

Occupational therapy students rating with the N-ACIS: Do they agree?

Av Tore Bonsaksen, Maria Fouad, Cecilia Celso, & Hiromi Nakamura-Thomas

Tore Bonsaksen er ergoterapeut og førstelektor ved Høgskolen i Oslo og Akershus, Fakultet for helsefag, Institutt for ergoterapi og ortopediingeniørfag. E-post: tore.bonsaksen@hioa.no

Maria Fouad er forskningsassistent ved Høgskolen i Oslo og Akershus, Fakultet for helsefag, Institutt for ergoterapi og ortopediingeniørfag.

Cecilia Celso er ergoterapeut ved Oslo Universitetssykehus HF.

Hiromi Nakamura-Thomas is associate professor at Graduate School of Health, Medicine and Welfare, Saitama Prefectural University, Saitama, Japan.

Denne artikkelen ble mottatt 21.12.2015. Den ble godtatt for publisering 12.04.2016. Conflicts of interest: None

Occupational therapy students rating with the N-ACIS: Do they agree?

Abstract

Background: The Assessment of Communication and Interaction Skills (ACIS) has been shown to possess good psychometric properties, but the recently translated Norwegian version of the instrument (N-ACIS) has not yet been subjected to such investigation. Moreover, the amount of training needed in order to ensure that raters' ACIS scores are reliable has not yet been explored.

Methods: Twenty-six occupational therapy students (response rate 60.5 percent) participated in this study, which was performed in conjunction with a student training seminar. Each student performed two assessments with the N-ACIS, based on observations of students who had been given specific instructions for roleplaying during the performed activities. Pairs of students were assessed for interrater agreement on the N-ACIS total scale and subdomain scales with the intraclass correlation coefficient (ICC).

Results: The ICC measures were 0.92 (total scale score), 0.99 (physicality), 0.76 (information exchange), and 0.80 (relations).

Conclusion: A very brief introduction to the N-ACIS appears to be sufficient to achieve high interrater reliability on the aggregated scales – at least under artificial conditions. Remaining questions concern the level of interrater agreement at the item level, how item scores would correspond with an expert opinion, and interrater agreement in real life practice situations.

Keywords: Assessment of Communication and Interaction Skills, interrater reliability, intraclass correlation coefficient, Model of Human Occupation, observational assessment

INTRODUCTION

Occupational therapists need to have sufficient knowledge and skills in some areas, whereas they need to possess an expert level in others. One of the areas where an expert level of knowledge and skills is needed is occupational performance analysis (Thomas, 2012). When analyzing a person's performance, the occupational therapist may closely examine how the person executes the skills involved in the activity, i.e., the single action components needed to perform the activity successfully (Crepeau, 2003). Persons with skill limitations may suffer from diminished capacity to perform important and valued occupations, and may resultantly have reduced opportunity to participate in society the way they would like to.

Given the importance of skills for occupational performance, and ultimately for participation in society, occupational therapists have developed a range of tools for assessing skills during occupational performance. One tool developed with a theoretical basis in the Model of Human Occupation (MOHO; Kielhofner, 2008) is the Assessment of Communication and Interaction Skills (ACIS; Forsyth, Salamy, Simon, & Kielhofner, 1998). The ACIS consists of 20 discrete skills. It can be used as a measure of the overall communication and interaction level of an individual, or it can be used to identify specific skills or broader skill areas (domains) that may need to be addressed as part of an intervention. The skills are organized as belonging to three different domains: physicality, information exchange, and relations (Forsyth et al., 1998). An overview of the

PHYSICALITY	INFORMATION EXCHANGE	RELATIONS
contacts	articulates	collaborates
gazes	asserts	conforms
gestures	asks	focuses
maneuvers	engages	relates
orients	expresses	respects
postures	modulates	
	shares	
	speaks	
	sustains	
All items are scored 1-4, where 1 = deficit, 2 = ineffective, 3 = questionable, 4 = competent.		

Table 1: Skills assessed with the Assessment of Communication and Interaction Skills (ACIS) and their relationship to domains of interaction.

