The Pre-School Activities Inventory: A Standardized Assessment of Gender Role in Children

Susan Golombok and John Rust

The Pre-School Activities Inventory (PSAI) is a new psychometric scale for the assessment of gender role behavior in young children. Its design and test specification are reported, and the piloting and item analysis are described. Evidence of reliability is given, and several validation studies are reported, as are data on age standardization and norming. Some applications of the PSAI are considered.

A number of studies have demonstrated that children show gender-stereotyped toy and activity choices from as early as 18 months of age and that this pattern is well established by 3 years old (Caldera, Huston, & O'Brien, 1989; Fagot, 1974, 1978; Lyttton & Romney, 1991; Maccoby & Jacklin, 1974; O'Brien & Huston, 1985; Perry, White, & Perry, 1984). In spite of the variety of behaviors investigated and methods used, the findings are relatively consistent in showing that boys and girls respectively choose stereotyped masculine and feminine toys and activities.

Although many measures of gender-role behavior have been developed for children ages 5 and older (for a review, see Beere, 1979, 1990), few are available for younger children, and those that do exist have a number of difficulties associated with them. A major problem is their focus on the child's reported preferences for pictured toys, games, or activities, rather than the child's actual involvement in gender-typed play. In addition, little or no information is given about the reliability or validity of these tests. They are also somewhat outdated, which may be important given the changes that have taken place in attitudes toward children's gender-typed play over the past two decades.

It is now more acceptable for boys and girls to engage in a wide range of activities that are not traditionally associated with their gender. The particular toys, games, and activities that children like have also changed considerably over the years. A further drawback of existing tests is their failure to discriminate within the sexes (i.e., to differentiate between masculine and feminine boys and between masculine and feminine girls).

The Fauls-Smith Activity Preference Test (Fauls & Smith, 1956) requires children to choose between a masculine and a feminine activity in three sets of pictures according to the one that they prefer. The test produces a score along a single dimension of masculinity/femininity. No information is given about the reliability or validity of this short test. In the Toy Preference Test (De Lucia, 1963) children are presented with 24 pictures of pairs of toys and are asked to choose the one with which a pictured child of the same sex as the subject would like to play. A single measure of masculinity/femininity is obtained according to the gender appropriateness of the child's toy choices. Although this test is presented by its authors as a measure of gender-role identification, it does not ask for the child's own toy preferences and is thus more accurately described as a measure of knowledge of gender-role stereotypes. Parallel-forms reliability was found to be .57, and test-retest reliability coefficients ranged from .67 to .72 when an opposite-sex experimenter administered the test, and from .13 to .21 with a same-sex experimenter. Construct validity was measured using a list of pairs of games adapted from the Checklist of Games and Play Activities (Rosenberg & Sutton-Smith, 1964). The correlation between these two measures was found to be .64 with an opposite-sex experimenter and -.13 with a same-sex experimenter.

Although a version of this test in which children are asked to select the toys with which they themselves would most like to play has now been developed (Newman & Carney, 1981), no data are available for preschool children.

The most widely used measure is the Sex Role Learning Index (SERLI, Edlebrock & Sugawara, 1978). This has been designed to measure the child's knowledge of gender role stereotypes ("gender-role discrimination"), the child's desire to adhere to gender role stereotypes ("gender-role preference"), and the child's desire to adhere to his or her own ideas of what is gender-role appropriate ("gender-role confirmation"). Boys and girls are presented with 20 pictures of male or female individuals who are respectively engaged in a variety of masculine and feminine activities, and 20 pictures of masculine and feminine objects relating to these activities. The children are asked to sort the objects according to their appropriateness for boys or girls and then to rank the activities in order of preference. In scoring for gender-role preference, gender-role stereotypes are used to define what is sex appropriate, and in scoring for gender-role confirmation, the child's own classification of the objects is used. With a same-sex experimenter, test-retest reliability coefficients ranged between .65 and .69 for gender-role discrimination, between .84 and .90 for gender-role preference, and between .51 and .69 for gender-role confirmation. The reliability...
coefficients were consistently lower with an opposite-sex exper-
iment. The It Scale for Children (Brown, 1956) was admin-
tered together with the SERLI to obtain a measure of construct
validity. Although some relationship was found between the
tests for boys, this was not the case for girls.

