The Measurement of Gender Role Behaviour in Pre-School Children: a Research Note

Susan Golombok* and John Rust[†]

Abstract—The development and construction of a new psychometric scale, the Pre-School Activities Inventory (PSAI) for the assessment of gender role behaviour in young children, is reported. The test specification, piloting and item analysis are described. Data are also given on reliability and on the various validation procedures used, as well as on age standardization and norming

Keywords Gender role, measurement, pre-school children, gender identity

The aim of the present research was to develop a reliable and valid psychometric questionnaire for the assessment of gender role behaviour in pre-school children. Although many measures have been developed for children aged 5 years or older (for a review see Beere, 1990), few are available for younger boys and girls. Unlike existing tests, the Pre-School Activities Inventory (PSAI) has been designed to discriminate both within and between the sexes so that variation among as well as between boys and girls can be assessed. As pre-school children tend to be unreliable reporters, the inventory is completed by the child's mother or other caretaker. Following the trend in recent measures of gender role behaviour in older children, such as the Children's Personal Attributes Questionnaire (Hall & Halberstadt, 1980) and the Children's Sex Role Inventory (Boldizar, 1991), the PSAI also includes items relating to the child's personality characteristics.

Two sources of information were utilized for the development of the test specification. Firstly the literature on sex-typing in pre-school children was reviewed and ways in which boys and girls have been empirically shown to differ from each other were identified, together with ways in which boys and girls respectively had been described as "girlish' or "boyish". Secondly, a questionnaire was administered to 27 mothers asking them to identify 10 aspects of their son's or daughter's behaviour which they

* The City University, London.
† Goldsmith's College, University of London
Requests for reprints to · Dr Susan Golombok, Clinical and Health Psychology Research Centre, The City University, Northampton Square, London EC1, U.K.

felt was typical for their sex, and 10 ways in which they felt their behaviour was occasionally more characteristic of a child of the opposite sex. This was done to ensure that common belief systems for the attribution of cross sex behaviour within each sex were sampled as well as those which were traditionally thought to discriminate between sexes.

From these responses a test specification was constructed which consisted of three content categories: (i) toys, (ii) activities, and (iii) personality characteristics of the child. From an initial item pool of 57 toys, 47 games and 49 personality characteristics, 30 items per category were identified. Of these 30, 10 were traditionally masculine items, 10 were traditionally feminine and 10 were neutral. In the selection of these items, careful attention was paid to the balancing of desirable with possibly undesirable attributes. This was done with the aim of balancing the final version of the PSAI with respect to the social desirability of the items. The 90 item initial pilot version of the questionnaire was designed to be completed by a parent or caretaker who was in close contact with the child.

The pilot study was carried out on a sample of 32 boys and 43 girls from two day-care nurseries and two play-groups in the London area. The mean age was 43 months (s.d. 9.75 months), and did not differ between boys and girls. The mothers were asked to complete the pilot questionnaire and were given the opportunity to comment on any of the questions. Some items were dropped from further analysis on the basis of these comments, for example, where the activity was very dependent on the weather or the time of year. Comparisons between boys and girls for each item showed that 40 of the 90 items successfully discriminated between the sexes. Some items were omitted because all of the respondents tended to endorse them in the same way.

Generally speaking, gender role items are of two types. Some items may show large overall differences between boys and girls, yet fail to differentiate the degree of sex-typing among children of the same sex. Other items may be considered characteristic of masculinity or of femininity in either boys or girls, yet the sexes do not differ on these items. An objective in the development of the PSAI was the maximization of variation in gender-typing within each sex. This allows the final scale to be used to examine variation within the sexes, as well as differences between the sexes. To achieve this end, we proceeded as follows.

An initial scale was constructed from those items in the "toys" section which discriminated significantly between boys and girls. This initial scale was then correlated with each of the "activities" section items for boys and girls separately. The items from the "activities" section which correlated significantly with the initial scale from the "toys" section for both sexes were used to construct an intermediate "activities" scale. This procedure was repeated in reverse to produce an intermediate "toys" scale. The combined intermediate "toys" and "activities" scales were then correlated with each of the items in the "personality characteristics" section to produce an intermediate "personality characteristics" scale. This provided three intermediate scales which discriminated masculine from feminine gender role behaviour within as well as between sexes. These intermediate scales were then combined and reduced using factor analysis to produce a unidimensional scale. This final scale had 24 items, 12 masculine and 12 feminine. This included 7 toy items, 11 activity items and 6 personality items.

Some subsequent changes to the questionnaire were made to a few of the items to ensure its application across a wide range of English-speaking populations. In particular, the word "aeroplane" (as used in England) was changed to "airplane", and the word "doll's pram" (as used in England) was altered to "doll's carriage". Further analysis on other samples have confirmed that these changes have had no detrimental effect on the factor structure of the questionnaire.