ACIS skills and domains is provided in Table 1.

As human occupations often are group based and performed in a variety of social settings and environments, communication and interaction skills are needed for establishing and maintaining relationships and for coordinating the actions of the persons involved to ensure that mutual goals are met (Forsyth et al., 1998; Kielhofner, 2008). Given the importance of these skills for a person's occupational performance and participation, occupational therapy students have been encouraged to learn about related assessment and intervention strategies, and using the ACIS has served as a starting point for such learning (Bonsaksen, Granå, Celso, Ellingham, & Myraunet, 2013; Bonsaksen, Myraunet, Celso, Granå, & Ellingham, 2011).

Decreased communication and interaction skills is by no means limited to specific diagnostic groups. By tradition, however, the

ACIS appears to have been used mainly in relation to persons with psychiatric disorders and intellectual/learning disabilities (Fuller, 2011; Simmons, Griswold, & Berg, 2010). In the first psychometric study introducing the ACIS to the worldwide community of occupational therapists (Forsyth, Lai, & Kielhofner, 1999), clients with different psychiatric disorders were compared.

The results showed that the three groups of clients had different levels of skills that mirrored the severity of their mental health problems. In addition, the analysis revealed the ACIS scale to be unidimensional, it measured the clients' skills in an appropriate way, it was able to separate study participants into six different levels of communication and interaction skills, and there was good level of consistency within and between raters (Forsyth et al., 1999). Later studies and reviews have continued to advocate for the validity of the ACIS,

DURATION	CONTENT
45 min	<i>Introduction</i> Introducing the ACIS. Its foundation in the MOHO, the content of each skill, assessment procedure, and scoring instructions
60 min	<i>Observation 1</i> Activity 1 (30 min), individual scoring (10 min) and group discussions concerning the scores and the observations they build on (20 min).
60 min	<i>Observation 2</i> Activity 2 (30 min), individual scoring (10 min) and group discussions concerning the scores and the observations they build on (20 min).
30 min	<i>Summary</i> Student experiences concerning the use, scoring, and reasoning with the ACIS are discussed in the whole class. Completed ACIS sheets are collected.

Table 2: Content and organization of the ACIS seminar.

most frequently based on data from persons with mental health problems (Fuller, 2011; Hsu, Pan, & Chen, 2008; Kjellberg, Haglund, Forsyth, & Kielhofner, 2003), and translated versions of the ACIS have been found to be useful for clinical practice as experienced by Scandinavian occupational therapists and occupational therapy students (Bonsaksen, Granå, et al., 2013; Bonsaksen et al., 2011; Kjellberg & Haglund, 2015; Nielsen & Andersen, 2006; Petersen & Hartvig, 2008). Research has continued to suggest, however, that measures of intrarater and interrater reliability of the ACIS should be obtained (Hsu et al., 2008).

Evidence that a measure has acceptable levels of validity and reliability is vital for the credibility of an instrument in the research community, and consequently for its use in occupational therapy practice (Laver-Fawcett, 2014). Such psychometric properties need to be re-established for translated versions of instruments previously validated in other languages (Streiner & Norman,

2008). So far, the psychometric properties of the Norwegian version (N-ACIS) have not been formally examined (Ellingham, Hussain, & Bonsaksen, 2014). Moreover, we do not know how much training is needed to obtain reliable N-ACIS scores: the original ACIS study used two days of training (Forsyth et al., 1999), whereas later studies have used one day (Kjellberg & Haglund, 2015) or as little as three hours of training (Bonsaksen et al., 2011; Haglund & Thorell, 2004). In the present study, we connected the shortcomings in the existing literature by performing a preliminary investigation of the interrater reliability of N-ACIS scores between pairs of occupational therapy students.

AIM OF THE STUDY

The aim of this study was to assess the level of correspondence between occupational therapy students' N-ACIS ratings. We also explored the correspondence between ratings related to the three subdomains: physicality, information exchange, and relations.

