

## 2. HISTORY OF THE ESI AND PRELIMINARY RESEARCH UNDERLYING THE DEVELOPMENT OF THE ESI

### 2.1 History of the Development of the ESI

The development of the ESI can be considered to have begun with the work of Doble and Magill-Evans (1992) when they developed an occupational therapy model of social interaction. In part, the development of that model was based on the Model of Human Occupation (as it was described at that time) (Kielhofner, 1985) where motor, process, and communication and interaction skills were viewed as comprising the performance subsystem, the foundation for occupational performance. Later, in the 1995 version of the Model of Human Occupation, motor, process, and communication and interaction as well as social interaction skills were operationally defined and described as *taxonomies of performance skills* that could be used to help the occupational therapist focus on describing, using a “standard language,” what actually was observed during a client’s occupational performance (Fisher & Kielhofner, 1995). At that time, performance skills were defined as observable elements of action and were clearly differentiated from underlying capacities or body functions.

**Performance skills (motor, process, and social interaction) must not be confused or equated with underlying physical, cognitive, or emotional capacities. Nor does the “concept of skill . . . directly imply a given amount of underlying capacity. It is purposefully disinterested in underlying capacity which to a large extent cannot be observed. Instead, attention is focused on the actual doing or performance” (Fisher & Kielhofner, p. 115).**

In Doble and Magill-Evans’ (1992) model, social enactment skills were defined as observable “behaviors and actions that enable us to communicate our needs and

intentions to others and to respond to the messages of others in a competent manner” (p. 146). In conjunction with the development of their model, they suggested a preliminary taxonomy of social interaction skills, which they referred to as social enactment skills (see Table A–1, Appendix A). Different from the taxonomy of communication and interaction skills developed by Simon (1989) and Salamy (1993), an unique feature of Doble and Magill-Evans’ taxonomy was that they broadened the concept of social interaction to include skills used to competently respond to messages from others.

During the next phase, beginning in 1992, Englund (1997; Englund, Bernspång, & Fisher, 1995) developed the Assessment of Social Interaction (ASI). To create the ASI, Englund started with the taxonomy of social enactment skills proposed by Doble and Magill-Evans (1992), and then, based upon a thorough review of the literature, including the first version of the Assessment of Communication and Interaction Skills (Simon, 1989), added a few additional skills (see Fisher & Kielhofner, 1995 for a description of the skills included in Simon’s taxonomy of communication and interaction skills, and Englund’s taxonomy of social interaction skills). Since Doble and Magill-Evans had not developed operational definitions for their suggested taxonomy of social enactment skills, Englund also created the first operational definitions for each social interaction skill included in the ASI. In that process, she adapted the methodology that formed the basis of the Assessment of Motor and Process Skills (Fisher, 2006a, 2006b) (and later, the School Version of the Assessment of Motor and Process Skills) (Fisher et al., 2005) to accommodate for the assessment of social interaction skills. More specifically, Englund adopted from the Assessment of Motor and Process Skills the 4-category rating scale and the basic format for defining each skill with a set of scoring examples.

The ASI was comprised of 25 social interaction items designed to evaluate observable interpersonal social actions (social interaction items) that one uses when he/she listens to others, asks for information, carries out conversations, and adapts his/her behavior based on the social situation and his/her social partners (see Table A–1, Appendix A). Englund’s (1997) goal was to develop a new type of occupational therapy assessment of social interaction skills designed to fill a gap in needed standardized assessments of a person’s quality of social interaction. More specifically, Englund felt that there was a need for a tool that enabled us to:

- Evaluate social interaction skill *without reliance on role play or other forms of contrived assessment*
- Evaluate *based on naturalistic and direct observation*, not using interview, self- or proxy-report, or videotaping
- Evaluate *using standardized*, rather than informal or nonstandardized assessment and *scoring procedures*
- Evaluate social interaction skills associated with *both the sending and the reception of messages*, not just the social interaction skills associated with sending skills

In 2002, Fisher (2002/2006, 2006c, 2009) began a process of refining and further clarifying the ASI items for inclusion in the Occupational Therapy Intervention Process Model (OTIPM) (Fisher, 1998, 2002/2006, 2009) (see also Section 4.1 and Table 4–1, Chapter 4 for further information about the OTIPM). As part of that process, Fisher implemented a careful review of the actions listed in (a) the Communication and (b) the Interpersonal Interactions and Relationships domains of the *International Classification of Functioning, Disability and Health (ICF)* (World Health Organization [WHO], 2001) to identify potential gaps in the content of the ASI. This resulted in several new items being added, and in others being revised (see Table A–1, Appendix A). The revised ASI was evaluated in a series of pilot studies, and the results provided the basis for the new *Evaluation of Social Interaction*, ESI.

