

Article

Validation of the Social Inclusion Scale with Students

Ceri Wilson^{1,2,*} and Jenny Secker^{1,2}

¹ Faculty of Health Social Care and Education, Anglia Ruskin University, Chelmsford, Essex, CM11SQ, UK;
E-Mails: ceri.wilson@anglia.ac.uk (C.W.), jenny.secker@anglia.ac.uk (J.S.)

² South Essex Partnership University NHS Foundation Trust, The Lodge, The Chase, Wickford, Essex, SS117XX, UK

* Corresponding author

Submitted: 11 June 2014 | In Revised Form: 26 August 2014 | Accepted: 5 September 2014 |

Published: 24 July 2015

Abstract

Interventions (such as participatory arts projects) aimed at increasing social inclusion are increasingly in operation, as social inclusion is proving to play a key role in recovery from mental ill health and the promotion of mental wellbeing. These interventions require evaluation with a systematically developed and validated measure of social inclusion; however, a “gold-standard” measure does not yet exist. The Social Inclusion Scale (SIS) has three subscales measuring social isolation, relations and acceptance. This scale has been partially validated with arts and mental health project users, demonstrating good internal consistency. However, test-retest reliability and construct validity require assessment, along with validation in the general population. The present study aimed to validate the SIS in a sample of university students. Test-retest reliability, internal consistency, and convergent validity (one aspect of construct validity) were assessed by comparing SIS scores with scores on other measures of social inclusion and related concepts. Participants completed the measures at two time-points seven-to-14 days apart. The SIS demonstrated high internal consistency and test-retest reliability, although convergent validity was less well-established and possible reasons for this are discussed. This systematic validation of the SIS represents a further step towards the establishment of a “gold-standard” measure of social inclusion.

Keywords

internal consistency; mental health; psychometrics; scale validation; social inclusion; Social Inclusion Scale; test-retest reliability; validity

Issue

This article is part of the special issue “Indicators and Measurement of Social Inclusion”, edited by Professor Peter Huxley (Bangor University, UK).

© 2015 by the authors; licensee Cogitatio (Lisbon, Portugal). This article is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1. Introduction

This paper describes the validation of the Social Inclusion Scale (SIS) with a sample of university students. By way of introduction the concept of social inclusion and previous development of the SIS are outlined. The methods used for the present study are then described, covering participant recruitment and power analysis, the materials selected for comparison with the SIS (Berry, Rodgers, & Dear, 2007; Huxley et al., 2012; Tennant et al., 2007) and the procedures for data

collection and analysis. Results are presented in relation to internal consistency, test-retest reliability and convergent validity (one aspect of construct validity).

1.1. Social Inclusion

Social inclusion is defined in the European Union as having the opportunities and resources to participate fully in economic, social and cultural life and to enjoy a standard of wellbeing that is considered normal in the society in which we live (Commission of the European

Communities, 2000). Three dimensions of inclusion have been identified: experiences of friendships, feeling worthwhile through meaningful activities and hopefulness (Davidson, Stayner, Nickou, Styron, & Chinman, 2001). Social inclusion is a multidimensional concept encompassing physical aspects (e.g., housing), psychological aspects (e.g., a sense of belonging), social aspects (e.g., friendships), and occupational aspects (e.g., leisure: Le Boutillier & Croucher, 2010). Social inclusion can increase mental health and reduce mental illness, help to promote recovery, and provide both mental and physical health gains (e.g., Boardman, 2003; Waddell & Burton, 2006; Whiteford, Cullen, & Baingana, 2005).

Social inclusion overlaps with the following concepts: social quality; social participation; citizenship; social capital; social networks; wellbeing and quality of life. Understanding of the concept of social inclusion and its relationship to these concepts is at an early stage (McKenzie, Whitely, & Weich, 2002). There is also debate about the relationship between social inclusion and social exclusion. Although often viewed as opposite ends of a single continuum, when conceptualisations of social exclusion are examined these commonly identify dimensions that focus on the one hand on political, economic and social structures, and on the other hand on social relationships based on mutual acceptance (Burchardt, Le Grand, & Piachaud, 2002; Huxley & Thornicroft, 2003). Thus Secker (2010) argues that it is more helpful to separate out the two concepts and to think of exclusion as operating on a structural level through barriers that work to exclude individuals and groups from full participation in society, akin to the complex disadvantage described by Jenkins (2011), while inclusion operates on an individual or group level and relates to the extent to which people are accepted and feel they belong within different social contexts.

As yet there is no “gold-standard” measure of social inclusion (Huxley et al., 2012; Secker, Hacking, Kent, Shenton, & Spandler, 2009). This is problematic, as social inclusion is proving to play a key role in the establishment of mental health and wellbeing; which in turn has led to interventions (such as community-based arts-for-health activities) aimed at increasing social inclusion; which require evaluation with a systematically developed and validated measure of social inclusion.

1.2. Development of the Social Inclusion Scale (SIS)

In order to address this problem, Secker et al. (2009) recently developed the SIS. Secker and colleagues initially derived concepts associated with social inclusion from the literature (bonding and bridging social capital; social acceptance; neighbourhood cohesion and engagement in leisure and cultural activity; citizenship; perceived security of housing tenure; and occupational

activity), and modified questions from measures of these concepts to produce the SIS. This was then piloted with (and modified by) arts and mental health service users and a service user research group. The revised measure had 22 items and three subscales: Social Isolation, Social Relations and Social Acceptance, covering concepts relating to friendship and family, sense of belonging and social opportunities.