Methods

EDUCATION CONTEXT AND TRAINING

The study was conducted at the occupational therapy education program at Oslo and Akershus University College in Oslo, Norway. Approximately 250 students are enrolled in the program, and approximately 70 students graduate on an annual basis (Bonsaksen, Kvarsnes, & Dahl, 2015). The education program is an undergraduate program with a duration of three years encompassing 12 study modules (Oslo and Akershus University College, 2011).

The second year of the education program starts with a 10 week study module named «Mental health and participation». During this module, the students are introduced to the N-ACIS in a half-day seminar. The organization and content of the seminar is outlined in Table 2. The seminar instructors (Authors 1 and 3) are local experts in using the assessment, having used it in clinical practice and in research over the last six years (Bonsaksen, Celo, Myraunet, Granå, & Ellingham, 2013; Bonsaksen, Granå, et al., 2013; Bonsaksen et al., 2011). The first author has also been involved in translating the ACIS user's manual into Norwegian (N-ACIS) (Ellingham et al., 2014).

SAMPLE AND PROCEDURE

A total of 58 students entered the second year of the education program in the autumn of 2015, and 55 of these participated in the ACIS seminar. They had no previous knowledge of, or experience with, the ACIS, but were somewhat familiar with the MOHO (Kielhofner, 2008). In the seminar,

the students were organized in 12 groups. In each of these groups, one student had been given the task of role-playing a person who had specified problems related to communication and interaction skills (the «student-client»). In some of the groups, the student-client was encouraged to behave with decreased skills related to the physicality domain. The student-client could then, for example, play out behaviors like inappropriate hugging (contacts), having prolonged eye-contact (gazes), and using excessive hand movements during interaction (gestures). In other groups, the student-client was encouraged to play out decreased skills related to the information exchange domain. This could, for example, include behaviors like not sharing relevant facts with the group (shares), monopolizing the conversation (sustains), and expressing rapidly shifting emotions with the group members (expresses). In yet other groups, the student-client was encouraged to play out decreased skills related to the relations domain, for example by using offensive language (conforms), being easily distracted (focuses), and interacting in non-relating ways (relates).

Two group activities were performed during the seminar, followed by individually performed N-ACIS ratings and a subsequent discussion concerning the ratings and the observations they were based on (see Table 2). The students were given examples of activities that they could choose from, but were also given the option of independently selecting activities during which to observe the student-client. The chosen group activities included: building with Legos, playing Twister

(a highly physical game where the participants assume various body postures according to given procedures), playing board games (such as Scrabble), visits to a café, group compositions (like drawing), and doing a quiz.

As 12 participants in the seminar role-played a person with decreased communication and interaction skills, there were 43 eligible students for inclusion as participants. At the end of the seminar, 26 students returned their completed N-ACIS sheets, yielding a response rate of 60.5 percent. Given that the study only collected anonymous data, no specific information about the participants' background characteristics is available. Previous research, however, has described the study cohort (from which this sample was drawn) as relatively young (mean age 23 years) and predominantly female (81 percent) (Bonsaksen et al., 2015).

MEASURES

The ACIS (Forsyth et al., 1998) is a 20 item assessment of communication and interaction skills. It is theoretically based on the MOHO (Kielhofner, 2008), defining skills as «...observable, goal-directed actions that a person uses while performing [an activity]» (p. 103). It consists of 20 discrete skills that are part of one of three subdomains: physicality, information exchange, and relations. Based on observation from activity situations requiring some social interaction, each skill is rated on a 1 - 4 scale (1 = deficit, 2 = ineffective, 3 = questionable, 4 = competent). Generally, each rating reflects the extent to which the observed skill supports or hinders the flow and completion of the activity as well as the ongoing social interacti-

on. In result, the sumscore of the total scale ranges from 20 (lowest skill level) to 80 (highest skill level). In this study, to adjust for unequal number of items on the three subdomains, all scale scores (total scale and subdomains) were divided with the number of items belonging to the scale, resulting in scale scores ranging from 1 (lowest) to 4 (highest). Following this procedure, direct comparisons across the different scales are made meaningful – for example, lower scores in the relations domain compared to the scores in the physicality and information exchange domains would indicate a lower level of skill in the relations domain. The Norwegian translation of the ACIS (N-ACIS) was used in the study (Ellingham et al., 2014). Table 1 shows the skills and domains of the ACIS, which are the same as in the employed N-ACIS.

