; language transfer; and ASL comprehension. I recognize that the labels applied to the first two factors are theoretically different than the latter two; discourse fluency and basic ASL communication could just as easily be (and initially were) labelled as “high proficiency” and “low proficiency”. My goal in using an EFA was to see what underlying components emerged in relation to the more general concept of proficiency, but after closely examining factors one and two it was clear that there were too many possible differences in what the questions were asking to actually provide a more appropriate label. It is quite likely that the self-assessment was not always asking the right questions, and was not sensitive enough to draw out more of the explicit factors that underlie student proficiency development. That being said, the latter two (language use context and comprehension) are important areas of language skill development. The fact that they are two of the factors that most account for variance indicates that they are possibly not as well-integrated into ASL instruction as they could be. This is a key

finding, as these are also two areas of concern reported by teacher-participants. It seems that instructors are right to be concerned about their students' ASL comprehension skills, and to wonder whether current testing practices are providing accurate information on how students would perform outside of the classroom.

These results included students at different levels of the ASL program. Since it is possible that students at different levels may have responded differently to questions of proficiency, the factor analysis was run again, this time examining the total participants as two separate groups: first-year students were analysed as one group, and then second, third, and fourth-year students as a second group. Second, third, and fourth were kept together because there were not enough participants at each of those levels separately to make up a large enough sample size.

When second, third, and fourth-year were analysed, the same four factors were identified, and they occurred in the same order of weight, with slightly different percentages of variance attributed to each factor. Three questionnaire items loaded under different factors: Question #9 and #14, which had previously clustered under factor 4 (comprehension), and #11 (previously under new context) now group with factor 2. Question #18 moved from factor 1 (high proficiency) to factor 4 (comprehension). This seems to indicate a shift in proficiency as ASL level increases. Things that were once new or strictly comprehension-based become integrated with basic skill; things that were previously high-level become more clearly identified as skills that call for comprehension. The 'new context' factor becomes less heavily weighted, which may indicate that new contexts are introduced as ASL level increases.

Analysing first-year participants alone resulted in a pattern matrix that was very similar. Questionnaire items loaded under the same factors, the difference being that the weighting of the factors (the amount of variance each one explained) changed. The order for all of the participants as a whole is: ASL discourse fluency; basic ASL communication; new context language use; and ASL comprehension. For first-year alone, the factor identified as 'new language context' accounted for 3.15% of the variance, rather than 1.36%.

4.5 Quantitative Results Summary

The results of the quantitative analyses provide triangulation for the findings of the qualitative results. More specifically, when it comes to ASL proficiency the results of the proficiency section factor analysis show that ASL comprehension and language transfer are two critical areas that account for a large percentage of the diversity in student proficiency. The questionnaire analysis also resulted in several other findings: little evidence to support teacher-participants belief in a test-method effect; a significant feedback difference between Deaf and hearing instructors; a reported drop in new material as ASL level increases; and a complex and unclear relationship between grades, learning, and test benefit.

While students report that what they are assessed on is related to what they learn in class, the relationship between assessment, grades, and perceptions of learning are not clear. What is more clear is that students report a significant drop in new course content as ASL level increases. This may indicate that the same or similar material is covered again (but more deeply) in higher levels, or that there is an uneven distribution of course material over the course of the program arc. Student responses indicated that there is a

significant difference in amount of feedback provided between Deaf and hearing instructors. More feedback does not necessarily mean better feedback (Lepper, Drake, & O'Donnell-Johnson, 1997). The type of feedback being offered by each instructor, and student responses to that feedback, would require further investigation.

4.5.1 Addressing Research Question #3: How well do students understand the construct of proficiency, and how do they respond to the current testing practice?

The results of the questionnaire indicate that students generally respond well to the current testing practices, at least from a practical and technical standpoint. There appear to be no problems with the testing system and most student-participants agreed that the assessments are beneficial to their learning process, even if they also indicated that they feel their signing would improve at the same rate without assessment.

Whether students understand what it means to be proficient in ASL is very much open for debate. The questionnaire results suggest that most student-participants believe they understand ASL phonology, syntax, and the use of facial grammar. This is true even of first-year (first semester) students, even though teacher-participants report that these concepts are not introduced until later levels. While it is not impossible for a first-year student to have a grasp of ASL linguistics, it is unlikely, particularly given the comment by a teacher-participant that even by fourth-year, students' knowledge of ASL linguistics is 'layman'.

The truly interesting follow up question is why did these responses occur in the way they did? To recap: ASL assessments are all expressive (productive) only. They measure a student's ability to produce signs and to communicate information in ASL. Understanding of linguistics - the grammatical rules of ASL - is not assessed outside of

those performances. If it is never assessed, then how are students so sure of what they know? Do students believe they have a deeper understanding of ASL than they actually do? Without assessments that measure areas of student proficiency outside of their ability to produce (usually rehearsed) ASL performances, it is impossible to compare what students *believe* they know with what they are able to *show* they know. In this context much of the true picture of student proficiency is unaccounted for.

This concludes the presentation and discussion of results. This aim of this chapter has been to offer detailed results from both the qualitative and the quantitative sections of the research while remaining focused on those findings that best offer answers to the research questions. I have chosen to combine the results and discussion rather than allotting them to separate chapters. This format is better able to address and contextualize the results as they are presented. This also allows different elements from the research to be drawn together in order to provide a deeper analysis of why the results may have occurred in the way they did, and what this might mean about ASL assessment in this context.