The SIS has so far been validated with 88 arts and mental health project users (encompassing a range of mental health difficulties). The SIS as a whole ($\alpha = 0.85$) and its separate subscales (Social Isolation = 0.76; Social Acceptance = 0.76; Social Relations = 0.70) demonstrated good internal consistency (Secker et al., 2009). Secker et al. (2009) also assessed how the scale related to two measures of similar constructs: the Clinical Outcomes in Routine Evaluation (CORE: a measure of overall mental health; Evans et al., 2000) and a measure of empowerment (Schafer, 2000). The SIS was significantly correlated with both of these measures (with poor mental health associated with low levels of social inclusion, and high levels of social inclusion associated with high levels of empowerment). In addition, a shortened version of the SIS (12 items) has been shown to be responsive to change over time brought about by a specific intervention designed to promote social inclusion (Margrove, SE-SURG, Heydinrych, & Secker, 2013). However, test-retest reliability and construct validity of the SIS requires assessment (Huxley et al., 2012; Secker et al., 2009). Furthermore, Huxley et al. (2012) and Secker et al. (2009) have pointed out the need to validate the SIS in the general population.

1.3. Study Aims

The aim of the present research was to validate the full and shortened versions of the SIS in a non-clinical population of university students, by establishing its internal consistency, convergent validity and test-retest reliability. If the measure demonstrates its suitability for use in this population, this will add to the evidence base for the SIS and contribute towards establishing a “gold-standard” measure of social inclusion.

2. Methods

2.1. Recruitment

Following receipt of ethical approval from the Anglia Ruskin University (ARU) Faculty Research Ethics Committee, students at ARU were invited to take part via faculty emails lists, posters and flyers around the university campuses. The inclusion criteria required participants to be aged 18 and over, and a student at ARU (full-time or part-time, undergraduate or postgraduate).

2.2. Power Analysis

Cohen's Statistical Power Analysis (Cohen, 1988, 1992) is one of the most popular approaches in the behavioural sciences in calculating the required sample size (e.g., Cappelleri & Darlington, 1994; Chuan, 2006). According to Cohen Statistical Power Analysis for Pearson Product Moment Correlation, a sample of 85 participants is necessary for a power of 0.80 ($B = 0.2$), for a medium effect size ($r = 0.30$) and a significant alpha of 0.05. The number of participants completing both questionnaire packs exceeded this target (see Section 3.1).

2.3. Materials

2.3.1. The Social Inclusion Scale (SIS: Secker et al., 2009)

The SIS was developed during the national study of arts and mental health (Hacking, Secker, Spandler, Kent, & Shenton, 2008; Secker et al., 2009). The measure has 22 items and three subscales: Social Isolation, Social Relations and Social Acceptance. Two items fitted well into two subscales and were therefore included in both scales when calculating subscale scores; three items did not fit into any subscale and were included as separate items (see Table 1). The concepts underpinning the subscales are as follows: social isolation refers to the amount of contact an individual has with people and society; social acceptance refers to a person's sense of being accepted within their social contexts; and social relations refers to relationships between people. This measure has demonstrated good internal consistency and concurrent validity amongst arts and mental health project users (Secker et al., 2009). The scale consists of statements in which participants choose the option on a Likert scale ("*Not at all*", "*Not particularly*", "*Yes a bit*" and "*Yes definitely*") that best describes their relationships with other people over the last month. The overall score is the sum of each item; the score of each subscale is the sum of each item in that subscale. As the SIS was originally developed for use in a mental health sample, the wording of three of the questions in this scale was not suitable for the present population. Therefore, the wording of two questions was modified: "*I have been involved in a group not just for mental health*" was changed to "*I have been involved in a group not just for my university studies*" (which can be modified depending on the group of interest); and "*I have felt some people look down on me because of my mental health needs*" was changed to "*I have felt some people look down on me because of how I am*". One question was removed due to lack of relevance to the population sample: "*my social life has been mainly related to mental health, or people who use mental health services.*"

Table 1. SIS items and subscales.

Subscale	Item
<i>Social Isolation</i>	I have felt terribly alone and isolated
	I have felt accepted by my friends
	I have been out socially with friends
	I have felt I am playing a useful part in society*
<i>Social Relations</i>	I have friends I see or talk to every week*
	I have felt I am playing a useful part in society*
	I have felt what I do is valued by others
	I have been to new places
	I have learnt something about other cultures
	I have been involved in a group not just for my university studies
	I have done some cultural activity
<i>Social Acceptance</i>	I have felt some people look down on me because of how I am
	I have felt unsafe to walk alone in my neighbourhood in daylight
	I have friends I see or talk to every week*
	I have felt accepted by my neighbours
<i>Individual Items</i>	I have felt accepted by my family
	I have felt clear about my rights
	I have felt free to express my beliefs
<i>Individual Items</i>	I have felt insecure about where I live
	I have done a sport, game or physical activity
	I have helped out at a charity or local group

* Items in more than one subscale.

2.3.2. Selection of Comparison Measures

Convergent validity (one aspect of construct validity) is when several different methods for obtaining the same information about a given trait or concept produce similar results. Validity is usually expressed as a correlation coefficient between two sets of data (levels ≥ 0.7 are generally accepted as representing good validity, e.g., Litwin, 1995). In order to assess the convergent validity of the SIS, correlations would need to be carried out between scores on the SIS and scores on other already validated measures of social inclusion. As previously discussed, there is no "gold-standard" measure of social inclusion. Following a literature search, currently available measures of social inclusion (and related concepts) were identified. Those measures deemed most suitable were the Satisfaction with Opportunities (Sat Opps) and Perceived Opportunities (Perceived Opps) subscales of the SCOPE, the 15-item "Big 7" ACPQ, the Perceived Time and Perceived Enjoyment subscales of the ACPQ. A number of other measures were located; however, the majority were unsuitable