The distribution of scores on the final version is bimodal for the sexes combined, but shows a large amount of within gender variation with significant overlap. The factor analysis of the final scale demonstrated that the items were generally parallel for the combined and separate boy and girl samples in the pilot study. The gender role scale came out clearly in all cases as an unrotated first factor. The second factor was easily identified as the result of acquiescence. The third factor was due to age.

Item analysis statistics were subsequently replicated on several other groups in the UK (data collected by the authors), the USA (data collected by Dr William Freidrich at the Mayo Clinic, Rochester) and the Netherlands (data collected by Dr Cohen-Kettensis at the Academisch Ziekenhuis, Utrecht). The data for the various groups were comparable. All items make a contribution to the discrimination between sexes, and also towards the variation in sex-typing within either one or both sexes.

Scoring

Each item has a score of one to five, representing the response categories "never", "hardly ever", "sometimes", "often" and "very often". The PSAI is scored by first adding the male items, subtracting the female items and then applying a transformation. The total score is such that a higher score indicates more masculine behaviour and a lower score more feminine behaviour. The population mean score for data on 2161 children collected to date is 51.10. Of these, 1166 are boys with a mean score of 61.66 (s.d. = 9.40), and 926 are girls with a mean score of 38.72 (s.d. = 9.66).

Test–retest reliability was examined on a follow-up sample of 15 boys and 18 girls from the four day-care nurseries/play-groups in the pilot study. The mothers made a second rating of their child one year later (within a range of two months).* The pooled test–retest reliability for boys and girls analysed separately was .64 ($N = 33$). Split-half reliability has been calculated for the total sample to date and is .66 for boys ($N = 1260$) and .80 ($N = 1070$) for girls.

The PSAI has been validated on a group of 45 boys and 57 girls attending day care in London in five different centres (mean age 45.7 months, s.d. = 7.51). The inventory was completed by the mother while the day care teachers carried out ratings of the boys on a six point scale ranging from a score of 1 for "much more boyish than average", through "more boyish than average", "slightly more boyish than average", "slightly less boyish than average", "less boyish than average" to a score of 6 for "much less boyish than average". The same procedure was followed for girls, but with "girlish" substituted for "boyish" throughout. For girls, the correlation between the inventory score and the teachers ratings was .48 ($p < .0002$), while for

---

* There were no significant age effects on scores over time.

boys the correlation was .37 ($p < .01$). The partial correlations between the PSAI scores and the validation ratings with age partialed out were .47 ($p < .0003$) for girls and .36 ($p < .02$) for boys, showing that age did not have any confounding influence on these validities.

The questionnaire was standardized across several samples, including (a) the pilot group, (b) the validation group considered above, (c) a sample of 939 boys and 704 girls (mean age 35.79 months, s.d. = 14.88) in the UK obtained through the popular magazine *Practical Parenting,* (d) 178 boys (mean age 44.51, s.d. = 11.33) and 170 girls (mean age 43.34, s.d. = 11.32) from pre-schools in the Netherlands (the questionnaire having been administered in Dutch translation), and (e) 96 boys, mean age 51.41 months (s.d. = 10.35) and 115 girls, mean age 47.54 months (s.d. = 11.23) from pre-schools in Minnesota, U.S.A. The means and standard deviations of the scores for the various groups are given in Table 1.

Table 1. Means and standard deviations of PSAI scores for boys and girls in the standardization groups

Group	N	boys mean	s.d.	N	girls mean	s d
Pilot study	32	60 21	9 56	43	41 41	10 53
Validation study	45	61 06	8.66	57	43 91	11 06
Magazine study	918	60 36	10 16	748	40 31	10 52
The Netherlands	176	58 11	10 09	165	39.11	8 19
Minnesota, USA	94	60 19	9 12	107	37 73	8 41

Sex-typing increases during the pre-school years, so that children before their second birthday are relatively less sex-typed, while by the time they are of five years of age, sex typing is very marked. For boys between the ages of 2 and 6 the correlation of the PSAI with age is .20 ($N = 1061$, $p < .0001$) while for girls it is $- .24$ ($N = 926$, $p < .0001$). As with most developmental scales, it is often necessary to interpret PSAI scores in terms of the average scores for the child's age cohort. In this particular case the standardization procedures necessary to do this must be carried out separately for boys and girls as the distribution of scores for all children is bimodal around the two foci represented by the male and female averages respectively. For boys and girls separately, various models were fitted to the regression of the score on age, and in both cases a straight line was found to remove all significant variance due to age. An age transformation formula that adjusts for age effects is available (Golombok & Rust, in press). This presents the child's PSAI score in relation to the normative group of children of the same age.

Although children's knowledge of gender role stereotypes, as measured by their understanding of the toy and activity preferences of boys and girls in general, have sometimes been considered to reflect children's own preferences, Eisenberg (1983) points out that the two are not equivalent and should not be treated as so. The PSAI is a measure of gender role adoption rather than knowledge about gender role stereotypes. It could be argued that observation of children's behaviour, either in

a laboratory or a natural setting, would provide a more valid measure of gender role adoption than parental report. However, observational assessments only allow short episodes of behaviour to be examined and, particularly when this is carried out in a laboratory, the behaviour observed may not be typical of the child's normal repertoire. In addition, direct observation cannot easily be carried out with large numbers of children. Instead, by using ratings by a parent or other care giver, the PSAI provides an assessment based on observation of the child's behaviour over a long period of time in a variety of settings.