Early research on ASL tended to focus on providing linguistic analyses of the language in order to document and eventually prove its status as a naturally occurring and fully-formed human language. Studies of the linguistics of ASL continue into the present day, and have been joined (as outline in the literature review) by research that investigates the cognitive processing of ASL. Recently, studies have begun to explore the pedagogy of ASL in terms of how it is taught as a second language. The present study focuses on what it means for a hearing learner of ASL to be proficient in the language at different stages of learning, how that proficiency is assessed, and how students respond to

the assessments. ASL is being taught to an increasing number of hearing learners. The type of research that I have engaged in here is not only timely, but necessary. It is crucial that the construct of proficiency for hearing learners become better understood, defined, and operationalized in order to strengthen the validity of ASL assessments. This will ensure that the inferences about individuals' language abilities that are drawn from assessments are adequately and appropriately (Messick, 1989b) supported.

Chapter 5: Conclusion

This thesis was guided by three research questions: 1) What constitutes ASL proficiency for hearing learners acquiring the language in this context? 2) How is proficiency determined through current assessment practices? 3) How well do students understand the construct of proficiency, and how do they respond to current testing practices? Before addressing final conclusions, there are a few limitations that need to be mentioned.

5.1 Limitations

1. This study is context-dependant. While the number of participants (both teacher-participants and student-participants) was high, the results cannot be generalized to a wider population. The outcomes of this study may be quite different if recreated within another program.
2. Results from the proficiency section of the questionnaire were skewed towards first-year students, since they are the most heavily represented portion of the ASL program population. A higher number of responses (perhaps collected over time, if necessary) from more students in later years would be useful for subsequent research. Based on present results, further research would also need to refine the proficiency section in order to continue teasing apart where differences in student proficiency and proficiency development occur. Had I had time to pilot this section and run an initial EFA first, I would have been better equipped to then re-work the proficiency statements, which may have lead to improved results.

3. I was not able to complete the classroom observations, which would have added more information to the qualitative section of this study, and may have provided some insights into teacher-participants classroom practices.

4. Since this study used quantitative data to collect student responses it was able to provide a snapshot of the student population. This could be strengthened with qualitative research (for example, interviews or focus groups) that includes students at each level of the program.²

5. As an instructor in the program I was afforded more access to teacher-participants than other ‘outsiders’ may have been. During interviews teacher-participants were open and candid with me. Because of my own involvement in the program and my relationship to my participants, at times I had to think carefully about what information to include in the research. I had to consider not just the privacy concerns of my participants but also the reputation of the program. I have done my best to balance my own research needs with concerns for my participants and the program – without misrepresentation or cutting out critical information from participant responses. I was not at arm’s length from this research. While I do not see this as a limitation per se it is important to point out that another researcher, without the same connection to the program and participants, may not have ended up with the same results.

5.2 Conclusions About ASL Proficiency Development

While the results of this research were not able to completely identify what constitutes proficiency development for hearing learners of ASL, results did reveal that there are some aspects of ASL proficiency which are recognized by all teacher-

² I had originally intended to include student-participant interviews in this study, but due to time constraints and insufficient participants this was not possible.

participants (facial expressions, connectivity, good hold-movement-hold, and low visual noise). Additionally, comparison of qualitative and quantitative data from this research reveals that teachers express concern over some of the same proficiency aspects that actually separate students: ASL comprehension, and the ability to transfer language to situations outside of the classroom. Both of these are critical skills for language learning and language development. This reveals an emerging ASL proficiency construct for hearing learners that can be expanded and developed into a stronger definition of ASL proficiency for this group of learners. This will require additional research into ASL proficiency for hearing learners in order to better understand the acquisition patterns and learning needs of this group.

This research shows that ASL comprehension and language transfer warrant further investigation and improvement. A first step would be addressing the lack of outcomes at each level by creating clear and achievable goals for proficiency that begin in first year and are systematically built and expanded upon until the end of fourth year. These goals could be developed in-house, but may benefit from examining the standards that have already been developed by associations such as the ASLTA. If the ASL program considered in this study continues to expand, clear learning outcomes will be increasingly necessary. Course outcomes can provide unification and standardization across levels with different instructors, and ensure that each level connects to and expands on the level before it in as seamlessly a way as possible. Course outcomes also ensure that key concepts are taught and assessed by all instructors, so that students are introduced to the same material no matter who the instructor may be. While the program is currently textbook-free, and allows a classroom curriculum to be created by each

teacher, there may come a point where that is no longer pragmatic or sustainable, and this research suggests that that point may have arrived.

5.3 Conclusions About ASL Assessment Validity

The assessment practices of the program currently permit a lot of individual instructor flexibility. This allows assessments to be connected to what is taught in class, but it occurs on a class by class basis. The assessments, and the inferences drawn from them, are not generalizable beyond the classroom. Also, without formal assessment of ASL comprehension, current assessment strategies are only examining half of the skills necessary to function in a visual language. The assessment practices in the program currently rely on what Bachman (2000) terms ‘communicative’, ‘performance’, or ‘alternative’ assessment types. He notes that performance assessments are often associated with concerns about the reliability and validity of their measurement qualities, and argues that they “need to satisfy the same standards for assessment as do any other type of assessment” (p. 14). Learning ASL as an L2 or M2 can be difficult; learners are brought into a context where they are dependent on their eyes and hands more than their ears and voices. However, until there is strong evidence that ASL comprehension is objectively ‘harder’ than spoken language comprehension, this is an area that needs to be included in language assessment at every level. Student-participant results show that almost half of all students believe that they are assessed on both expressive and receptive ASL ability, even though they are not.