DATA PREPARATION AND ANALYSIS

When performing assessments with the ACIS, it is not uncommon to find that the activity chosen for the assessment did not provide information about all of the 20 skills (Bonsaksen et al., 2011). In such cases, the item is not rated (Ellingham et al., 2014; Forsyth et al., 1998). In the dataset for this study, a total of 46 missing data points (representing 4.4 percent of the maximum data that could be obtained) were detected. In preparing the dataset for analysis, it was decided, in line with previous research using other scales (e.g., Bonsaksen, Lerdal, & Fagermoen, 2012), that the N-ACIS scale scores (total scale and the three subdomain scale scores) could be constructed while tolerating a certain level

	No assessment forms returned	1 assessment form returned by the group	2 assessment forms returned by the group	3 assessment forms returned by the group	4 assessment forms returned by the group
Number of groups who returned this number of assessment forms	2	2	2	4	2
Total number of participants	0	2	4	12	8
Participants removed prior to analysis (to assess rating consistency between sets of two raters only)	0	2 (no other assessment for measuring agreement)	0 (paired data achieved for all data)	4 (3rd participant data removed – selected at random leaving paired data for analysis)	0 (paired data achieved for all data)
Participants included in the analysis	0	0	4 (2 sets of paired data)	8 (4 sets of paired data)	8 (4 sets of paired data)
On each ACIS form, two assessments (completed at two separate occasions) were completed. 20 participants (10 pairs), each participant providing two ACIS assessments, resulted in a total of 40 ACIS assessments to be included in the analysis.					

Table 3. Overview of the obtained data from the participants, accounting for missing data.

of missing responses on single skill items. In this study, it was pragmatically decided that up to 25 percent missing responses on items belonging to each scale was acceptable. In effect, the N-ACIS total scale was based on the mean of a minimum of 15 items, whereas the physicality, information exchange, and relations subdomain scores were based on the mean of a minimum of 5, 7, and 4 items respectively.

Each of the twenty-six participating students provided scores from two subsequent observations of the student-client, according to protocol. The 12 groups had an unequal number of participants. In two groups, the students did not return any completed assessments. In two other groups, one participant per group returned the completed N-ACIS sheets. From another two groups, two completed assessments were returned from each group. Three

completed assessments were returned from four groups, whereas the remaining two groups returned completed assessments from four participants.

Participants belonging to a group where only one assessment was returned for analysis were excluded (as they had no other assessment against which to measure agreement). To be able to use as much of the data as possible, we decided to assess the level of rating consistency between sets of two raters only. Thus, we also excluded the data provided by the third student rater (randomly selected) in the four groups where three participants had returned their assessments. In the groups consisting of four participants, we re-arranged the dataset so that two and two student raters were assessed for rating consistency. As a result, the dataset subjected to the interrater agreement analysis consisted of

a total of 40 ACIS assessments coming from 20 raters (10 pairs of raters) who had each performed the assessment at two occasions. Table 3 shows an overview of the obtained data and explains how parts of the data material was excluded from the analysis.

In order to estimate the level of agreement between the student raters, the intraclass correlation coefficients (ICC) were produced (Shrout & Fleiss, 1979; Streiner & Norman, 2008). This method of estimating interrater agreement is not sample dependent and is able to account for several different sources of error simultaneously (Kielhofner, 2006). Judging by the number of journal articles reporting ICCs as a measure of interrater reliability, this method has become more common in the field of occupational therapy over the last years (e.g., Donohue, 2007; Stigen & Page, 2012). As it was the reliability of the specific