## **2.2 Preliminary Evidence of Validity and Reliability Based on the Assessment of Social Interaction (ASI) and Early Research Versions of the ESI**

The pilot studies that involved the use of the ASI as well as earlier versions of the ESI provided good preliminary evidence for both validity and reliability of the emerging ESI. Those results are summarized below.

### 2.2.1 Evidence based on the ASI

Englund, Bernspång, and Fisher (1995) implemented a pilot study of the original ASI. Their participants included 16 adults, 24 to 75 years of age ( $M = 36.1$ ,  $SD = 16.9$ ). Six participants were described as having good health, 5 had psychiatric disorders, and 5 were healthy adults who role-played persons with psychiatric disorders.

Englund et al. (1995) used Rasch analysis methods to analyze their data (Bond & Fox, 2007; Linacre, 1993). When one uses Rasch analyses, several aspects of validity and reliability can be evaluated. **Internal scale validity** refers to the idea that a set of items included in a test must be unidimensional (i.e., test only one construct). One method of evaluating internal scale validity is to ensure that at least 95% of the items demonstrate statistical goodness of fit to the specified Rasch model. Goodness of fit is evaluated by use of mean square (*MnSq*) and standardized (*z*) goodness-of-fit statistics (Bond & Fox). Englund et al. found that 21 of 25 items (84%) demonstrated acceptable goodness of fit to the many-faceted Rasch (MFR) model of the ASI ( $MnSq \leq 1.4$ ,  $z < 2$ ). The four items that did not demonstrate acceptable fit were *Touches*, *Places Self*, *Completes*, and *Thanks* (see Table A–1, Appendix A). Based on their analysis of the reasons for the misfit, Englund et al. recommended revision of these four items.

When they evaluated **person response validity**, Englund et al. (1995) used person-related Rasch goodness-of-fit statistics in a similar manner to determine if the 16 participants had patterns of scores that matched the expectations of the MFR model of the ASI. In this case, 14 of the 16 participants (87.5%) demonstrated acceptable goodness of fit. To further evaluate person response validity, they examined the hierarchical ordering of the participant social interaction quality of performance measures, and found that the participants were ordered according to their expected rankings. Not surprisingly, the two participants who failed to demonstrate acceptable fit were ones who had role-played persons with psychiatric disorders.

### 2.2.2 Evidence based on the early research versions of the ESI

In 2006, after Fisher's (2002/2006) initial revision of the ASI items to create the ESI items (i.e., refining the original items and adding content from ICF), Asplund and Forsberg (2006) evaluated 30 persons from northern Sweden who were judged to have overall good social competence. These persons were between 20 to 95 years of age ( $M = 44.6$ ,  $SD = 19.9$ ), and 19 were women and 11 were men. When their data were

subjected to MFR analysis, their results revealed that both raters demonstrated **rater reliability** ( $MnSq \leq 1.4, z < 2$ ), and that there was no clinically meaningful difference in their rater severities (differences = 0.2 logit). More specifically, Asplund and Forsberg used goodness-of-fit statistics for raters who co-scored the same persons to evaluate if the two raters scored the participants in a manner expected by the MFR model of the ESI.

**The term *logits* refers to log-odds probability units (Bond & Fox, 2007), and in the ESI, they represent linearized increments of quality of social interaction along the social interaction measurement scale.**

Typically, when **internal scale validity** is evaluated, the evaluation is based, in part, on goodness-of-fit statistics for the items included in a test. When Asplund and Forsberg (2006) evaluated internal scale validity, however, they recognized that the ESI scale is constructed based on items that vary in difficulty as well as different intended purposes of social interaction that, like items, also vary in their challenges. Therefore, Asplund and Forsberg also considered the importance of ensuring that both the ESI items and the intended purposes of social interactions work together to define the same underlying construct.