As discussed in Chapter 2, validity can be approached in a variety of ways. In this context it is construct validity which needs to be strengthened first so that the tests can actually claim to act as a measure of a particular attribute (Cronbach & Meehl, 1955)

(proficiency), with the result that the interpretations made of the basis of those tests are appropriate (Bachman & Palmer, 1996). Chapter 2 also highlighted research which showed that hearing learners are not the same as deaf users of ASL. Much of the research on ASL to date examines deaf signer. Hearing learners need to be studied as a distinct group, and material, teaching methods, and testing practices developed that support their learning needs. The construct of proficiency must be defined for *this group specifically*.

While there may not be one agreed upon definition of L2 proficiency in the broader field of SLA research, a definition - specific to ASL - needs to be selected and followed consistently. Teachers working within the program need to come to a consensus regarding how they determine proficiency, and what the expectations for proficiency are at each level. Developing a clear construct definition for proficiency will allow measurements of proficiency to be more generalizable across and between levels. This in turn will provide evidence and justification for assessment scores. Any assessments created also need to be regularly examined for evidence of validity, so that it is ensured that tests are measuring what they purport to measure.

5.4 Consequences of Current Assessment Practices

ASL assessments are intended to discern what students can do with ASL. They are an opportunity for students to show their language ability, and for instructors to offer meaningful feedback that will aid the students in increasing their proficiency and thinking more critically about their language use and language learning. Assessments are also an opportunity to show students where they stand within a proficiency construct; to point out ‘you are here’ as a way of indicating their current threshold of language ability, indicating how far they have come from the beginning, and also how far they may still

have to go to reach their intended language goal(s). In this sense, a student's grade should be representative of what they have achieved on that criterion scale. That scale needs to be carefully developed, and should include a much broader range of ASL skills than are currently being assessed.

If instructors are teaching to the test, or overcompensating for a perceived effect of student anxiety or lack of practice, then the results of the test are not a valid indication of language proficiency. This may be at least part of the explanation as to why teachers are concerned about students' ability to function in contexts outside of the classroom.

When the assessment construct or content is too narrowly defined, the intended relationship between test and outcome falls apart. If the goal of a course is developing proficiency to level x, with proficiency defined as w, x, y, z, but only the specific structures of w and x are assessed, then the test content is too narrow. This, combined with instances of coaching or teaching to the test, further threaten the validity of the assessment (Shepard, 1997). Without a defined construct for proficiency, what assessments are actually measuring is unclear. Teachers each develop their own assessments, unsure of whether these are supplying them with valid results.

Consequently, students complete a course assuming that the grade they received is an indication of their skill level when in fact what the grade actually signifies is unknown.

As assessment practices within the program currently stand, teachers are not able to provide students with "reasonable assurances that the inferences and decisions we make on the basis of their test scores are useful" (Bachman, 2005, p. 2). Without tracking grades from one class to the next, it is impossible to say that students with high assessment scores perform substantially better than those with lower scores. Since

instructors are assessing proficiently differently, a high final grade in one class will not necessarily mean a student is better prepared for the proficiency expectations of the next class.

5.5 Future Directions

In the future, research should include other ASL programs (teachers and students) to provide a better idea of how different programs (as well as practices and responses) across Canada compare to each other. Expanding the interviews to include teacher-participants from other ASL programs, and a random sampling of student-participants, would provide information on similarities and differences across programs and help clarify curricular focus and assessment practices.

In the literature review it was briefly mentioned that there has been very little research done that analyzes the pragmatics of ASL. Though this was only touched on, further research in this area could prove to be highly important for examining and understanding ASL second language proficiency development in hearing learners, particularly at later stages of language learning.

Two secondary findings of this research also need to be pursued in greater depth and detail. The first is the reported difference (by students) in the amount of feedback given by deaf compared with hearing instructors. Why are the feedback amounts significantly different? Are there differences in the type of feedback that these two groups tend to give? How do these difference effect students? The second finding, which could be pursued in a follow-up study, is method effect/testing anxiety. While student responses did not reveal strong evidence for a test method effect, they did indicate that students are more nervous during a video test than during a written one. Since a visual

language requires a visual assessment strategy, further exploration of the effects of video testing is important.

The body of research associated with ASL as a second language for hearing learners is small, but growing. There is ample potential to engage in research that may be used to develop new teaching methods, gather more information about L2/M2 acquisition, and create ASL assessments. There are only a handful of universities in Canada that currently offer ASL as a minor, but interest in the language continues to grow. There is an opportunity for future research to be carried out within this program – to use it as the grounds from which to more deeply investigate the construct of proficiency for hearing learners of a second modality second language.

This thesis allowed me to engage more deeply with principles and practices in second language assessment in general, and to begin to think about how these may be applied to ASL. Through this process I have become more confident in my role as an instructor and test creator, but I also see with increasing clarity the difficulties and intricacies involved in evaluating second-language proficiency successfully. Reading other people's research and writing on similar, or related, matters and subsequently engaging in my own research has also convinced me that it is crucial to continue to explore what it means for hearing learners to be (or become) proficient in ASL and how this proficiency is being assessed.