Pair of raters	Observation #	PHYSICALITY		NFORMATION EX.		RELATIONS		ACIS TOTAL	
		Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2	Rater 1	Rater 2
1	1	3.00	3.20	2.50	2.78	1.00	1.00	2.22	2.42
	2	2.33	2.08	2.22	2.33	1.00	1.20	2.00	1.98
2	1	3.33	3.33	3.11	3.11	2.40	2.40	3.00	3.00
	2	1.83	2.00	3.11	3.22	2.20	2.20	2.50	2.60
3	1	3.33	2.92	3.57	3.88	2.75	3.25	3.29	3.42
	2	*	*	2.89	3.75	*	*	*	*
4	1	4.00	4.00	3.33	3.00	*	*	3.40	3.35
	2	4.00	4.00	3.00	3.11	2.80	3.20	3.26	3.40
5	1	2.50	2.17	2.00	2.38	2.20	2.00	2.20	2.21
	2	2.83	2.83	3.00	2.33	2.00	2.60	2.70	2.55
6	1	4.00	4.00	3.33	3.11	3.40	3.00	3.55	3.35
	2	4.00	4.00	3.11	2.89	3.60	2.60	3.50	3.15
7	1	1.67	2.50	1.56	2.67	1.40	2.40	1.55	2.55
	2	4.00	4.00	3.89	4.00	3.80	3.60	3.90	3.90
8	1	1.33	1.17	2.11	2.56	1.60	1.60	1.75	1.90
	2	1.67	1.67	2.22	2.33	1.80	2.00	1.95	2.05
9	1	4.00	4.00	3.67	2.22	3.60	1.60	3.75	2.60
	2	4.00	4.00	3.11	2.56	3.40	2.20	3.45	2.90
10	1	1.33	1.50	2.22	1.67	1.80	1.80	1.85	1.65
	2	1.17	1.33	2.00	2.11	1.25	1.00	1.58	1.60
		ICC = 0.99		ICC = 0.76		ICC = 0.80		ICC = 0.92	

Table 4. Scores on skills, subdomains, and N-ACIS total scale provided by pairs of student raters. Note: * indicates missing data.

raters that was being studied, a mixed-effect model was employed, treating N-ACIS scale scores as fixed factors and raters as random factors. Also, given the interest in the raters' consistency in agreement, not their absolute agreement; the consistency type was used. The ICC is interpreted similar to well-known measures of reliability, like Cronbach's alpha. For scale consistency, a ≥ 0.70 is generally considered acceptable, a ≥ 0.80 is good, whereas a ≥ 0.90 is considered excellent (Field, 2005). For satisfactory intrarater/interrater reliability, a ≥ 0.75 is generally required. A confidence interval (95 percent CI)

was constructed around the ICC average measure, and the level of statistical significance was set at $p < 0.05$.

ETHICS

All of the students were informed about the study by the researchers (Authors 1 and 3) and volunteered to participate. As these researchers also had the role of seminar instructors, it was emphasized that study participation was voluntary and there would be no negative consequences (e.g., related to the teaching and learning experience, or to assignment marks) for persons who opted not to participate. Conversely,

participation in the study had no benefit for those who chose to take part. All data was collected anonymously, therefore approval from the Norwegian Data Protection Official for Research was not required.

Results

Table 4 shows the scale scores of all participants after 1) adjusting for missing scores on single skill items and 2) having re-arranged the dataset to allow for pairwise comparisons between two and two student raters who had observed the same student-client at the same two occasions (see data analysis section for details). The

number of pairwise comparisons varied across domains as a result of missing data on some of the scales.

There was excellent consistency in agreement related to the ACIS total score between pairs of student raters (ICC = 0.92, 95 percent CI [0.79, 0.97], $p < 0.001$). There was virtually perfect inter-rater agreement related to the physicality subdomain scores (ICC = 0.99, 95 percent CI [0.96, 0.99], $p < 0.001$). In the information exchange subdomain, the scores of pairs of raters showed a lower level of consistency (and with a wide confidence interval), yet within the limits of what is generally considered acceptable (ICC = 0.76, 95 percent CI [0.38, 0.90], $p < 0.01$). There was good level of consistency (although with a wide confidence interval) between pairs of raters when concerned with the relations subdomain scores (ICC = 0.80, 95 percent CI [0.47, 0.93], $p < 0.001$).