for the following reasons: they were developed for an evaluation of a specific service (e.g., Bates, 2005; Davis & Lindley, 1999; Marino-Francis & Worrall-Davies, 2010; Williamson & Allen, 2006); they were developed specifically for use in, and/or only validated within, a different country and/or culture (e.g., Lev-Wiesel, 2003; Van Brakel, Anderson, & Mutatkar, 2006); were developed for use in a specific population not relevant to the present research (e.g. Sibley et al., 2006, developed a measure to evaluate social inclusion/exclusion as a result of a physical disability); consisted of a measure of exclusion as opposed to inclusion (e.g., De Jong, Gierveld, & van Tilburg, 2006); had shown poor validity/reliability (e.g., Lelieveldt, 2004); had not been validated at all (e.g., Stickley & Shaw, 2006); or were designed as an interview and were not suitable for adaptation to a self-report questionnaire (e.g., Gordon et al., 1999; Mezey et al., 2012). The Resource Generator-UK (RG-UK) (Webber & Huxley, 2007), a measure of social capital, was also located. The RG-UK has been validated in a UK general population sample (and demonstrated good test-retest reliability, internal consistency, and concurrent validity). The measure asks respondents about access to a fixed list of social resources that represent multiple domains of social capital and their relationship to the person through whom they could access that resource. It is a 27-item self-report questionnaire with four subscales (domestic resources, expert advice, personal skills, and problem solving) with a yes/no response format. However, one key difference between this measure and the SIS is that it has no questions relating to emotional support (which means it is not entirely comparable). As well as the collection of such “objective” data, understanding is also needed of whether individuals actually feel included (i.e., “subjective” data). How individuals feel is thought to be a critical factor in their social inclusion (e.g., Sayce & Morris, 1999), and mental health service users have acknowledged that their internal world affects whether they feel socially included or excluded (Smyth, Harries, & Dorer, 2011). Prince and Prince (2002) have also found that individuals with mental health difficulties experience fear and rejection and lack a sense of connection and belonging, despite being physically involved in their community. Such individuals are, therefore, not socially included unless they feel included (Onken, Craig, Ridgway, Ralph, & Cook, 2007; Pinfold, 2000). It is also worth noting that the problem solving scale reliability was poor ($r = 0.35$). Therefore, this measure was also not included in the present validation study.

It is also important to assess whether the SIS significantly relates to validated measures of similar constructs. Therefore, the SIS was compared with a validated measure of mental wellbeing (the WEMWBS), given that greater social inclusion is associated with greater mental wellbeing. The selected measures are described below.

2.3.3. The Social and Community Opportunities Profile (SCOPE: Huxley et al., 2012)

The SCOPE includes both objective and subjective measures of social inclusion. The SCOPE short version (48 items) has shown good internal consistency (overall), good test-retest reliability (in a student sample at two week follow-up), and was able to discriminate between mental health groups. The SCOPE consists of various response formats (for example individual items consist of yes/no responses, or open response). In addition to these individual items, there are two subscales: SatOpps and Perceived Opps. SatOpps items ($n = 11$) are measured on a seven-point Likert Scale (ranging from “*delighted*” to “*terrible*”: $\alpha = 0.77$) and Perceived Opps items ($n = 5$) are measured on a five-point Likert scale (ranging from “*opportunities are extremely restricted*” to “*there are plenty of opportunities*”: $\alpha = 0.62$). As the reliability and consistency of the SatOpps and Perceived Opps subscales have been assessed (and found to be acceptable) only these items were included in the present study (16 items).

2.3.4. The Australian Community Participation Questionnaire (ACPQ: Berry et al., 2007)

The original 67-item ACPQ is made up of 14 factors, which consist of different types of community participation. However, only seven types of community participation (contact with immediate household, extended family, friends, and neighbours; organised community activities; religious observance; and active interest in current affairs) were found to be significantly associated with fewer symptoms of distress (and this finding was later replicated: Berry & Welsh, 2010). These were named the “Big 7”. Upon contacting the author the 15-item version of the ACPQ (which was also used by Huxley et al. (2012) as a comparative measure to assess the construct validity of the SCOPE) was provided. This measure contains three overarching dimensions of community participation: informal social connectedness, civic engagement, and political participation. This measure was found to significantly correlate with both the SatOpps and Perceived Opps subscales of the SCOPE short version. In a further study that has employed the ACPQ (Berry & Shipley, 2009), subjective perceptions about community participation were added (including questions about thoughts and feelings about participation), which were recommended for use by the author. These questions tell you whether the respondent feels he/she is spending sufficient time participating irrespective of how much time he/she is actually spending participating. This is, therefore, a measure of perceived gap between the degree of inclusion desired and the degree that is thought to actually occur. Although internal consistency and relation to a related construct (distress) has been assessed for

the shortened version of the ACPQ; the test-retest reliability has not been assessed. It is also worth noting that this was validated in an Australian sample, which may mean it is not as suitable for a UK sample. Therefore, the present research assessed the internal consistency and test-retest reliability of the measure in a UK sample.

2.3.5. The Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS: Tennant et al., 2007)

The WEMWBS measures positive affect, psychological functioning and interpersonal relationships. Mental wellbeing is more than the absence of mental illness, and the scale covers only positive aspects of mental health. A measure of mental wellbeing was chosen in preference to a measure of mental ill health, in order to correspond with the increasing emphasis in health policy on promoting positive mental health (Department of Health, 2011). The WEMWBS is an ordinal scale consisting of 14 positively phrased statements rated on Likert scales: “None of the time”, “Rarely”, “Some of the time”, “Often” and “All of the time”. The overall score is the sum of each item with a higher score reflecting higher mental wellbeing. This scale has demonstrated high internal consistency, construct validity, discriminant validity, and test-retest reliability across a range of populations (e.g., Bartram, Yadegarfar, Sinclair, & Baldwin, 2011; Clarke et al., 2011, Tennant et al., 2007).