References

- Alderson, C. J. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Alderson, C. J. (2009). The micropolitics of research and publication. In J. C. Alderson (Ed.), *The politics of language education: Individuals and institutions* (pp. 222-236). (New perspectives on language and education). Bristol: Multilingual Matters.
- Ashton, G., Cagle, K., Kurz, K. B., Newell, W., Peterson, R., & Zinza, J. E. (2014). *Standards for learning American Sign Language*. American Council of the Teaching of Foreign Languages (ACTFL): Publication forthcoming. Retrieved August 10, 2014 from http://asltla.org/wp-content/uploads/2014/07/National_ASL_Standards.pdf
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing*, 17(1), 1-42.
- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly*, 2(1), 1-34.
- Bachman, L. F., & Clark, J. L. D. (1987). The measurement of foreign/second language proficiency. *Annals of the American Academy of Political and Social Science*, 490, 20-33.
- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.
- Bachman, L. F., & Savignon, S. J. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *The Modern Language Journal*, 70, 380-390.
- Baker, B. A. (2012). Individual differences in rater decision-making style: An exploratory mixed-methods study. *Language Assessment Quarterly*, 9, 225-248.
- Battison, R. (1978). *Lexical borrowing in American Sign Language*. Silver Spring, MD: Linstok Press.
- Bellugi, U., Klima, E. S., & Siple, P. (1975). Remembering in signs. *Cognition*, 3, 93-125.
- Bochner, J. H., Christie, K., Hauser, P. C., & Searls, J. M. (2011). When is a difference really different? Learners' discrimination of linguistic contrasts in American Sign Language. *Language Learning*, 61 (4), 1302-1327.
- Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement: Issues and Practice*, 22(4), 5-12.
- Buck, G. (2009). Challenges and constraints in language test development. In J. C. Alderson (Ed.), *The politics of language education: Individuals and institutions* (pp. 166-184). (New perspectives on language and education). Bristol: Multilingual Matters.
- Byrnes, H. (1987). Proficiency as a framework for research in second language acquisition. *The Modern Language Journal*, 71(1), 44-49.

- Canale, M., & Swaine, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Carroll, B. J. (1978). *Guidelines for the development of communicative tests*. London: Royal Society of Arts.
- Cerney, J. (2007). *Deaf education in America: Voices of children from inclusion settings*. Washington, DC: Gallaudet University Press.
- Chapin, P. G. (1988). American Sign Language and the liberal education. *Sign Language Studies*, 59, 109-113.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Charmaz, K. (2008). Constructionism and the grounded theory method. In J. A. Holstein & J. F. Gubrium (Eds.), *Handbook of constructionist research* (pp. 397-412). New York: Guilford Press.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1980). *Rules and representations*. Oxford: Basil Blackwell.
- Cooper, S. B., Reisman, J. I., & Watson, D. (2008). The status of sign language instruction in institutions of higher education: 1994-2004. *American Annals of the Deaf*, 153(1), 78-88.
- Cooper, S. B., Reisman, J. I., & Watson, D. (2011). Sign language program structure and content in institutions of higher education in the United States, 1994-2004. *Sign Language Studies*, 11(3), 298-328.
- Council of Europe. (2001). *Common European Framework of References for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

- Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Los Angeles, CA: Sage.
- Creswell, J. W., & Plano Clark, V. L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.
- Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed.) (pp. 443-407). Washington, DC: American Council on Education.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Davies, A. (1990). *Principles of language testing*. Cambridge, MA: Basil Blackwell Ltd.
- De Jong, J. H. A. L. (1990). Response to Masters: Linguistic theory and psychometric models. In J. H A. L de Jong & D. K. Stevenson (Eds.), *Individualizing the assessment of language abilities* (pp. 71-82). Clevedon: Multilingual Matters.
- Dey, I. (2007). Grounding categories. In A. Bryant & K. Charmaz (Eds.), *The SAGE handbook of grounded theory* (pp. 167-190). (2010 paperback edition). Los Angeles and London: SAGE Publications.
- Djonov, E. & Zhao, S. (2014). From multimodal to critical multimodal studies through popular discourse. In E. Djonov & S. Zhao (Eds.), *Critical multimodal studies of popular discourse*, (pp. 1-14). New York: Routledge.
- Doré, C. (2014). *Exploring teacher perceptions of a textbook-free ASL curriculum at a Canadian university: A qualitative combined methods study*. Unpublished master's thesis. Carleton University, Ottawa Ontario, Canada.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University

Press.

Edwards, R. A. R. (2012). *Words made flesh: Nineteenth-century deaf education and the growth of deaf culture*. New York, NY: New York University Press.

Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.

Emmorey, K. (1993). Processing a dynamic visual-spatial language: Psycholinguistic studies of American Sign Language. *Journal of Psycholinguistic Research*, 22(2), 153-187.

Emmorey, K., & Corina, D. (1990). Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills*, 71, 1227-1252.

Fairbairn, S. B., & Fox, J. (2009). Inclusive achievement testing for linguistically and culturally diverse test takers: Essential considerations for test developers and decision makers. *Educational Measurement: Issues and Practice*, 26(1), 10-24.

Furman, N., Goldberg, D., & Lusin, N. (2010). *Enrollments in languages other than English in United States institutions of higher education fall 2009*. New York: Modern Language Association of America. ERIC database: ED513861.

Gorsuch, R. (1983). *Factor Analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Haug, T. (2005). Review of sign language assessment instruments. *Sign Language & Linguistics*, 8, 59-96.

Hole, R. (2007). Working between languages and cultures: Issues of representation,

- voice, and authority intensified. *Qualitative Inquiry*, 13, 696-710.
- Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly*, 8, 229-249.
- Hulstijn, J.H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15, 422-433.
- Isabella, L. A. (1990). Evolving interpretations as a change unfolds: How managers construe key organizational events. *The Academy of Management Journal*, 33(1), 7-41.
- Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29(1), 24-49.
- Jacobowitz, E. L. (2005). American Sign Language teacher-preparation programs in the United States. *Sign Language Studies*, 6(1), 76-110.
- Jewitt, C. (2007). A multimodal perspective on textuality and contexts. *Pedagogy, Culture, & Society*, 15, 275-289.
- Kanda, J. & Fleischer, L. (1988). Who is qualified to teach American Sign Language? *Sign Language Studies*, 59, 183-194.
- Kathleen, S., & Kitao, K. (1996). Testing communicative competence. *The Internet TESL Journal*, 2(5). Retrieved July 26, 2014, from <http://iteslj.org/Articles/Kitao-Testing.html>
- Kelle, U. (2007). The development of categories: Different approaches in Grounded

