

Reliability of the Leisure Boredom Scale for Use with High School Learners in Cape Town, South Africa

Lisa Wegner

University of the Western Cape, Department of Occupational Therapy

Alan J. Flisher

University of Cape Town, Department of Psychiatry and Mental Health

Martie Muller

Carl Lombard

Medical Research Council, Biostatistics Unit

The reliability of the Leisure Boredom Scale (LBS) for use in South Africa has not been established. In order to address this situation, two studies were undertaken to document the reliability of the LBS in Grades 8 and 11 learners attending high schools in Cape Town, South Africa. In Study 1, a random sample of learners ($N = 117$) attending four independent high schools completed the LBS on two occasions 10 to 14 days apart. In Study 2, a multistage sampling procedure produced a sample of learners ($N = 621$) from 39 government high schools who completed the LBS. In Study 1, test-retest reliability using Cohen's Kappa was found to be moderate to fair for 9 items. The observed agreement ranged from 38.8% to 66.6% for the remaining 7 items. Internal consistency using Cronbach's alpha was 0.76 (Study 1, Time 1), 0.87 (Study 1, Time 2) and 0.76 (Study 2). It was concluded that the LBS has satisfactory reliability for use among high school learners in Cape Town.

KEYWORDS: *Leisure Boredom Scale, adolescents, reliability, South Africa*

Introduction

To improve measurement quality in research, consideration should be given to the reliability of the measurements that are used. Reliability refers to "the degree to which a measurement produces systematic or reproducible variation" (Shrout, 1995, p. 213). Critically examining factors that may affect reliability, such as the variability of the trait being studied and respondents' understanding of questions improves the quality of the measurement. Using several methods to determine reliability provides a more comprehensive picture of the measurement's reliability than using one method alone.

The test-retest method can be used to determine reliability by administering the measurement to the same respondents on two separate occa-

Address correspondence to Lisa Wegner, Department of Occupational Therapy, University of the Western Cape, Private Bag X17, Bellville 7535, South Africa.

The authors gratefully acknowledge the South African Community Epidemiology Network on Drug Use and the Faculty of Health Sciences, University of Cape Town, for funding this study and Janet Evans for coordinating the fieldwork.

sions. A limitation of this method is that the second measurement is often affected by systematic changes in the respondents that have occurred during the interval between measurements. Internal consistency is another method of establishing reliability in questionnaires, by determining the degree to which patterns of responses to items are empirically related. Items relating to the same underlying construct can be considered as replications of one another. Thus where questionnaires include items relating to different constructs, reliability may be underestimated (Shrout, 1995).

Previous studies have shown that the Leisure Boredom Scale (LBS) is a reliable and valid method of determining subjective perceptions of leisure boredom. Iso-Ahola & Weissinger (1990) used the LBS to investigate leisure boredom among college students in America, in three studies ($N = 171$, $N = 164$, $N = 344$). Cronbach's alpha was used to determine reliability and was found to be 0.85, 0.88 and 0.86 for the three studies respectively, indicating high internal consistency. To determine construct validity, theoretically meaningful constructs that included factors relevant to leisure behavior, such as "depth of leisure boredom," "leisure satisfaction," and "frequency of leisure participation," were correlated with the LBS (Iso-Ahola & Weissinger, p. 11). The results showed evidence of significant correlation between the constructs and the LBS as predicted ($p < 0.01$), providing strong support for construct validity of the LBS.

The LBS has not been used previously in South Africa, therefore the reliability of the LBS with South African populations has not been established. Furthermore, no studies have determined the test-retest reliability of the LBS. In order to address this situation, two studies were undertaken to document the test-retest reliability and internal consistency of the LBS with high school learners in Cape Town, South Africa.

Methods and Procedures

Sampling Procedures

The data were provided by a pilot study (Study 1) and a main study (Study 2). These studies focused on the degree of leisure boredom experienced by high school learners, and the extent to which leisure boredom was associated with use of selected substances and demographic variables (Wegner et al., 1999). In Study 1, respondents were randomly drawn from all learners in the Grade 8 and 11 classes at four independent (private) high schools, producing a sample of 117 learners. All of these schools were situated in urban parts of Cape Town. One school was for boys only, one for girls only and two were for boys and girls. As the schools were private, the majority of learners were from middle to upper class families of all ethnic backgrounds.