2.4. Hypotheses

It was firstly hypothesised that Cronbach's alpha for the SIS as a whole, its three subscales, and the short-form version would be ≥ 0.70 demonstrating good internal consistency (e.g., Carmines & Zeller, 1991; Litwin, 1995). Secondly, it was hypothesised that SIS scores at both time points (one-to-two weeks apart) would be significantly positively correlated (correlation coefficient ≥ 0.70 ; demonstrating good test-retest reliability). The third hypothesis was that scores on the SIS would be significantly correlated with scores on the selected measures of related concepts (with correlation coefficients ≥ 0.70), demonstrating good convergent validity.

2.5. Procedure

An invitation email with the information sheet attached was sent to students at ARU via faculty email lists and a research student mailbase. The information sheet outlined the purpose of the study, what participation would involve, and reassured of the right to refuse to participate (with no action required if they did not wish to participate). Participants were asked to email the researcher if they would like to take part. The researcher then emailed Questionnaire Pack 1 to the participant (consisting of the SIS, the SCOPE Sat Opps and Perceived Opps subscales, the 15-item “Big 7”

ACPQ, the Perceived Time and Perceived Enjoyment subscales of the ACPQ, and the WEMWBS). The first page of the questionnaire also asked for demographic information (age, gender and ethnicity). Receipt of a completed questionnaire was taken as consent to participate (as explained in the information sheet). Upon receipt of the questionnaire each participant was allocated a participant code that was used on all questionnaire materials. A password-protected computer database (which matched participants' email addresses and participant codes) was used in order to email Questionnaire Pack 2 (identical to Questionnaire Pack 1) to each participant seven days after receiving their Questionnaire Pack 1. The code was also used to match up Questionnaire Packs 1 and 2 in the data analysis. The researcher requested receipt of the completed Questionnaire Pack 2 within seven days. If it was not received within seven days one email reminder was sent. Upon receipt of Questionnaire Pack 2 participants were thanked for their participation. At the end of data collection a prize draw took place in which ten participants were randomly selected to receive a £20 thank-you voucher (and were notified via email).

2.6. Data Analysis

Prior to data analysis, the questionnaire data were checked for normality via histograms, box-plots, and skewness and kurtosis z-scores ($< \pm 1.96$). Where data were normally distributed, Pearson's correlations were conducted. Where data were non-normally distributed, Spearman's rho correlations were conducted.

In order to calculate the internal consistency of the SIS, Cronbach's alpha was calculated for the scale as a whole and its three subscales. Cronbach's alpha ≥ 0.7 is generally considered as acceptable (e.g., Carmines & Zeller, 1991; Litwin, 1995; Nunnally, 1967; Zait & Berteau, 2011).

In order to assess the test-retest reliability of the SIS, questionnaire scores (total and subscale scores) at time point one were correlated with scores at time point two. Assessing correlation coefficients is the standard measure of test-retest reliability, with r values ≥ 0.7 considered as indicating good test-retest reliability (e.g., Carmines & Zeller, 1991; Litwin, 1995).

In order to assess the convergent validity of the scale, correlations were carried out between SIS scores (overall and subscales) and SCOPE scores (SatOpps and Perceived Opps subscales: 16 items), “Big 7” ACPQ and Perceived Time and Perceived Enjoyment scores (29 items), and WEMWBS scores (14 items).

3. Results

3.1. Sample

103 participants completed the questionnaires at time

1, and 95 participants completed the questionnaires at time 2 (92.23% response rate at time 2). The mean number of days between questionnaires was 9.67 ($SD = 3.08$), with a minimum of seven days and a maximum of 19 days. Out of the 103 participants at time one 22 participants were male (21.6%) and 80 participants were female (78.4%), with one participant not stating their gender. Participant's mean age was 31.37 ($SD = 13.04$) ranging from age 18 to 66 (three participants did not state their age). Four participants did not state their ethnicity, the remaining 99 participants defined their ethnicity as follows: 65.7% White British ($n = 65$); 6.1% British Asian ($n = 6$); 5.1% White ($n = 5$); 4% African ($n = 4$); 4% Asian ($n = 4$); 3% Mixed ($n = 3$); 3% Black ($n = 3$); 2% White Irish ($n = 2$); 2% White European ($n = 2$); 2% Southern European ($n = 2$); 1% Eastern European ($n = 1$); 1% Middle Eastern ($n = 1$); and 1% Non-British Irish ($n = 1$). The make-up of the 95 participants who completed both questionnaire packs was as follows: 20.2% male ($n = 19$) and 79.8% female ($n = 75$), mean age = 31.87 ($SD = 13.34$), 67% White British.

3.2. Internal Consistency

Cronbach's alpha was calculated for scores at time 1 in order to assess the internal consistency of the measures. Cronbach's alpha for the full-version SIS was 0.80 (indicating high internal consistency). Internal consistency was subsequently assessed for the SIS with the three individual items which did not fit into the subscales removed. Cronbach's alpha remained unchanged. Cronbach's alpha for the Social Isolation subscale was 0.65, for the Social Acceptance subscale Cronbach's alpha was 0.54, and for the Social Relations subscale Cronbach's alpha was 0.71. The 12-item short-form version of the SIS demonstrated high internal consistency (0.75). See Table 2 for internal consistency statistics for each questionnaire measure.