- Theory. In A. Bryant & K. Charmaz (Eds.), *The SAGE handbook of Grounded Theory* (pp. 119-213). Los Angeles, CA: SAGE.
- Kemp, M. (1998). Why is learning American Sign Language such a challenge? *American Annals of the Deaf*, 143(3), 255-259.
- Klein, W. (1998). The contribution of second language acquisition research. *Language Learning*, 48, 527-550.
- Klima, E., & Bellugi, U. (1979). *The signs of language*. Massachusetts: Harvard University Press.
- Kline, P. (1993). *The handbook of psychological testing*. New York, NY: Routledge.
- Knigga, M. E. (2012). *The role of language program directors in the articulation of American Sign Language foreign language programs* (Unpublished doctoral dissertation). West Lafayette, IN: Purdue University
- Kramsch, C. (1997). Guest column: The privilege of the nonnative speaker. *PMLA*, 112, 359-369.
- Kress, G., Jewitt, C., Bourne, J., Franks, A., Hardcastle, J., Jones, K., & Reid, E. (2005). *A multimodal perspective on teaching and learning*. London: RoutledgeFalmer.
- Lane, H., Hoffmeister, R., & Bahan, B. (1996). *A journey into the Deaf-world*. California: Dawn Sign Press.
- Lado, R. (1961). *Language testing*. London: Longman.
- Lange, D. L. (1982). The problem of articulation. In T. V. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 113-137). Skokie, IL: National Textbook Co.
- Lepper, M. R., Drake, M. F., & O'Donnell-Johnson, T. (1997). Scaffolding techniques of

- expert human tutors. In K. Hogan & M. Pressley (Eds.), *Scaffolding student learning: Instructional approaches & issues* (pp 108-144). Cambridge, MA: Brookline Books.
- Liddell, S. K., & Johnson, R. E. (1989). American Sign Language: The phonological base. *Sign Language Studies*, 64, 195-278.
- Llurda, E. (2004). Non-native speaker teachers and English as an international language. *International Journal of Applied Linguistics*, 14, 314-323.
- Long, M. H., Gor, K., & Jackson, S. (2012). Linguistic correlates of second language proficiency: Proof of concept with ILR 2-3 in Russian. *Studies in Second Language Acquisition*, 34, 99-126.
- Lupton, L. K., & Zelaznik, H. N. (1990). Motor learning in sign language students. *Sign Language Studies*, 67, 153-174.
- Maller, S. J., Singleton, J. L., Supalla, S. J., & Wix, T. (1999). The development and psychometric properties of the American Sign Language proficiency assessment (ASL-PA). *Journal of Deaf Studies and Deaf Education*, 4, 249-269.
- Mayberry, R. I., & Fischer, S. D. (1989). Looking through phonological shape to lexical meaning: The bottleneck of non-native sign language processing. *Memory and Cognition*, 17, 740-754.
- McDermid, C. A. (2012). *Adult second language pragmatic enrichment: The case of ASL*. Unpublished doctoral dissertation. York University, Toronto Ontario, Canada.
- McKee, R. L., & McKee, D. (1992). What's so hard about learning ASL?: Students' and teachers' perceptions. *Sign Language Studies*, 75, 129-158.

- McNamara, T. & Roever, C. (2006). *Language testing: The social dimension*.
Massachusetts: Blackwell Publishing, Ltd.
- Medgyes, P. (1994). *The non-native teacher*. London: Macmillan.
- Messick, S. (1989a). Validity. In R. Linn (Ed.), *Educational Measurement*. New York: Macmillan, pp. 13-103.
- Messick, S. (1989b). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Metzger, M., & Bahan, B. (2001). Discourse analysis. In C. Lucas (Ed.), *The sociolinguistics of sign languages* (pp. 112-144). Cambridge: Cambridge University Press.
- Miller, K. R. (2008). American Sign Language: Acceptance at the university level. *Language, Culture and Curriculum*, 21, 226-234.
- Morford, J. P., Grieve-Smith, A. B., MacFarlane, J., Staley, J., & Waters, G. (2008). Effects of language experience on the perception of American Sign Language. *Cognition*, 109, 41-53.
- Munby, J. (1978). *Communicative syllabus design*. Cambridge: Cambridge University Press.
- Newell, W. J. (1995). Competencies important to teaching ASL: Perceptions between groups. *Sign Language Studies*, 89, 303-330.
- Newkirk, D., Klima, E., Peterson, C. & Bellugi, U. (1980). Linguistic evidence from slips

- of the hand. In V. Fromkin (Ed.), *Errors in linguistic performance* (pp. 165-197). New York, NY: Academic Press.
- North, B. (1995). The development of a common framework scale of descriptors of language proficiency based on a theory of measurement. *System*, 23, 445-465.
- Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language Testing*, 6(1), 1-13.
- Olsen, G. (1988). Acceptance of American Sign Language: An American ground swell. *Sign Language Studies*, 59, 100.
- Paesani, K., & Barrette, C. (2005). The role of the language program director within a three-dimensional model of articulation. In K. Paesani & C. Barrette (Eds.), *Language program articulation: Developing a theoretical foundation* (pp. 2–20). Boston, MA: Thomson Heinle.
- Peterson, R. (2009). *The unlearning curve: Learning to learn American Sign Language*. Burtonsville, MD: Sign Media Inc.
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: SAGE.
- Pichler, D. C. (2011). Sources of handshape error in first-time signers of ASL. In G. Mathur & D. J. Napoli (Eds.), *Deaf around the world: The impact of language* (pp. 96-121). Oxford: Oxford University Press.
- Pienemann, M. (1985). Learnability and syllabus construction. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition* (pp. 23-75). Bristol, UK : Multilingual Matters .