In Study 2, the study population was defined as all learners in the Grade 8 and 11 classes attending government (non-private) high schools in Cape Town. These schools were situated in urban areas, and admitted boys and girls from all socio-economic and racial/ethnic backgrounds. Postal (zip)

codes were used to stratify schools as there is a high degree of homogeneity in terms of race/ethnicity and social class within each postal code area; thus producing a representative sample. Thirty-nine schools were randomly selected, such that the proportion of selected schools within a stratum was proportional to the total number of learners in that stratum. Within each stratum, the probability of selection of a school was proportional to the number of learners in the school. Forty learners were randomly selected from the combined class lists of two randomly selected classes from each participating grade. An additional five learners were selected to replace absentees. This multistage sampling procedure produced a sample of 2946 learners. This sample was divided into smaller subgroups that completed questionnaires pertaining to different research questions. Randomly selecting seven to eight learners from every group of forty learners produced a sample of 621 learners for Study 2, who completed the Leisure Boredom Scale.

Measurements

The LBS is a self-report questionnaire used to measure "individual differences in perceptions of boredom in leisure" (Iso-Ahola & Crowley, 1991, p. 264). The LBS consists of 16 items (see Table 1) to which subjects respond on a 1 to 5 scale. Total scores can range from 16 to 80 with higher numbers indicating greater boredom. The LBS was adapted to make it more understandable for South African adolescents by changing some of the wording in three items. For example the phrase "*spinning my wheels*" was changed to "*bored and hanging around.*" The LBS was translated from English into the other main languages spoken in Cape Town (Afrikaans and Xhosa). To ensure accuracy in translation, the LBS was then back-translated into English by different people who had Afrikaans or Xhosa as their home language.

Procedures

Permission to carry out the studies was obtained from the Western Cape Education Department and principals of the selected schools. The learners were assured of confidentiality and anonymity, and for this reason only members of the research team (and no teachers) were present during administrations. It was hypothesised that this would promote more accurate completion of the LBS. Learners received explicit written and oral instructions and were able to complete the LBS in their home language.

In order to determine test-retest reliability in Study 1, the LBS was administered to the same learners on two occasions. The interval between administrations was between 10 and 14 days. During the first administration, learners were asked to write the number of their questionnaire on a piece of paper and insert this into an envelope, which they then sealed and signed over the seal. On the outside of the envelope they wrote their name and grade. At the second administration the envelopes were handed back and

learners were asked to copy the number from the paper inside the envelope, onto their new questionnaire.

Analysis

Descriptive statistics were used to obtain demographic profiles of the two samples, and LBS scores (means and standard deviations for items and total scale) for Study 1 (Time 1 and 2) and Study 2. Three measures of agreement were used to calculate test-retest reliability—the observed agreement, Cohen's kappa (κ) (Fleiss, 1981) and the concordance correlation coefficient (Lin, 1989). Observed agreement is the percentage of cases where there is agreement between the two administrations in relation to the total number of administrations. Kappa is the indication of the agreement between the two administrations beyond that which would be expected by chance (Cohen, 1960). If $\kappa = 1.00$ there is perfect agreement. Weighted kappa was reported as the LBS has ordered response categories (1 to 5), which were collapsed into three categories: 1 and 2 (strongly disagree and disagree), 3 (neutral), 4 and 5 (agree and strongly agree). Confidence intervals (95%) were calculated for kappa. If the 95% confidence interval did not include 0, this indicated that kappa was significantly different from 0. The descriptive terms of Landis and Koch (1977) were used to describe the kappa values: $<0 \sim$ poor; $0-0.20 \sim$ slight; $0.21-0.40 \sim$ fair; $0.41-0.60 \sim$ moderate; $0.61-0.80 \sim$ substantial; and $0.81-1.0 \sim$ almost perfect.

Kappa is adversely affected by inconsistencies between the two administrations where the proportions are low ($< 5\%$) (Maclure & Willett, 1987). Thus the observed agreement was used as the preferred measure of agreement where any of the marginal and/or row totals were less than 5%. Bowker's test of symmetry was used to check whether the marginal proportions were homogeneous. Where Bowker's test of symmetry was significant ($p < 0.05$), meaning that on a 95% significance level there was a difference between the proportions at Time 1 and Time 2, the observed agreement was the preferred measure of agreement.