Discussion

This study showed that occupational therapy students, after only minimal training, managed to score the N-ACIS in a way that resulted in a very high level of interrater agreement on the scales between pairs of student raters. The reliability coefficients indicated almost perfect agreement between the student raters on the physicality scale, good to excellent agreement on the relations and total score scales, and acceptable agreement on the information exchange scale.

Hsu and coworkers (2008) limited their research on the Chinese version of the ACIS to ascertain the validity of the scale, and encouraged future studies to

«...examine interrater and intrarater reliability as well as concurrent and predictive validity» (p. 184). The original psychometric study of the ACIS (Forsyth et al., 1999) had a more extensive scope, and it indicated a good level of reliability among the raters using the ACIS. Similar positive results concerning rater reliability were set forth by Kjellberg and coworkers (2003), who reported from a study in which the Swedish translation of the ACIS has been used. However, the Rasch methodology employed in the Forsyth (1999) and Kjellberg (2003) studies is different from the ICC measures produced with the classical test-theory approach in the present study. For example, Forsyth and coworkers (1999) found that the *pattern of scores* was consistent with the applied Rasch model for all except three of the 52 raters, whereas the present study documented the level of *score agreement between pairs of raters*. The few studies in the area, and the different analytical approaches used in them, indicate that comparisons should be made with caution.

Given the very brief training that was provided to the students, our study may indicate that the ACIS has a very intuitive appeal for users, and that little training is required for using it with reliable results. This mirrors the relatively scarce amount of training provided to raters in previous studies (Bonsaksen et al., 2011; Haglund & Thorell, 2004; Kjellberg et al., 2003). However, it should be noted that this study only examined the level of correspondence between student raters, and not the extent to which these ratings were correct or justified (i.e. in correspondence

with an expert opinion). Further studies may ensure that interrater reliability is assessed with reference also to an expert opinion, in order to minimize the error potential associated with lack of adequate training (Kielhofner, 2006).

Considering the students' ratings in a «face value» perspective, there were different levels of interrater correspondence between the three subdomain scales. Physicality was, apparently, the easiest subdomain to agree on, whereas the relations and the information exchange subdomains showed lower measures of agreement. Although skills, according to the MOHO (Kielhofner, 2008) and ACIS (Forsyth et al., 1998) definitions are observable actions, skills in the physicality subdomain may be more clearly observable than skills in the other two domains, and therefore easier to agree on. It may be harder to agree on observations of «expresses» and «sustains» (in the information exchange subdomain) and «relates» and «respects» (in the relations subdomain), compared to observations of «contacts» and «gestures» (in the physicality subdomain). The latter skills clearly include physical action and body movements, whereas the former skills are more open to interpretation. Some interpretation may be needed to determine if a certain skill was performed as part of a person's behavior, before eventually assessing the quality of the skill performance.

Across all skills in all of the ACIS subdomains, however, observers are continually challenged to weigh the relative importance of several specific skill observa-

tions with a view to how they contribute to, or detract from, the successful completion of the activity and the ongoing social interaction (Ellingham et al., 2014; Forsyth et al., 1998). For example, a person may ask questions (the skill «asks» in the information exchange subdomain) several times during a half-hour observation, sometimes rather appropriately, at other times not so appropriately. How such multiple observations are transformed into a single, abstracted score for the skill «asks» is a matter of skilled judgement on the part of the observer. The observer should also be aware of the limitations of his or her skills assessment, as its validity may not extend beyond the current situation (Haglund & Thorell, 2004).