- Poizner, H. (1983). Perception of movements in American Sign Language: Effects of linguistic structure and linguistic experience. *Perception & Psychophysics*, 33, 215-231.
- Quinto-Pozos, D. (2011). Teaching American Sign Language to hearing adult learners. *Annual Review of Applied Linguistics*, 31, 137-158.
- Redden, E. (2007). More Students, More Languages. Inside Higher Ed. Retrieved June 6, 2014, from <https://www.insidehighered.com/news/2007/11/14/languages>
- Rosen, R. (2010). American Sign Language as a foreign language in U.S. high schools: State of the art. *The Modern Language Journal*, 92, 10-38.
- Roush, D. (1999). *Indirectness in American Sign Language: Requests and refusals*. Department of ASL, Linguistics and Interpretation Working Papers 1999. Washington, DC: Gallaudet University Press.
- Savignon, S. J. (1983). *Communicative competence: Theory and classroom practice*. Reading, MA: Addison-Wesley.
- Saldaña, J. (2013). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage
- Sacks, O. (1989). *Seeing voices*. Los Angeles, CA: University of California Press.
- Schornstein, R. A. (2005). Teaching ASL in the university: One teacher's journey. *Sign Language Studies*, 5, 398-414.
- Shepard, L. A. (1997). The centrality of test use and consequences for test validity. *Educational Measurement: Issues and Practice*, 16(2), 5-24.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.

- Siple, P., Fischer, S., & Bellugi, U. (1977). Memory for nonsemantic attributes of American Sign Language sign and English words. *Journal of Verbal Learning & Verbal Behavior*, 16, 561-574.
- Slobin, D. I. (2008). Breaking the molds: Signed languages and the nature of human language. *Sign Language Studies*, 8, 114-130.
- Spolsky, B. (1968). Language testing: The problem of validation. *Language Testing*, 2(2), 88-94.
- Spolsky, B. (1978). Introduction: Linguists and language testers. In B. Spolsky (Ed.), *Approaches to language testing* (pp v-x). Arlington, VA: Center for Applied Linguistics.
- Stedman, C. H. (1984). Testing for competency: A pyrrhic victory? *Journal of Teacher Education*, 35(2), 2-5.
- Stevenson, D. K. (1981). Beyond face validity: The multitrait-multimethod matrix and the convergent and discriminant validity of oral proficiency tests. In A. S. Palmer, P. J. M. Groot, & G. A. Trosper (Eds.), *The construct validation of tests of communicative competence* (pp. 37-61). Washington, DC: Teachers of English to Speakers of Other Languages.
- Supalla, T., Newport, E., Singleton, J., Supalla, S., Coulter, G., & Metlay, D. (in press). The test battery for American Sign Language morphology and syntax. San Diego, CA: DawnSign Press.
- Taub, S. F. (2001). *Language from the body: Iconicity and metaphor in American Sign Language*. Cambridge: Cambridge University Press.
- Tevenal, S. & Villanueva, M. (2009). Are you getting the message?: The effects of

- SimCom on the message received by deaf, hard of hearing, and hearing students. *Sign Language Studies*, 9, 266-286.
- Thoryk, R. (2010). A call for improvement: The need for research-based materials in American Sign Language education. *Sign Language Studies*, 11, 100-120.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14, 281-294.
- Trompenaars, F., & Hampden-Turner, C. (1998). *Riding the waves of culture: Understanding cultural diversity in global business*. New York: McGraw Hill.
- Valli, C., Lucas, C., & Mulrooney, K. J. (2005). *Linguistics of American Sign Language: An introduction* (4th ed.). Washington, DC: Gallaudet University Press.
- Verhoeven, L. & Vermeer, A. (1992). Modeling communicative second language competence. In L. Vermeer & J. H. A. L. de Jong (Eds.), *The construct of language proficiency* (pp 163-173). Amsterdam: John Benjamins Publishing Company.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Warner, R. M. (2013). *Applied statistics: From bivariate through multivariate techniques* (2nd ed.). Thousand Oaks, CA: Sage.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Hampshire: Palgrave Macmillan.
- Wilcox, S., & Wilcox, P. (1991). *Learning to see: American Sign Language as a second*

language. New Jersey: Regents/Prentice Hall.

Wood, D. (2001). In search of fluency: What is it and how can we teach it? *The Canadian Modern Language Review*, 57, 573-589.

Zhang, B. (2010). Assessing the accuracy and consistency of language proficiency classification under competing measurement models. *Language Testing*, 27(1), 119-140.

Appendices

Appendix A: Ethics Certificate

Carleton University Research Office
Research Ethics Board
1325 Dunton Tower
1125 Colonel By Drive
Ottawa, ON K1S 5B6 Canada
Tel: 613-520-2517
ethics@carleton.ca

Ethics Clearance Form

This is to certify that the Carleton University Research Ethics Board has examined the application for ethical clearance. The REB found the research project to meet appropriate ethical standards as outlined in the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition* and, the *Carleton University Policies and Procedures for the Ethical Conduct of Research*.

X New clearance

☐ **Renewal of original clearance**

Original date of clearance:

Date of clearance	7 October 2013
Researcher	Josée-Anna Tanner, Master's student
Department	School of Linguistics and Language Studies
Supervisor	Prof. Janna Fox, School of Linguistics and Language Studies
Project number	100385
Title of project	American Sign Language testing: Examining the construct of proficiency

Clearance expires: **31 May 2014**

All researchers are governed by the following conditions:

Annual Status Report: You are required to submit an Annual Status Report to either renew clearance or close the file. Failure to submit the Annual Status Report will result in the immediate suspension of the project. Funded projects will have accounts suspended until the report is submitted and approved.