The concordance correlation coefficient was used to calculate test-retest reliability of the LBS as a whole (Lin, 1989). This evaluated the degree to which measurement pairs fell on the 45-degree line, and was constructed from two components—accuracy and precision. The accuracy component measured how far the best-fit line deviated from the 45-degree line whereas the precision component (the Pearson correlation coefficient) measured how far each point (before, after) deviated from the best-fit line. The concordance correlation coefficient was interpreted on the same scale as the Kappa statistic.

Internal consistency of the LBS was determined by calculating Cronbach's alpha coefficient (α) for each item and for the total scale, as well as the item-total correlation for each item. Optimal item-total correlation was considered to be between 0.2 and 0.5 (Jessor & Jessor, 1977). Items with

TABLE 1
Test-retest Reliability of the Leisure Boredom Scale (Study 1, N = 117)

	Missing <i>n</i>	Proportion at Time 1 (%)			Proportion at Time 2 (%)			Bowker's Test of Symmetry (<i>p</i>)	Observed Agreement %	Weighted Kappa (95% CI)
		1-2	3	4-5	1-2	3	4-5			
1. For me, leisure time just drags on and on.	3	72.81	22.81	4.39	71.05	22.81	6.14	0.14	57.2	—
2. During my leisure time, I become highly involved in what I do.	3	12.28	35.96	51.76	6.14	35.96	57.89	0.20	43.7	0.45 (0.30-0.6)
3. Leisure time is boring.	3	78.07	17.54	4.38	81.58	15.79	2.63	0.64	66.6	—
4. If I could leave school now and have enough money, I would have plenty of exciting things to do for the rest of my life.	6	28.83	20.72	50.45	24.33	20.72	54.95	0.74	39.0	0.41 (0.26-0.56)
5. During my leisure time, I feel like I'm just bored and hanging around.	6	20.72	42.34	36.94	16.21	45.95	38.84	0.42	36.8	0.52 (0.38-0.67)
6. In my leisure time, I usually don't like what I'm doing, but I don't know what else to do.	5	63.40	21.43	15.18	68.75	25.89	5.36	0.03*	50.0	—
7. Leisure time gets me aroused and going.	4	11.50	50.44	38.05	15.92	39.82	44.24	0.04*	38.8	—
8. Leisure experiences are an important part of my quality of life.	3	7.02	17.54	75.44	2.63	24.56	72.81	0.11	59.4	—
9. I am excited about leisure time.	3	3.51	28.95	67.54	2.63	33.33	64.03	0.79	53.0	—
10. In my leisure time, I want to do something, but I don't know what to do.	8	34.87	35.78	29.36	42.20	38.53	19.27	0.09	34.2	0.52 (0.39-0.65)

TABLE 1
(Continued)

	Missing <i>n</i>	Proportion at Time 1 (%)			Proportion at Time 2 (%)			Bowker's Test of Symmetry (<i>p</i>)	Observed Agreement %	Weighted Kappa (95% CI)
		1-2	3	4-5	1-2	3	4-5			
11. I waste too much of my leisure time sleeping	8	61.46	24.77	13.76	59.63	25.69	14.68	0.97	45.0	0.50 (0.36-0.64)
12. I like to try new leisure activities that I have never tried before.	6	11.71	23.42	64.86	6.31	32.43	61.26	0.12	48.1	0.38 (0.21-0.54)
13. I am very active during my leisure time.	5	12.50	41.96	45.53	8.04	41.96	50.00	0.58	50.2	0.32 (0.18-0.46)
14. Leisure time activities do not excite me.	5	68.75	26.79	4.46	78.57	17.86	3.57	0.09	59.0	—
15. I do not have many leisure activities available to me.	5	67.86	21.43	10.71	68.75	23.21	8.03	0.57	53.0	0.48 (0.31-0.64)
16. During my leisure time, I almost always have something to do.	11	17.92	22.64	59.44	15.10	19.81	65.09	0.12	46.0	0.47 (0.30-0.63)

Note: 1 = Strongly disagree
 2 = Disagree
 3 = Neutral
 4 = Agree
 5 = Strongly agree