This study was performed in conjunction with students' training in using the ACIS. In fact, the data are based on the students' two training observations. Owing to the small amount of data available, it was not possible to assess whether the interrater reliability increased from the first to the second observation. Hopefully, more assessment experience and the ability to discuss experiences related to the transformation of activity observations into skill ratings, should translate into more skilled assessment practice – and consequently, into improved correspondence between raters. Further studies may not only assess how much training is necessary to have raters produce similar and correct scores, but may also consider how more training, and perhaps various modes of training (e.g., video rating, discussion with expert raters, etc.), may increase the reliability and correctness of scores.

STUDY LIMITATIONS

The study is limited in several ways. The participants were occupational therapy students, and thus not fully qualified therapists. This sample characteristic is, therefore, a limitation in itself. The small sample size, and thus the small amount of data, limits the generalizability of the results. Moreover, it prohibited a meaningful comparison of interrater correspondence between the first and the second observation, which would otherwise be an interesting line of inquiry. Also, the data as based on observations of fellow students roleplaying a character with skills limitations related to communication and interaction. It is possible that this «laboratory experience» setting detracts from the reliability of the results, but there appears to be no reported study using a similar study design, against which our study results can be directly compared.

Interrater reliability was assessed, whereas reliability in the sense of «correctness» was not. Thus, we do not know the extent to which the students' ratings were justified. The students were instructed not to compare their scores before discussing them, and were similarly instructed not to change their individual scores as a result of the discussion. However, we had limited control over the students' behaviors in this respect. It is also possible that those who did not agree to participate in the study were the ones whose scores diverted most from the other group members' scores. If so, such a sample bias would contribute to inflate the measures of interrater agreement.

The study addressed interrater correspondence on the scale

level. As the scales consist of several items, it is possible that multiple inconsistencies between raters have cancelled each other out in the analysis, and that this effect has played a part in producing the very high measures of interrater reliability. Further studies may examine interrater reliability at the more detailed level concerning each of the 20 specific skills listed in the ACIS.

CONCLUSION

The design of the study indicates that the produced results concerning interrater correspondence should be considered preliminary. However, we found a high level of correspondence between the students' ratings with the ACIS after only a very brief training session. We interpret these results as promising in terms of the comprehensibility and feasibility of the ACIS for new and inexperienced users, as well as when considering the amount of training needed for observers.

References

- Bonsaksen, T., Celo, C., Myraunet, I., Granå, KE., & Ellingham, B. (2013). Promoting academic-practice partnerships through students' practice placement. *International Journal of Therapy and Rehabilitation*, 20(1), 33-39.
- Bonsaksen, T., Granå, KE., Celo, C., Ellingham, B., & Myraunet, I. (2013). A practice placement design facilitating occupational therapy students' learning. *Journal of Mental Health Training, Education and Practice*, 8(4), 169-180. doi: 10.1108/JMH-TEP-08-2012-0028
- Bonsaksen, T., Kvarsnes, H., & Dahl, M.