Changes to the project: Any changes to the project must be submitted to the Carleton University Research Ethics Board for approval. All changes must be approved prior to the continuance of the research.

Adverse events: Should any participant suffer adversely from their participation in the project you are required to report the matter to the Carleton University Research Ethics Board. You must submit a written record of the event and indicate what steps you have taken to resolve the situation.

Suspension or termination of clearance: Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Andy Adler, Chair
Carleton University Research Ethics Board

Louise Heslop, Vice-Chair
Carleton University Research Ethics Board

Appendix B: Teacher-Participant Interview Questions

- 1) How long have you been teaching ASL and what level(s) do/have you taught?
- 2) How important is testing and assessment as part of your curriculum?
- 3) How do you decide what to test? What does a test in the level(s) you teach look like?
- 4) What is important for proficient signing at (level taught most frequently)?
*At each interview, this was followed/clarified with “what does a high, mid, and low proficiency student look like at the level you teach most often?” and “what should learners in that level be able to do in ASL?”
- 5) How do you assess proficiency, and do all linguistic aspects of ASL need to be tested equally?
- 6) Do the tests give an accurate depiction of student proficiency?
- 7) How do you grade assessments?
- 8) How do you explain the assessment process to students?
- 9) What are the benefits of testing?
- 10) Do you perceive any issues with the current testing process, or how could it be improved?

Appendix C: Student Questionnaire

American Sign Language Testing: Exploring the Construct of Proficiency

ASLA Student Survey-All Levels

Fall 2013

1. I am: Female ☐ Male ☐
2. Age 17-18 ☐ 19-20 ☐ 21-25 ☐ 26-29 ☐ 30-35 ☐
36 or over ☐
3. I am in First year ☐ Second year ☐ Third year ☐ Fourth year ☐
Post-degree year (i.e. I already have a degree) ☐
4. My major is _____ (e.g. English, Engineering, Business, etc).
5. a) My overall average (CGPA) so far is _____ (78%, 92%, 66%, etc)
or b) I am a first year university student. My high school or college average was _____
6. Why are you taking this course? (Check all that apply)
Required for the ASL minor ☐	Personal interest ☐
It fits my timetable ☐	It was suggested by a friend ☐
It was suggested by an academic advisor ☐	I want to minor in a language ☐
I need a language credit ☐	I wanted something to keep my CGPA up ☐
Other _____ ☐	
7. How much of the course material is new to you?
All ☐	Almost all ☐	Most ☐	Half ☐
Some ☐	A small amount ☐	None ☐	
8. Have you taken a university or college language class prior to this?
No ☐	Yes, but did not complete it ☐	Yes, completed it successfully ☐
9. I am in ASL level: 1010 ☐ 1020 ☐ 1110 ☐ 2010 ☐ 2020 ☐ 3010 ☐ 3020 ☐ 4010 ☐
10. Before entering the ASL program here, did you have any prior experience with ASL or another visual language? No ☐ Yes ☐

In the next section, please respond by circling the number (from 1-6) that best answers the question. If you do not have an answer for the question, or do not understand what the question is asking, leave it blank.

General Academic Skills

11. My overall academic skills are:	Very poor					Excellent
	1	2	3	4	5	6
12. How often do you interact with your professors, outside of class hours?	Never					Very frequently
	1	2	3	4	5	6
13. How often do you contact or meet with TAs/CAs in a typical semester?	Never					Very frequently
	1	2	3	4	5	6

Test Format

14. The ASL video system for testing is easy to use	Completely disagree					Completely agree
	1	2	3	4	5	6
15. I feel just as prepared for a video test as I do for a written test	Completely disagree					Completely agree
	1	2	3	4	5	6
16. Not being able to “look over my answers” at the end of a video test is a concern to me	No concern					Large concern
	1	2	3	4	5	6
17. I know how to prepare for a video test	Not at all					Yes, definitely
	1	2	3	4	5	6
18. I feel more nervous during a video test than I do during a written one	Not at all					Yes, definitely
	1	2	3	4	5	6
19. The length of time allotted for the video tests is adequate	Not adequate					More than adequate
	1	2	3	4	5	6

	Very uncomfortable					Very comfortable
20. I am comfortable being filmed	1	2	3	4	5	6

21. In my current ASL course, I am formally assessed on (check one):	-My ability to sign ☐ -My ability to understand what someone else signs ☐ -An equal balance of both me signing and understanding what other people sign ☐
--	--

22. At the end of an ASL test, I feel confident that I have communicated clearly	Not at all					Yes, definitely
	1	2	3	4	5	6

Connection to Classroom Learning

23. ASL video tests are beneficial to my learning process	Negative benefit					Very beneficial
	1	2	3	4	5	6

24. Feedback on my tests is important to me	Unimportant					Very important
	1	2	3	4	5	6

25. How much feedback on your tests are you given?	None					A lot
	1	2	3	4	5	6

26. In this course, I feel I am doing:	Very poorly					Very well
	1	2	3	4	5	6

27. My test grades accurately reflect how I feel I am doing in this course	Not at all					Yes, definitely
	1	2	3	4	5	6

28. My test grades accurately reflect how much I feel I am learning in this course	Not at all					Yes, definitely
	1	2	3	4	5	6