The aim of the present research was to develop a reliable and
valid psychometric questionnaire for the assessment of gender-
role behavior in preschool children. Unlike existing tests, it 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. Another deviation from existing measures is
its focus on actual behavior rather than on preferences; the
questionnaire was designed to measure the child's frequency of
play with respect to a variety of toys, games, and activities.

As preschool children tend to be unreliable reporters, the
Pre-School Activities Inventory (PSAI) is completed by the
child's mother or other caretaker. Following the trend in recent
measures of gender-role behavior in older children, such as the
Children's Personal Attributes Questionnaire (Hall & Halber-
stadt, 1980) and the Children's Sex-Role Self Concept Inventory
(Kurdek & Siesky, 1980; Stericker & Kurdek, 1982), the inven-
tory also includes items relating to the child's temperamental
characteristics.

The proper construction of a psychometric questionnaire
of gender-role behavior for preschool children will provide a re-
search tool that can be applied to a wide variety of future inves-
tigations on the development of gender-role behavior in young
boys and girls. Examples are the influences of different types of
family structure and the effects of prevailing gender stereotypes
on gender development. It will allow meaningful compar-
isons between different groups of children and, by providing
normative data, will allow individual children to be assessed.

Method

A questionnaire was completed by 27 mothers who were asked to
identify 10 aspects of their son's or daughter's behavior that they felt
were typical for his or her sex, and 10 ways in which they felt his or her
behavior was occasionally more characteristic of a child of the oppo-
site sex. The results of this questionnaire study, together with a review
of the literature, formed the basis of the test specification, which had
three content categories: (a) Toys, (b) Activities, and (c) Temperamental
Characteristics of the child. The initial item pool contained 153 items
that were reduced to 90 when balanced across the test specification.
Thirty items per category were selected. Of these 30, 10 were tra-
ditionally masculine items, 10 were traditionally feminine, and 10 were neu-
tral. In the selection of these items, careful attention was paid to the
balancing of desirable with possibly undesirable attributes of both mas-
culine and feminine behavior. This 90-item initial version of the ques-
tionnaire, designed to be completed by a parent or caretaker who was
in close contact with the child, was piloted on a sample of 32 boys and
43 girls whose ages ranged between 2 and 5 years, from two day-care
nurseries and two play groups in the London area.

Subsequent item analysis maximized the within-sex variance of the
items while retaining only items that also discriminated between the
sexes. Because of the bimodality of the scale structure, factor analysis
was carried out separately for boys and girls. Following item analysis,
the final scale had 24 items—12 masculine and 12 feminine. This
included 7 Toy items, 11 Activity items, and 6 Character items. The
factor analysis of the final scale demonstrated that the items were gen-
erally parallel for the combined and separate boy and girl samples in
in the pilot study. The mothers made a second rating of their child 1 year later (within a range of 2 months). For the combined group of boys and girls, the test–retest reliability was .84. However, this statistic is spuriously high as the distribution of scores for boys and girls combined is bimodal, and the correlation is thus confounded with the difference between the sexes in mean scores. When the sexes were analyzed separately, the test–retest reliability was .62 for boys and .66 for girls. The pooled test–retest reliability across the sexes was .64 (n = 33).

Split-half reliability was calculated for the total sample to date. For the sexes combined, this was .88 (n = 2,330); for boys .66 (n = 1,260) and for girls .80 (n = 1,070).

### Validity

The PSAI has been validated on a group of 45 boys and 57 girls attending daycare in London in five different centers (mean age = 45.7 months, SD = 7.51). The inventory was completed by the mother while the daycare teachers carried out blind ratings of the boys on a 6-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), whereas for 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.

Of some further interest were the partial correlations between the item scores and the nursery teachers’ ratings with age partialed out. For boys, the items most likely to predict a teacher’s rating of boyishness were *guns* (r = .35, p < .02), *avoiding risks* (r = -.34, p < .02), *swords* (r = -.34, p < .02), and *trains, cars, and airplanes* (r = .30, p < .05). For girls, the items that best predicted a teacher’s rating of girlishness were *guns* (r = -.37, p < .005), *trains, cars, and airplanes* (r = -.33, p < .01), and *pretending to be a female character* (r = .29, p < .05).