- (2015). Who wants to go to occupational therapy school? Characteristics of Norwegian occupational therapy students (in press). *Scandinavian Journal of Occupational Therapy*.
- Bonsaksen, T., Lerdal, A., & Fagermoen, MS. (2012). Factors Associated with Self-Efficacy in Persons with Chronic Illness. *Scandinavian Journal of Psychology*, 53(4), 333-339. doi: 10.1111/j.1467-9450.2012.00959.x.
- Bonsaksen, T., Myraunet, I., Celo, C., Granå, KE., & Ellingham, B. (2011). Experiences of occupational therapists and occupational therapy students in using the Assessment of Communication and Interaction Skills in mental health settings in Norway. *British Journal of Occupational Therapy*, 74(7), 332-338. doi: 10.4276/03080221X1309951366117
- Crepeau, EB. (2003). Analyzing occupation and activity: A way of thinking about occupational performance. In EB Crepeau, ES Cohn & BAB Schell (Eds.), *Willard & Spackman's Occupational Therapy* (10 ed., pp. 189-198). Philadelphia: Lippincott Williams & Wilkins.
- Donohue, MV. (2007). Interrater reliability of the Social Profile: Assessment of community and psychiatric group participation. *Australian Journal of Occupational Therapy*, 54(1), 49-58. doi: 10.1111/j.1440-1630.2006.00622.x
- Ellingham, B., Hussain, RA., & Bonsaksen, T. (2014). *User's manual for the assessment of communication and interaction skills [Brukermanual for kartlegging av ferdigheter i kommunikasjon og interaksjon; in Norwegian]*. Oslo: Oslo and Akershus University College of Applied Sciences.
- Field, A. (2005). *Discovering Statistics Using SPSS* (2 ed.). London: Sage Publications.
- Forsyth, K., Lai, JS., & Kielhofner, G. (1999). The Assessment of Communication and Interaction Skills (ACIS): Measurement Properties. *British Journal of Occupational Therapy*, 62(2), 69-74.
- Forsyth, K., Salamy, M., Simon, S., & Kielhofner, G. (1998). *A User's Guide to the Assessment of Communication and Interaction Skills (ACIS)*. Version 4. Chicago, USA: University of Illinois.
- Fuller, K. (2011). The effectiveness of occupational performance outcome measures within mental health practice. *British Journal of Occupational Therapy*, 74(8), 399-405. doi: 10.4276/03080221X13125646371004
- Haglund, L., & Thorell, LH. (2004). Clinical perspective on the Swedish version of the Assessment of Communication and Interaction Skills: Stability of assessments. *Scandinavian Journal of Caring Sciences*, 18, 417-423.
- Hsu, WL., Pan, AW., & Chen, TJ. (2008). A psychometric study of the Chinese version of the Assessment of Communication and Interaction Skills. *Occupational Therapy in Health Care*, 22(2-3), 177-185. doi: 10.1080/07380570801991818
- Kielhofner, G. (2006). *Research in occupational therapy: Methods of inquiry for enhancing practice*. Philadelphia, PE: F. A. Davis Company.
- Kielhofner, G. (2008). *A Model of Human Occupation. Theory and Application*. (4 ed.). Baltimore, MD: Lippincott Williams & Wilkins.
- Kjellberg, A., & Haglund, L. (2015). Utilization of the Swedish version of the Assessment of Communication and Interaction Skills (e-pub ahead of print). *British Journal of Occupational Therapy*. doi: 10.1177/0308022615580328
- Kjellberg, A., Haglund, L., Forsyth, K., & Kielhofner, G. (2003). The measurement properties of the Swedish version of the assessment of communication and interaction skills. *Scandinavian Journal of Caring Sciences*, 17(3), 271-277.
- Laver-Fawcett, A. (2014). Routine standardised outcome measurement to evaluate the effectiveness of occupational therapy intervention: essential or optional? *Ergoterapeuten*, 57(4), 28-37.
- Nielsen, KK., & Andersen, J. (2006). ACIS' feasibility in relation to people with mental illness (ACIS' anvendelighed - i forhold til mennesker med sindslidelser; in Danish). *Ergoterapeuten* (Danish Journal of Occupational Therapy), 5, 21-23.
- Oslo and Akershus University College. (2011). *Bachelor Programme in Occupational Therapy*. Oslo: Oslo and Akershus University College.
- Petersen, K., & Hartvig, B. (2008). A process for translating and validating Model of Human Occupation assessments in the Danish context. *Occupational Therapy in Health Care*, 22(2-3), 139-149. doi: 10.1080/07380570801991784
- Shrout, PE., & Fleiss, JL. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 85(2), 420-428.
- Simmons, CD., Griswold, LA., & Berg, B. (2010). Evaluation of social interaction during occupational engagement. *American Journal of Occupational Therapy*, 64(1), 10-17.
- Stigen, L., & Page, J. (2012). Inter-rater reliability of the Norwegian version of the Perceive, Recall, Plan and Perform System of Task Analysis. *Ergoterapeuten*, 55(4), 40-47.
- Streiner, DL., & Norman, GR. (2008). *Health measurement scales - a practical guide to their development and use* (Vol. 4). Oxford: Oxford University Press.
- Thomas, H. (2012). *Occupation-based activity analysis*. Thorofare, NJ: SLACK Inc.