Cronbach's alpha was also calculated using "scale if item deleted" to assess whether the SIS subscales would have greater internal consistency if any items were deleted (see Table 3). As can be seen in Table 3, the internal consistency of the Social Isolation subscale did not improve with any items deleted. However, this was not the case with the other two subscales. Cronbach's alpha for the Social Relations subscale was found to increase upon the removal of two of its items ("*I have felt some people look down on me because of how I am*" and "*I have felt unsafe to walk alone in my neighbourhood*"). When both items were removed, Cronbach's alpha increased further to 0.74. The Social Acceptance subscale demonstrated the poorest internal consistency (0.54), however this increased to 0.63 upon removal of one item ("*I have friends I see or talk to every week*").

Table 2. Cronbach's alpha for each questionnaire at time 1.

Measure	Cronbach's alpha
SIS Total Mean	0.80*
Social Acceptance (SIS subscale)	0.54
Social Relations (SIS subscale)	0.71*
Social Isolation (SIS subscale)	0.65
SIS 12-Item Short Form	0.75*
SatOpps (SCOPE subscale)	0.80*
Perceived Opps (SCOPE subscale)	0.54
15-item "Big 7" (ACPQ)	0.80*
Perceived Time (ACPQ)	0.65
Perceived Enjoyment (ACPQ)	0.47
WEMWBS	0.90*

* ≥ 0.70 .

Table 3. Cronbach's alpha using "scale if item deleted".

SIS subscale	Cronbach's alpha
Social Isolation Total (5 items)	0.65
Item 1 deleted	0.59
Item 2 deleted	0.61
Item 3 deleted	0.65
Item 4 deleted	0.57
Item 5 deleted	0.55
Social Relations Total (8 items)	0.71
Item 1 deleted	0.65
Item 2 deleted	0.65
Item 3 deleted	0.69
Item 4 deleted	0.69
Item 5 deleted	0.68
Item 6 deleted	0.65
Item 7 deleted	0.72*
Item 8 deleted	0.72*
Social Acceptance Total (5 items)	0.54
Item 1 deleted	0.63*
Item 2 deleted	0.38
Item 3 deleted	0.50
Item 4 deleted	0.44
Item 5 deleted	0.39

* improved internal consistency when removed.

3.3. Test Re-Test Reliability

Some of the questionnaire data at time points 1 and 2 were non-normally distributed; therefore, Spearman's correlations were carried out for mean questionnaire scores at times 1 and 2 (see Table 4). Alpha was adjusted to 0.005 in order to account for multiple correlations ($0.05/11=0.005$: to account for 11 correlations).

3.4. Convergent Validity

In order to assess the convergent validity of the SIS, Spearman's correlations were carried out between the SIS (overall mean scores, subscale mean scores, and short-form mean scores) and all other measures at

time 1 (see Table 5). Significant alpha was adjusted to 0.01 in order to account for five sets of scores (SIS overall mean, Social Acceptance mean, Social Isolation mean, Social Relations mean, and short-form mean) being correlated with each measure ($0.05/5 = 0.01$). For the Perceived Opps subscale of the SCOPE, lower scores reflect greater social inclusion and so a negative correlation with SIS scores was expected. Negative correlations were also expected between SIS scores and Perceived Time (ACPQ) scores, as lower scores on this scale also represent greater social inclusion.

Table 4. Test-retest reliability assessment for each measure at times 1 and 2.

Measure (T1 and T2)	r_s	p
SIS Total Mean	0.80*	<0.001**
Social Acceptance (SIS subscale)	0.72*	<0.001**
Social Relations (SIS subscale)	0.71*	<0.001**
Social Isolation (SIS subscale)	0.80*	<0.001**
SIS 12-item Short Form	0.73*	<0.001**
SatOpps (SCOPE subscale)	0.83*	<0.001**
Perceived Opps (SCOPE subscale)	0.78*	<0.001**
15-item "Big 7" (ACPQ)	0.88*	<0.001**
Perceived Time (ACPQ)	0.66	<0.001**
Perceived Enjoyment (ACPQ)	0.67	<0.001**
WEMWBS	0.79*	<0.001**

* $r_s \geq 0.70$; ** $p < 0.005$

4. Conclusions

Social inclusion is a multi-dimensional concept that involves having the opportunities and resources to participate fully in economic, social and cultural life, and enjoying a standard of wellbeing considered normal in the society in which we live (Commission of the European Communities, 2000). The SIS assesses multiple aspects of social inclusion, including psychological aspects (such as a sense of belonging), social aspects (such as relationships with family and friends), occupational aspects (such as engagement with culture and leisure activities), and physical aspects (such as security with one's living environment). The SIS measures the respondents' feelings and perceptions of their social inclusion in relation to these aspects, as an individual cannot be considered socially included unless they feel socially included (e.g., Onken et al., 2007; Pinfold, 2000). Raising levels of social inclusion has become an important part of the mental health strategy as it has been found to increase mental health and wellbeing. Increasing concerns regarding levels of social inclusion in mental health populations have led to a number of interventions (such as participatory arts projects) aimed at increasing social inclusion. Outcome evaluation of such interventions is essential for assessing whether they are successful in achieving their aims. Many participatory arts project evaluations as-