### Standardization

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, SD = 14.88) in the UK obtained through the popular magazine *Practical Parenting*; (d) 178 boys (mean age = 44.51, SD = 11.33) and 170 girls (mean age = 43.34, SD = 11.32) from preschools in the Netherlands (the questionnaire having been administered in Dutch translation); and (e) 96 boys, mean age = 51.41 months (SD = 10.35) and 115 girls, mean age = 47.54 months (SD = 11.23) from preschools in Minnesota. These samples provided a breadth of representation across ages, preschools, nationality, and both urban and rural populations. Although, as would be expected in samples of this size, there were some significant differences between the groups, these did not reflect any large-scale variation in the nature of the samples. The means and standard deviations of the scores for the various groups are given in Table 2.

#### Age Related Effects

Sex-typing increases during the preschool years, so that children before their second birthday are relatively less sex-typed, whereas by the time they are of 5 years of age, sex-typing is very marked. Consequently, we would expect the scores on the PSAI to show this pattern. For boys between the ages of 2 and 6 years the correlation of the PSAI with age is .20 (n = 1061, p < .0001), whereas for girls it is -.24 (n = 926, p < .0001). The increasing differentiation is also demonstrated by the increase in the signal detection “d’ prime” function with age, which measures the extent of the overlap between the distribution of scores for boys and girls, with a larger score reflecting less overlap. For children ages 2 to 3 years, this function is 1.86, whereas for children age 5 years, it almost doubles to 3.24. The mean scores and their standard deviations for age bands 24 to 29 months, 30 months to 35 months, 36 to 47 months, 48 to 59 months, and 60 to 71 months for each sex are given in Table 3. A narrower 6-month age band was used for under-3s because of the more rapid change in these earlier months. It can be seen that a score of 40 for girls roughly represents the average femininity score of a girl at just before her third birthday. Younger girls have higher (less feminine) and older girls lower (more feminine) scores. For boys, a score of 60 roughly represents the average masculinity score for a boy as he approaches his third birthday. Younger

---

**Table 2**

<table>
<thead>
<tr>
<th>Groups</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSAI = Pre-School Activities Inventory.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

**Means and Standard Deviations of PSAI Scores for Boys and Girls in the Standardization Groups**

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>n</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot study</td>
<td>32</td>
<td>60.21</td>
<td>9.56</td>
<td>43</td>
<td>41.41</td>
<td>10.53</td>
</tr>
<tr>
<td>Validation study</td>
<td>45</td>
<td>61.06</td>
<td>8.66</td>
<td>57</td>
<td>43.91</td>
<td>11.06</td>
</tr>
<tr>
<td>Magazine study</td>
<td>918</td>
<td>60.36</td>
<td>10.16</td>
<td>748</td>
<td>40.31</td>
<td>10.52</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>176</td>
<td>58.11</td>
<td>10.09</td>
<td>165</td>
<td>39.11</td>
<td>8.19</td>
</tr>
<tr>
<td>Rochester, Minnesota</td>
<td>94</td>
<td>60.19</td>
<td>9.12</td>
<td>107</td>
<td>37.73</td>
<td>8.41</td>
</tr>
</tbody>
</table>

*Note. PSAI = Pre-School Activities Inventory.*

---

### Table 3

**Means and Standard Deviations of PSAI Scores for Boys and Girls Separately Across Age Bands**

<table>
<thead>
<tr>
<th>Age band (in months)</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>18–23</td>
<td>105</td>
<td>59.53</td>
</tr>
<tr>
<td>24–29</td>
<td>179</td>
<td>58.54</td>
</tr>
<tr>
<td>30–35</td>
<td>170</td>
<td>60.07</td>
</tr>
<tr>
<td>36–47</td>
<td>398</td>
<td>62.35</td>
</tr>
<tr>
<td>48–59</td>
<td>229</td>
<td>63.86</td>
</tr>
<tr>
<td>60–71</td>
<td>85</td>
<td>64.87</td>
</tr>
</tbody>
</table>

*Note. PSAI = Pre-School Activities Inventory.*
boys have lower (less masculine) scores, and older boys have higher (more masculine) scores.

**Age Standardization**

As with most developmental scales, it is often necessary to interpret PSAI scores in terms of the average score 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. The linear regression equation is in these circumstances the best transformation for age adjustment. The age adjustment was of the same format as that used for standardization, which was carried out to approximate a mean of 60 and a standard deviation of 10 for boys, and a mean of 40 with a standard deviation of 10 for girls. Thus the standardized age-adjusted scores consist of three components: a scaling element for the initial PSAI score, an adjustment factor for age, and a constant term. They are as follows:

For boys, \( b = a \times 1.075 - \text{Age} \times .177 + 0.88 \)

For girls, \( b = a \times 1.026 + \text{Age} \times .218 - 8.33 \),

where \( b \) is the age standardized score, \( a \) is the initial PSAI score, and Age is the age of the child in months.