29. I feel I sign better in class/during practice than I do on the tests	Not at all					Yes, definitely
	1	2	3	4	5	6

30. The tests are directly related to what I am learning in class	Not at all					Yes, very
	1	2	3	4	5	6

31. My signing would improve at the same rate, even if there were no tests	Completely disagree					Completely agree
	1	2	3	4	5	6

Assessing Linguistic Understanding

32. I can explain the function of the key phonological features of ASL (movement, location, hand configuration, etc)	Not at all					Yes, easily
	1	2	3	4	5	6

33. I can explain the basic aspects of ASL structure/syntax:	Not at all					Yes, easily
	1	2	3	4	5	6

34. I can describe the function of key ASL grammatical features as they appear in questions and sentence topicalization	Not at all					Yes, easily
	1	2	3	4	5	6

35. I know what aspects of ASL I am being tested on	No					Yes
	1	2	3	4	5	6

36. I understand what I need to do/practice in order to improve my signing proficiency	No					Yes
	1	2	3	4	5	6

37. I know what will be expected of me in the next level, if I choose to continue	No idea					Yes, definitely
	1	2	3	4	5	6

Assessing Understanding of the Grading Process

38. How have you done on your ASL tests this semester?	Much lower than expected					Much higher than expected
	1	2	3	4	5	6

39. Were you surprised by your grades?	Not at all					Very surprised
	1	2	3	4	5	6

40. How did you feel about your test results?	Very bad					Excellent
	1	2	3	4	5	6
41. The instructor is clear in explaining how tests will be graded	Not clear					Very clear
	1	2	3	4	5	6
42. I understand the grading process for ASL tests	Not at all					Yes, completely
	1	2	3	4	5	6
43. When I see my grade for an ASL test, I understand why I obtained that mark	No					Yes
	1	2	3	4	5	6

Self-Evaluation of ASL Proficiency

At my current ASL level, I can:

1. Recognize familiar and very basic signs, such as those concerning myself, my family, and my immediate surroundings.	Not at all					Yes, easily
	1	2	3	4	5	6
2. Understand simple statements or questions in ASL (“Hello”, “How are you?”, “What’s your name?”, “It’s cold today”, etc).	Not at all					Yes, easily
	1	2	3	4	5	6
3. Create a short monologue describing myself in ASL.	Not at all					Yes, easily
	1	2	3	4	5	6
4. Make myself understood using short, simple signed sentences and gestures when I am describing things I am very familiar with.	Not at all					Yes, easily
	1	2	3	4	5	6
5. Ask and answer questions in ASL about personal details in a simple way.	Not at all					Yes, easily
	1	2	3	4	5	6
6. Understand high-frequency signed phrases that are used often in the classroom.	Not at all					Yes, easily
	1	2	3	4	5	6

7. Use a series of signed sentences to describe everyday situations/places (i.e. daily routines, my home, my educational background, etc).	Not at all					Yes, easily
	1	2	3	4	5	6
8. Understand a native signer who is signing to me slowly and carefully (i.e. someone who is deliberately adapting their signing to suit you).	Not at all					Yes, easily
	1	2	3	4	5	6
9. Understand most short, simple, personal ASL narratives about basic subjects (shopping, travel, work, etc).	Not at all					Yes, easily
	1	2	3	4	5	6
10. Buy clothes in a store where all the staff only sign.	Not at all					Yes, easily
	1	2	3	4	5	6
11. Describe my favourite hobby in ASL in some length, using appropriate vocabulary.	Not at all					Yes, easily
	1	2	3	4	5	6
12. Understand fingerspelling when it is slow and deliberate, and when the word or name being spelled is familiar to me.	Not at all					Yes, easily
	1	2	3	4	5	6
13. Give a brief but accurate signed description of a photograph or painting while looking at it.	Not at all					Yes, easily
	1	2	3	4	5	6
14. Repeat back part of what someone has signed to confirm mutual understanding.	Not at all					Yes, easily
	1	2	3	4	5	6
15. Give directions in sign to someone on the street with only slight hesitations and rephrasing.	Not at all					Yes, easily
	1	2	3	4	5	6
16. Understand the main points of a news broadcast in ASL.	Not at all					Yes, easily
	1	2	3	4	5	6

18. Appropriately interject in a signed conversation in order to give my opinion on a topic that I am somewhat familiar with.	Not at all					Yes, easily
	1	2	3	4	5	6
19. Follow complex and extended signing, as long as the topic is familiar to me (for example, a signed lecture or speech on a topic that I am familiar with).	Not at all					Yes, easily
	1	2	3	4	5	6
20. Understand fingerspelling at full speed, as long as I know what the context is.	Not at all					Yes, easily
	1	2	3	4	5	6
21. Develop well-structured points and conclusions as part of an argument/discussion in ASL, without having to obviously search for expressions.	Not at all					Yes, easily
	1	2	3	4	5	6
22. Express myself in ASL on a wide range of topics fluently and spontaneously, using clear descriptions and themes.	Not at all					Yes, easily
	1	2	3	4	5	6
23. Understand a native signer who is signing to me quickly, using as many colloquialisms as they would when communicating with another native signer.	Not at all					Yes, easily
	1	2	3	4	5	6
24. Create properly structured and coherent discourse in ASL on almost any topic, using a wide range of connectors and varying organizational patterns.	Not at all					Yes, easily
	1	2	3	4	5	6
25. Maintain consistent grammatical control of complex signing, even when my attention is otherwise engaged (thinking ahead, making note of others' reactions, etc).	Not at all					Yes, easily
	1	2	3	4	5	6
26. Take part in a formal job interview in ASL where I am required to give accurate information about my interests, qualifications, and experience.	Not at all					Yes, easily
	1	2	3	4	5	6

Thank you for your time!