sess the effect of their interventions on wellbeing and quality of life (e.g., Hillman, 2002), of which there are already validated measures. However, these projects also aim to improve social inclusion of which there is presently no "gold-standard" validated measure. It has been acknowledged that there is a lack of evidence about the efficacy of such interventions in increasing service users' social inclusion, largely due to a lack of suitable measure (e.g., Bates & Repper, 2001; Dorer, Harries, & Marston, 2009). Therefore, a validated "gold-standard" measure of social inclusion is required. The SIS had already been partially validated (having demonstrated good internal consistency, concurrent validity, and responsiveness to change in a mental health population), but required further reliability and validity testing (e.g. test-retest reliability and construct validity), in both mental health service user and general population samples. The present research study aimed to validate the SIS in a sample of university students, by establishing its internal consistency, convergent validity and test-retest reliability. It was firstly hypothesised that the SIS as a whole, its three subscales, and the 12-item short-form version would demonstrate high internal consistency. As expected, the SIS as whole (Cronbach's alpha = 0.80), the Social Relations subscale (Cronbach's alpha = 0.71) and the short-form version (Cronbach's alpha = 0.75) demonstrated high internal consistency. The Social Isolation (0.65) and Social Acceptance (0.54) subscales, however, did not demonstrate acceptable internal consistency. The Social Acceptance subscale (which showed the lowest internal consistency) did demonstrate improved internal consistency (0.63) with one item deleted. Therefore, in future use of the SIS in a non-clinical population, the item "*I have friends I see or talk to every week*" should be retained, but should only be used in the calculation of the Social Isolation subscale not in the Social Acceptance subscale. It is also worth considering the complete removal of the three items which did not fit into the three subscales in previously reported analysis (Secker et al., 2009). This is due to the present finding that the internal consistency of the SIS remained unchanged upon the removal of these items.

The ACPQ was validated in a UK sample for the first time, with the "Big 7" demonstrating high internal consistency (0.80) and test re-test reliability (0.88). However, the Perceived Time and Perceived Enjoyment ACPQ subscales did not demonstrate acceptable internal consistency or test-retest reliability (<0.70). This requires further investigation in order to assess whether these scales are suitable for use with UK samples.

Secondly, it was hypothesised that SIS scores at both time points (one-to-two weeks apart) would be significantly positively correlated with correlation coefficients ≥ 0.70 ; demonstrating good test-retest reliability. As expected the SIS as a whole, its three subscales and the short-form version demonstrated high test-retest reliability.

Table 5. Spearman's correlations between the SIS and other measures.

Measure	SIS Total Mean		Social Acceptance		Social Isolation		Social Relations		SIS Short Form	
	r_s	p	r_s	p	r_s	p	r_s	p	r_s	p
SatOpps	0.634	<0.001**	0.591	<0.001**	0.595	<0.001**	0.520	<0.001**	0.541	<0.001**
Perceived Opps	-0.360	0.001**	-0.282	0.004**	-0.384	0.001**	-0.303	0.002**	-0.274	0.007**
Big 7 ACPQ	0.589	<0.001**	0.476	<0.001**	0.518	<0.001**	0.432	<0.001**	0.571	<0.001**
Perceived Time	-0.309	0.002**	-0.177	0.075	-0.250	0.011	-0.286	0.004**	-0.242	0.019
Perceived Enjoyment	0.437	<0.001**	0.354	0.001**	0.395	<0.001**	0.351	<0.001**	0.482	<0.001**
WEMWBS	0.674	<0.001**	0.602	<0.001**	0.521	<0.001**	0.584	<0.001**	0.610	<0.001**

* $r_s \geq 0.70$; ** $p < 0.01$.

The third hypothesis was that scores on the SIS would be significantly correlated with scores on measures of related concepts, and that correlation coefficients would be greater than 0.70; demonstrating convergent validity. As expected the SIS as a whole was significantly positively correlated with SatOpps (SCOPE), Big 7 ACPQ, Perceived Enjoyment (ACPQ) and WEMWBS scores, and significantly negatively correlated with Perceived Opps (SCOPE) and Perceived Time (ACPQ) scores. However, none of the correlation coefficients were >0.70. Scores on the Social Acceptance SIS subscale and the short-form version of the SIS were significantly correlated (in expected directions) with all measures except the Perceived Time (ACPQ) subscale. Again all coefficients were <0.70. Scores on the Social Isolation and Social Relations SIS subscales were significantly correlated (in expected directions) with all measures. However, again all coefficients were <0.70.

The results demonstrate that the SIS as whole has high internal consistency and test-retest reliability, but convergent validity is less clear. This may be due to the lack of a "gold-standard" measure to use as a comparison: the comparative measures used were the most suitable measures available but have each only been partially validated (except for the WEMWBS). Therefore, it is not clear whether the lower than expected correlation coefficients are due to which measure. The WEMWBS is the only comparative measure used which has been thoroughly validated in various populations, and had the highest correlation coefficient with the SIS as a whole (0.674), only just falling short of the 0.70 cut-off for high validity. Following completion of another research project (validating the SIS with a mental health service user sample), the contrasted-groups approach will also be used to further assess the construct validity of the SIS. This is where two groups thought to be high and low in the construct being measured are compared on the measure of the construct. The mean scores of the two groups should differ significantly in the expected direction if the instrument is valid (e.g. DeVon et al., 2007). It is hypothesised that mental health service users will have significantly lower social inclusion scores than the present sample. Following collection of questionnaire data from the mental health service user sample, an inde-

pendent t-test (or a Mann Whitney *U* test if the data are non-normally distributed) will be used to assess whether there is a significant difference between scores on the SIS for these samples (as will be reported in a forthcoming article following the completion of data collection).

It is important to acknowledge some limitations of the present research study. Firstly, the sample consisted of only students from one university. However, the large sample (exceeding the required sample size) was representative of a range of ages and ethnicities (due to the inclusion of all types of students). However, the sample was not equally representative of males and females. One further limitation is that not all participants who completed questionnaire pack 2 completed and returned it within 14 days. However, results remained unchanged with the nine participants who exceeded the 14 days removed from the analysis.