As the respondent does not usually observe the child in all situations, it remains possible that the ratings fail to take account of some behaviours or characteristics while over-emphasizing others. For example, children's behaviour at home and at school may differ so that the parent may not be aware of relevant behaviours which are engaged in only at school. Bias in responding may also result from the parent's own preferences and expectations regarding the child's sex-typed behaviour, or as a result of the child adapting his or her behaviour according to the parent's likes or dislikes. For example, a boy who likes playing with guns may refrain from doing so in the parent's presence if he knows he is likely to be punished. However, the high correlations found between mothers' and teachers' scores suggests that the PSAI is valid across different situations.

Acquiescence effects in most psychometric scales are neutralized by counterbalancing positively with negatively scoring items. In the PSAI the same neutrality is achieved by counterbalancing male with female items. Social desirability has been minimized as much as possible by a balancing of positive and negative degrees of desirability. A further effect which may contaminate scores on the PSAI is the availability of toys of a particular type, or the opportunity for particular types of activities. The toys and activities chosen for the PSAI items are ones which are considered to be widely available in industrial society, even if sometimes this may be only in second-hand or self-made form. Where items are not endorsed simply because a toy was not available or there was no opportunity for a particular activity, the scores will regress towards the population mean, which in most cases will result in girls receiving a less feminine rating and boys a less masculine rating. This effect should be kept in mind by researchers.

References

Beere, C. A (1990). *Gender roles a handbook of tests and measures.* Greenwood Press: New York.

Boldizar, J. (1991). Assessing sex typing and androgyny in children: the Children's Sex Role Inventory. *Developmental Psychology, 27*(3), 505–515.

Eisenberg, N. (1983) Sex-typed toy choices: what do they signify? In M. B. Liss (Ed.), *Social and cognitive skills sex roles and children's play* Academic Press: London

Golombok, S. & Rust, J. (in press). The pre-School Activities Inventory: a standardized assessment of gender role in children. *Psychological Assessment* (in press).

Hall, J. A & Halberstadt, A G (1980) Masculinity and femininity in children. development of the Childrens' Personal Attributes Questionnaire *Developmental Psychology*, **16**(4), 270–280

Appendix. Pre-School Activities Inventory

Name:

Age . . years . . . months

Sex M/F (delete as appropriate)

Instructions

This inventory is about the everyday activities of pre-school children It is in three sections: toy preferences, activities, and characteristics. Each question asks how frequently the child plays with particular toys, engages in particular activities or shows particular characteristics There are five possible answers: (N) never, (HE) hardly ever, (S) sometimes, (O) often or (VO) very often Answer each question by circling the response which best describes the child.

e.g N HE S O VO

Please answer all of the questions If you are unsure about which response best describes the child for any of the questions then please answer according to the response which seems most appropriate.

(Key N = Never, HE = Hardly Ever, S = Sometimes, O = Often, VO = Very Often)

PART 1: TOYS: Please answer these questions according to how often the child played with the following toys during the past month.

1. Guns (or used objects as guns) .	N HE S O VO
2 Jewellery .	N HE S O VO
3 Tool set . . .	N HE S O VO
4 Dolls, doll's clothes or doll's carriage .	N HE S O VO
5 Trains, cars or airplanes	N HE S O VO
6 Swords (or used objects as swords) .	N HE S O VO
7 Tea set	N HE S O VO

PART 2: ACTIVITIES: Please answer these questions according to how often the child engaged in the following activities during the past month.

1 Playing house (e.g. cleaning, cooking)	N HE S O VO
2 Playing with girls … .	N HE S O VO
3. Pretending to be a female character (e.g princess)	N HE S O VO
4. Playing at having a male occupation (e.g. soldier)	N HE S O VO
5. Fighting .	N HE S O VO
6 Pretending to be a family character (e.g. parent) .	N HE S O VO
7. Sports and ball games	N HE S O VO
8. Climbing (e.g. fences, trees, gym equipment) .	N HE S O VO
9 Playing at taking care of babies	N HE S O VO
10 Showing interest in real cars, trains and airplanes .	N HE S O VO
11 Dressing up in girlish clothes	N HE S O VO

PART 3: CHARACTERISTICS: Please answer these questions according to how often the child shows the following characteristics.

1. Likes to explore new surroundings .... N HE S O VO
2. Enjoys rough and tumble play .... N HE S O VO
3. Shows interest in snakes, spiders or insects ... . N HE S O VO
4. Avoids getting dirty . ... N HE S O VO
5 Likes pretty things .... N HE S O VO
6. Avoids taking risks . ... N HE S O VO

NOW PLEASE CHECK THAT YOU HAVE ANSWERED ALL THE QUESTIONS