To confirm that these transformations do indeed adjust for age effects, analyses of variance were carried out with the age bands as the independent variable. For both boys and girls, the results were not significant, the \( F \) statistic in both cases being less than 1. The Bartlett-Box \( F \) test for homogeneity of variance between these groups was also carried out. This test was not significant for either boys or girls. In spite of this, however, it is worthwhile noting that the standard deviation of the scores does tend to increase slightly with age for both boys and girls. The means of the standardized scores for the age bands are given in Table 4.

**Discussion**

The PSAI has been shown to be a generally reliable and valid psychometric technique for assessing sex-typed behaviors in young children. Although some further data on the test-retest reliability, for boys in particular, would be desirable, the robust values obtained for split-half reliability are encouraging. The results from the validation study are very reassuring, and those for girls are particularly so, given the usual low reliability of single rating scales of the type used for the criterion measure.

The raters and children in the various studies came from a wide variety of educational levels and socioeconomic backgrounds. Each will, of course, have held biases associated with the demographic characteristics of the catchment areas of the various nurseries and preschools involved. But these were themselves reasonably varied. The data from the *Practical Parenting* survey will contain a bias towards the characteristics of this magazine's particular readership, which presumably is well-informed and interested as far as parental issues are concerned. This sample was drawn from all over the UK. Only 2.6% of the respondents were not the mothers of the children being assessed (i.e., they were fathers, stepmothers, etc.). Of the female respondents, 95% were living with a male partner. There were no differences on PSAI scores within this *Practical Parenting* group in terms of geographical area or whether or not the mother worked outside the home.

In the application of the PSAI, some care will need to be taken in the choice of version, as the scale can be used in both age-unadjusted and age-adjusted forms. The former can be treated in most circumstances as a criterion-referenced measure of the absolute level of sex-typed behavior, independently of the child's age or sex. This allows the direct comparison of boys and girls with each other and also gives a measure of the extent to which the child engages in masculine or feminine activities. It does, however, have a bimodal distribution, so that care will need to be taken in the choice of appropriate statistics when data from both sexes is to be included in the same analysis. In some cases, data from boys and girls will need to be analyzed separately.

One particular situation that required such separate analysis was age standardization, where the standard deviation used in calculating standardized scores had to be from a normal distribution. It is for this reason that the age-standardized procedures have to be carried out separately for boys and girls. As with most developmental scales, there are many circumstances in which age-standardized scores are required. It is important to remember, however, that because of the separate age standardizations for boys and girls, it is not possible to compare or relate the age-standardized PSAI scores across sexes.

With the notable exception of researchers in the areas of androgyny (Bem, 1974; Spence, 1984) and gender-identity disorder (Green, 1987; Zucker & Green, 1992), the investigation of gender development has overwhelmingly focused on comparisons between the sexes rather than on individual differences within groups of boys and girls. The PSAI has been designed for both purposes, and it is hoped that it will be useful not only in investigating sex differences and similarities, but also in examining the variation that exists in the development of both male and female gender-role behavior.

**References**


**Table 4**

<table>
<thead>
<tr>
<th>Age band (in months)</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M</td>
</tr>
<tr>
<td>24-29</td>
<td>179</td>
<td>59.18</td>
</tr>
<tr>
<td>30-35</td>
<td>170</td>
<td>59.70</td>
</tr>
<tr>
<td>36-47</td>
<td>398</td>
<td>60.58</td>
</tr>
<tr>
<td>48-59</td>
<td>229</td>
<td>60.14</td>
</tr>
<tr>
<td>60-71</td>
<td>85</td>
<td>59.20</td>
</tr>
</tbody>
</table>

*Note.* PSAI = Pre-School Activities Inventory.
Appendix

Pre-School Activities Inventory

Name: ________________________________
Age: ___________ years ___________ months
Sex: M/F (delete as appropriate)

Instructions

This inventory is about the everyday activities of preschool 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


Received February 28, 1992
Revision received September 9, 1992
Accepted September 15, 1992