In conclusion, the SIS demonstrates high internal consistency and test-retest reliability in a sample of university students. The ability to reach a conclusion regarding the convergent validity of the SIS is limited by the lack of available fully validated and relevant measures to compare the SIS with. Further analysis upon completion of this study with a mental health service user sample, will provide further insight into the construct validity of the SIS through contrasted-groups analysis (i.e., comparison between general population and mental health service user scores). The establishment of the SIS as a "gold-standard" measure of social inclusion is progressing. If the SIS continues to be demonstrated as a reliable and valid measure, it will serve as a useful tool in outcome evaluation of participatory arts projects (and other interventions intended to improve social inclusion).

Acknowledgements

The present research study was funded by ARU Faculty of Health, Social Care & Education Scholarship and Research Funding.

Conflict of Interests

The authors declare no conflict of interests.

References

- Bartram, D. J., Yadegarfar, G., Sinclair, J. M. A., & Baldwin, D. S. (2011). Validation of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) as an overall indicator of population mental health and well-being in the UK veterinary profession. *The Veterinary Journal*, *187*, 397-398.
- Bates, P. (2005). *The Inclusion Web*. London: National Development Centre.
- Bates, P., & Repper, J. (2001). Social inclusion—A framework for evaluation. *A Life in the Day*, *5*(1), 18-23.
- Berry, H. L., Rodgers, B., & Dear, K. B. G. (2007). Preliminary development and validation of an Australian community participation questionnaire: Types of participation and associations with distress in a coastal community. *Social Indicators Research*, *64*, 1719-1737.
- Berry, H. L., & Shipley, M. (2009). *Longing to Belong: Personal Social Capital and Psychological Distress in an Australian Coastal Region* (Social Policy Research Paper No. 39). Canberra: Commonwealth of Australia.
- Berry, H. L., & Welsh, J. A. (2010). Social capital and health in Australia: An overview from the household, income and labour dynamics in Australia survey. *Social Science & Medicine*, *70*, 588-596.
- Boardman, J. (2003). Work, employment and psychiatric disability. *Advances in Psychiatric Treatment*, *9*, 327-334.
- Burchardt T., Le Grand, J., & Piachaud, D. (2002). Degrees of exclusion: Developing a dynamic, multidimensional measure. In J. Hills, J. Le Grand, & D. Piachaud (Eds.), *Understanding Social Exclusion* (pp. 30,40). Oxford: Oxford University Press.
- Cappelleri, J. C., & Darlington, R. B. (1994). The power analysis of cutoff-based randomized clinical trials. *Evaluation Review*, *18*(2), 141-152.
- Carmines, E. G., & Zeller, R. A. (1991). *Reliability and Validity Assessment*. California: Sage Publications.
- Chuan, C. L. (2006). Sample size estimation using Krejcie and Morgan and Cohen statistical power analysis: A comparison. *Jurnal Penyelidikan IPBL*, *7*, 78-86.
- Clarke, A., Friede, T., Putz, R., Ashdown, J., Martin, S., Blake, A., Adi, Y., Parkinson, J., Flynn, P., Platt, S., & Brown, S. (2011). Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): Validated for teenage school students in England and Scotland. A mixed methods assessment. *BMC Public Health*, *11*, 487.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillside, NJ: Erlbaum.
- Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin*, *112*(1), 155-159.
- Commission of the European Communities. (2000). *Social Policy Agenda: Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions*. Brussels: Commission of the European Communities.
- Davidson, L., Stayner, D., Nickou, T., Styron, M., & Chinman, M. (2001). "Simply to be let in": Inclusion as a basis for recovery. *Psychiatric Rehabilitation Journal*, *24*(4), 375-388.
- Davis, F. A., & Lindley, J. (1999). The Support Needs Questionnaire. Available from Fabian Davis, Oxleas NHS Trust 100255.3512@compuserve.com.
- De Jong Gierveld, J., & van Tilburg, T. G. (2006). A 6-item scale for overall, emotional, and social loneliness: Confirmatory tests on survey data. *Research on Aging*, *28*, 582-598.
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., Savoy, S. M., & Kostas-Polston, E. (2007). A Psychometric Toolbox for Testing Validity and Reliability. *Journal of Nursing Scholarship*, *39*(2), 155-164.
- Department of Health. (2011). *No Health without Mental Health: A Cross-Government Mental Health Outcomes Strategy for People of all Ages*. London: Department of Health.
- Dorer, G., Harries, P., & Marston, L. (2009). Measuring social inclusion: A staff survey of mental health service users' participation in community occupations. *British Journal of Occupational Therapy*, *72*(12), 520-530.
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health*, *9*(3), 247-255.
- Gordon, D., Levitas, R., Pantazis, C., Payne, S., Townsend, P., Bradshaw, J., Middleton, S., Bramley, G., Bridgwood, A., Maher, J., & Rowlands, O. (1999). *The Poverty and Social Exclusion Survey of Britain Questionnaire*. Bristol: Townsend Centre for International Poverty Research.
- Hacking, S., Secker, J., Spandler, H., Kent, L., & Shenton, J. (2008). Evaluating the impact of participatory art projects for people with mental health needs. *Health and Social Care in the Community*, *16*(6), 638-648.
- Hillman, S. (2002). Participatory singing for older people: A perception of benefit. *Health Education*, *102*(4), 163-171.
- Huxley, P., Evans, S., Madge, S., Webber, M., Burchardt, T., McDaid, D., & Knapp, M. (2012). Development of a social inclusion index to capture subjective and objective life domains (Phase II): Psychometric development study. *Health Technology Assessment*, *16*(1), 1366-5278.
- Huxley, P., & Thornicroft, G. (2003). Social inclusion, social quality and mental illness. *British Journal of*