**p* < 0.05

TABLE 2
Internal Consistency of the Leisure Boredom Scale

Item Number	Study 1, Time 1 (N = 107)*			Study 1, Time 2 (N = 107)*			Study 2 (N = 621)		
	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Deleted	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Deleted	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Del.
1.	2.1 (0.9)	.44	.74	2.0 (1.0)	.50	.86	2.5 (1.2)	.15	.76
2.	2.6 (1.0)	.41	.74	2.4 (0.8)	.60	.85	2.8 (1.2)	.41	.74
3.	1.9 (1.0)	.52	.73	1.8 (0.8)	.63	.85	2.3 (1.2)	.43	.74
4.	2.8 (1.2)	.10	.77	2.6 (1.2)	.27	.87	3.4 (1.4)	.04	.78
5.	3.2 (0.9)	.04	.77	3.3 (0.9)	.01	.88	2.5 (1.2)	.52	.73
6.	2.3 (1.0)	.40	.74	2.1 (0.8)	.63	.85	2.5 (1.2)	.45	.74
7.	2.8 (0.9)	.32	.75	2.7 (1.0)	.43	.86	2.8 (1.2)	.53	.73
8.	2.1 (0.9)	.49	.73	2.0 (0.8)	.67	.85	2.5 (1.1)	.42	.74
9.	2.2 (0.9)	.55	.73	2.2 (0.8)	.71	.85	2.5 (1.1)	.44	.74
10.	3.0 (1.0)	.51	.73	2.7 (0.9)	.51	.86	3.1 (1.2)	.35	.75
11.	2.4 (1.0)	.15	.76	2.4 (1.0)	.37	.87	2.3 (1.2)	.28	.75

TABLE 2
(Continued)

Item Number	Study 1, Time 1 (N = 107)*				Study 1, Time 2 (N = 107)*				Study 2 (N = 621)			
	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Deleted	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Deleted	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Deleted	Item Mean (Std. Dev.)	Item-total Correlation	Alpha if Item Deleted
12.	2.4 (0.9)	-0.02	.77	2.3 (0.8)	.48	.86	2.5 (1.2)	.18	.76	2.5 (1.2)	.43	.74
13.	2.6 (0.8)	.53	.73	2.5 (0.8)	.64	.85	2.7 (1.2)	.43	.74	2.5 (1.1)	.33	.75
14.	2.2 (0.8)	.43	.74	2.0 (0.8)	.64	.85	2.5 (1.1)	.32	.75	2.8 (1.2)	.47	.74
15.	2.3 (0.9)	.44	.74	2.2 (0.9)	.64	.86	2.8 (1.2)	.47	.74	2.8 (1.2)	.47	.74
16.	2.5 (1.0)	.44	.74	2.4 (0.9)	.49	.86	2.8 (1.2)	.47	.74	2.8 (1.2)	.47	.74
Scale totals	39.1 (6.9)		.76	37.5 (8.3)		.87	42.6 (9.0)		.76			

Note: *incomplete questionnaires = 10

scores falling out of this range were examined for possible exclusion from the LBS.

Results

Demographic Profile

Descriptive statistics of the demographic profile of the samples showed that in Study 1, 59% of the sample was in Grade 8 and 61.5% were boys. The median age of Grade 8 learners was 13 years, and 16 years in Grade 11. In Study 2, 51% of the sample was in Grade 11 and 60% were girls. The median age of Grade 8 learners was 14 years, and 17 years in Grade 11.

Test-retest Reliability

Using Landis and Koch's descriptive terms (1977), Cohen's Kappa fell into the moderate range for 7 items (range: 0.41 to 0.52) and the fair range for 2 items (range: 0.32 to 0.38) (Table 1). The observed agreement was preferred for 7 items (range: 38.8% to 66.6%). The concordance correlation on the LBS was 0.73 with 95% confidence interval (0.64 to 0.82), which showed substantial reliability (Study 1). The accuracy component was 0.96, which indicated very little deviance of the best-fit line from the 45-degree line. The precision component was 0.76, which indicated some scatter around the best-fit line.

Internal Consistency

The mean and standard deviation for individual items, and internal consistency of the LBS (item-total correlation and alpha) are reported in Table 2. Examination of the items showing suboptimal item-total correlation of either below 0.2 or above 0.5 revealed that in each case the alpha coefficient did not show a relative increase if the item was deleted; therefore removing the item did not improve internal consistency. The alpha coefficient ranged from 0.73 to 0.77 (Study 1, Time 1), 0.84 to 0.88 (Study 1, Time 2) and 0.73 to 0.78 (Study 2).