More specifically, when Asplund and Forsberg (2006) evaluated internal scale validity of the ESI, they found that all but one item (*Places Self*) demonstrated acceptable goodness of fit to the MFR model of the ESI ( $MnSq \leq 1.4, z < 2$ ). When four different intended purposes of social interaction were evaluated, all four demonstrated acceptable goodness of fit to the MFR model of the ESI. Listed in order of challenge of the intended purpose, from easiest to hardest, the four intended purposes were *Information/Service*, *Conversing socially/Small talk*, *Problem solving/Decision making*, and *Collaborating/Producing*. Familiarity of the social partner (known vs. unknown) also demonstrated acceptable fit, and social interaction with an unknown partner was slightly easier (0.32 logit) than was social interaction with a familiar social partner.

Finally, when **person response validity** was examined, Asplund and Forsberg (2006) found that 29 of the 30 persons demonstrated acceptable goodness of fit. Further analyses revealed that the ESI measures could be used to divide the sample of socially competent adults into at least four groups that differed significantly in mean quality of social interaction ( $F[3,26] = 43.09, p < .001$ ). Considered together, these results supported (a) the potential for development of a tool that could be scored in a reliable manner by trained raters, (b) the internal validity of the scale and the validity of the person quality of performance ESI measures, and (c) the sensitivity of the ESI measures.

In a subsequent study, Simmons, Griswold, and Berg (2010) analyzed the data for 128 persons, 4 to 73 years of age. Their participants included well adults and children, adults with neurological disorders (e.g., stroke, traumatic brain injury) or psychiatric disorders (e.g., depression, anxiety), and children with developmental disorders (e.g., autism, mental retardation). All but one person demonstrated acceptable infit goodness of fit, and all but five persons (96%) demonstrated acceptable outfit goodness of fit ( $MnSq \leq 1.4, z < 2$ ), supporting high **person response validity**. As expected, the well adults had significantly higher ESI quality of social interaction measures than did adults with disabilities ( $t[83] = 4.47, p < .001$ ).

The aspects of **reliability** of Rasch-generated measures for a test can be evaluated using separation indices and the standard error of measurement (*SE*). The separation index for persons indicates how reliably the participants in a sample would be hierarchically ordered if they were given a parallel form of the test (i.e., if they would have the same hierarchical order under both test forms) (Bond & Fox, 2007). The separation index for persons was 2.87, and the associated separation reliability coefficient was .89 (Simmons et al., 2010). The average person *SE* was 0.23 logit, supporting both reliable measurement and the potential for ESI measures to detect change over time (Simmons & Griswold, 2008).

When Simmons et al. (2010) evaluated **internal scale validity**, they examined data for six different purposes of social interaction and 27 items. All six intended purposes demonstrated goodness of fit to the MFR model of the ESI, and the hierarchy of intended purpose challenges was similar to the earlier findings of Asplund and Forsberg (2006). Again listed from easiest to hardest, they were: *Financial exchange*, *Conversing socially/Small talk*, *Problem solving/Decision making*, *Information sharing*, *Information seeking*, and *Collaborating/Producing*. The intended purpose

separation index was 3.55 (separation reliability,  $R = .93$ ) (Simmons & Griswold, 2008). When the items were examined, 24 of the 27 items (88.9%) demonstrated acceptable goodness of fit. Those not demonstrating acceptable fit were *Discloses*, *Concludes/Disengages*, and *Thanks* (Simmons et al.). These items have since been revised, and the result was the current version of the ESI.

In yet another study, Simmons and Griswold (2010) evaluated the *sensitivity* of the ESI quality of social interaction measures for detecting change a following community-based intervention program. The participants were 10 persons, 28 to 61 years of age, who had sustained a traumatic brain injury 7 to 19 years before implementation of the study. Following 8 weeks of compensatory and acquisitional occupation, implemented following the intervention process described in the OTIPM (Fisher, 1998, 2002/2006, 2009) (see Section 4.1 and Table 4–1, Chapter 4 for more information about the OTIPM), the participants demonstrated significant differences in mean quality of social interaction ( $t = 4.98, p = .001$ ). Six of the 10 participants had improvements that were  $\geq 0.34$  logit, a value equivalent to 2 *SE* based on the current ESI standardization sample (see Section 3.6.4, Chapter 3).

**The results of early pilot studies supported good potential for the continued development of a performance-based, naturalistic evaluation of social interaction that:**

- **Could be scored reliably by trained raters**
- **Demonstrated internal scale validity and validity of the person quality of social interaction measures**
- **Was sensitive enough to be an effective evaluation of change**

While there was no control group, it is important to stress that given the length of time post-injury, it was unlikely that the results were due to spontaneous recovery (cf. Wæhrens & Fisher, 2007). Moreover, the person who administered the ESI was blind to the fact that these 10 participants had received intervention.