- Psychiatry*, 182, 289-290.
- Jenkins, S. P. (2011). *Changing Fortunes: Income Mobility and Poverty Dynamics in Britain*. Oxford: Oxford University Press.
- Le Boutillier, C., & Croucher, A. (2010). Social inclusion and mental health. *British Journal of Occupational Therapy*, 73(3), 136-139.
- Lelieveldt, H. T. (2004). Helping citizens help themselves. Neighborhood improvement programs and the impact of social networks, trust and norms on neighborhood oriented forms of participation. *Urban Affairs Review*, 39, 531-551.
- Lev-Wiesel, R. (2003). Indicators constituting the construct of "perceived community cohesion". *Community Development Journal*, 38, 332-343.
- Litwin, M. S. (1995). *How to Measure Survey Reliability and Validity*. California: Sage Publications.
- Margrove, K. L., SE-SURG, Heydinrych, K., & Secker, J. (2013). Waiting list-controlled evaluation of a participatory arts course for people experiencing mental health problems. *Perspectives in Public Health*, 133(1), 28-35.
- Marino-Francis, F., & Worrall-Davies, A. (2010). Development and validation of a social inclusion questionnaire to evaluate the impact of attending a modernised mental health day service. *Mental Health Review Journal*, 15(1), 37-47.
- McKenzie, K., Whitely, R., & Weich, S. (2002). Social capital and mental health. *British Journal of Psychiatry*, 181, 280-283.
- Mezey, G., White, S., Thachil, A., Berg, R., Kallumparam, S., Nasiruddin, O., Wright, C., & Killaspy, H. (2012). Development and preliminary validation of a measure of social inclusion for use in people with mental health problems: The SInQUE. *International Journal of Social Psychiatry*, 59(5), 501-507.
- Nunnally, V. W. (1967). *Psychometric Theory*. New York: McGraw-Hill.
- Onken, S., Craig, C., Ridgway, P., Ralph, R., & Cook, J. (2007). An analysis of the definitions and elements of recovery: A review of the literature. *Psychiatric Rehabilitation Journal*, 31(1), 9-22.
- Pinfold, V. (2000). "Building up safe havens...all around the world": Users' experiences of living in the community with mental health problems. *Health and Place*, 6(3), 201-212.
- Prince, P. N., & Prince, C. R. (2002). Perceived stigma and community integration among clients of assertive community treatment. *Psychiatric Rehabilitation Journal*, 25(4), 323-331.
- Sayce, L., & Morris, D. (1999). *Outsiders Coming In? Achieving Social Inclusion for People with Mental Health Problems*. London: Mind Publications.
- Schafer, T. (2000). Empowerment: Towards a participatory model for the evaluation of the empowering therapeutic environment. *Mental Health Care*, 3, 233-237.
- Secker, J. (2010). Mental Health Problems, Social Inclusion and Social Exclusion: A UK Perspective. In D. Pilgrim, A. Rogers, & B. Pescosolido (Eds.), *The SAGE Handbook of Mental Health and Illness*. London, Thousand Oaks and Delhi: Sage Publications.
- Secker, J., Hacking, S., Kent, L., Shenton, J., & Spandler, H. (2009). Development of a measure of social inclusion for arts and mental health project participants. *Journal of Mental Health*, 18(1), 65-72.
- Sibley, A., Kersten, P., Ward, C. D., George, S., White, B., & Mehta, R. L. (2006). Measuring autonomy in disabled people: Validation of a new scale in a UK population. *Clinical Rehabilitation*, 20, 793-803.
- Smyth, G., Harries, P., & Dorer, G. (2011). Exploring mental health service users' experiences of social inclusion in their community occupations. *British Journal of Occupational Therapy*, 74(7), 323-331.
- Stickley, T., & Shaw, R. (2006). Evaluating social inclusion. *Mental Health Practice*, 9(10), 14-20.
- Tennant, R., Hiller, L., Fishwick, R., Platt, S., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK validation. *Health and Quality of Life Outcomes*, 5, 63.
- Van Brakel, W., Anderson, A. M., & Mutatkar, R. K. (2006). The participation scale: Measuring a key concept in public health. *Disability Rehabilitation*, 28, 193-203.
- Waddell, G., & Burton, A. K. (2006). *Is Work Good for Your Health and Well-Being?* London: The Stationary Office.
- Webber, M., & Huxley, P. (2007). Measuring access to social capital: The validity and reliability of the Resource Generator-UK and its association with common mental disorder. *Social Science & Medicine*, 65, 481-492.
- Whiteford, H., Cullen, M., & Baingana, F. (2005). Social Capital and Mental Health. In H. Hermann, S. Saxena, & R. Moodie (Eds.), *Promoting Mental Health: Concepts, Emerging Evidence, Practice*. Geneva: World Health Organization.
- Williamson, M., & Allen, A. (2006). *The Human Givens*. Exeter, UK: Mind South West.
- Zait, A., & Berteau, P. E. (2011). Methods for testing discriminant validity. *Management & Marketing*, 9(2), 217-224. Emerson (ed.) *Contemporary Field Research*. Boston: Little, Brown & Co. t09-126.

About the Authors



Dr. Ceri Wilson

Ceri Wilson is a Research Fellow of Mental Health within the Faculty of Health, Social Care and Education at Anglia Ruskin University. Prior to taking up this role Ceri completed a PhD within the Loughborough University Centre for Research into Eating Disorders. Her research interests centre on the promotion of mental wellbeing and social inclusion of mental health service users.



Dr. Jenny Secker

Jenny Secker is Emeritus Professor of Mental Health at Anglia Ruskin University and the South Essex Partnership University NHS Foundation Trust (SEPT). Her research interests centre on service developments aimed at supporting recovery and social inclusion.