Discussion

Test-retest Reliability

The highest observed agreement (66.6%) was for the item "*Leisure time is boring.*" This is the most direct, straightforward item in the LBS, and could be expected to show the least variation in responses over the two administrations. Nine of the items showed fair to moderate test-retest reliability as indicated by Cohen's kappa. A limitation of the test-retest design is that the second measurement may be affected by changes (psychological, physical, systematic and social) in the respondent (Shrout, 1995). Boredom may be regarded as an attitude, and as such could be described as a variable trait due to fluctuations that might occur depending on the respondent's mood

state, temporal changes, current situation and reason for participating in activities. Caldwell et al. (1999) found that boredom is influenced by the reason that adolescents engage in leisure activities being that they "want to," "have to" or "have nothing else to do." Therefore, test-retest reliability gives a conservative assessment of reliability as some of the differences between the two administrations could be ascribed to actual changes in leisure boredom that occurred between the two administrations.

The concordance correlation coefficient showed that the reliability of the LBS was substantial. The median difference between the two repeated evaluations of the scale for the Study 1 sample was 1 unit and this was consistent over the whole range of the scale showing good accuracy. The precision was the component that contributed to the scale falling into the substantial reliability category rather than almost perfect.

Internal Consistency

Close examination of the 16 items of the LBS revealed that each item related to the same underlying construct—leisure boredom. Thus the items were considered close replications of one another, excluding the bias, which may result when items do not all relate to the same construct (Shrout, 1995).

Nunnally (1978) suggested a reliability coefficient of 0.70 to be acceptable. Cronbach's alpha coefficients for the LBS in the three samples (0.76, 0.87 and 0.76) fell comfortably above this, and were found to be similar to other studies (0.85, 0.88 and 0.86) (Iso-Ahola & Weissinger, 1990). Although some of the items showed suboptimal item-total correlation, alpha would not improve sufficiently to justify the omission of these items from the scale. Also, items that showed suboptimal correlation with the total in one sample did not show the same result in the other samples. A final reason for not excluding any items was the small sample size ($N = 117$) in Study 1. Therefore the 16 items of the LBS should all remain part of the instrument.

Conclusion

The results of the reliability tests provided support for the reliability of the LBS. These studies have shown that the LBS has satisfactory test-retest reliability and internal consistency when used to document perceptions of leisure boredom among high school learners in Cape Town. These are the first studies in South Africa to have used the LBS as a subjective measurement of leisure boredom in adolescents. Documenting the reliability of the LBS with this population provides a baseline on which other studies can be based. Further studies should be done to establish reliability of the LBS with different populations in South Africa.

References

- Caldwell, L. L., Darling, N., Payne, L. L., & Dowdy, B. (1999). "Why are you bored?" An examination of psychological and social control causes of boredom among adolescents. *Journal of Leisure Research, 31*, 103-121.

- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions (2nd edition)*. New York: Wiley.
- Jessor, R. & Jessor, S. L. (1977). *Problem behaviour and psychosocial development: A longitudinal study of youth*. New York: Academic Press.
- Iso-Ahola, S. E. & Crowley, E. D. (1991). Adolescent substance abuse and leisure boredom. *Journal of Leisure Research*, 23, 260-271.
- Iso-Ahola, S. E. & Weissinger, E. (1990). Perceptions of boredom in leisure: conceptualization, reliability and validity of the Leisure Boredom Scale. *Journal of Leisure Research*, 22, 1-17.
- Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45, 255-268.
- Maclure, M. & Willett, W. C. (1987). Misinterpretation and misuse of the kappa statistic. *American Journal of Epidemiology*, 126, 1161-1169.
- Nunnally (1978). Assessing scale reliability with coefficient Alpha. In L. Hatcher, (1994), *A step-by-step approach to using the SAS System for factor analysis and structural equation modeling* (pp. 130-140). Cary, NC: SAS Institute Inc.
- Shrout, P. E. (1995). Reliability. In M. T. Tsuang, M. Tohen & G. E. Zahner, (Eds.). *Textbook in psychiatric epidemiology* (pp. 213-227). New York: Wiley-Liss, Inc.
- Wegner, L., Flisher, A. J., Watson, R. (1999). A description of a project to investigate leisure boredom and substance use among high school learners in Cape Town. *Urban Health and Development Bulletin*, 2, 78-80.