The problem-solving performance of young boys and girls was examined in three conditions in which good performance was indicated as being appropriate for boys, for girls, or for both boys and girls. The relations between task scores and five sex-role measures were also examined. As predicted, the boys performed significantly poorer when the task was indicated as one that girls did well. The girls' performance was not affected by the conditions of the study. The girls' problem-solving scores did not differ significantly from the boys' scores. There were some relations between the task scores and the sex-role measures for boys, but not for girls. Associations among the sex-role measures were not high. (Author)
Problem-Solving Performance of Young Boys and Girls
as a Function of Task Appropriateness and Sex Identity*

Dolores Gold and Charlene Berger

Sir George Williams Faculty of Arts
Concordia University

*This research was supported by a DGES-FCAC research grant from the government of Quebec.
The fact that females are the main caretakers of both female
and male children has profound effects on the development of sex identity
in children. The process of achieving masculine sex identity is consequently
both more difficult and more vulnerable than the process of achieving feminine
sex identity. Lynn (1969) has hypothesized that young boys initially identify
with their mothers and later must switch their identification to the male role.
To enable them to do this, male children develop conceptions of the male and
female role and reject the female role. The process of differentiating them-
selves from the female role, then becomes very important to young boys. Further-
more, Lynn has hypothesized that this process of establishing sex identity in
young boys facilitates their problem-solving ability. In contrast to male
children, according to Lynn, female children retain their initial identifica-
tion with their mothers, do not have to differentiate between sex roles as
greatly, do not receive the facilitation in the development of problem-solving
skills, and do not become as good problem solvers.

Some studies of the play activities and preferences of children
lend support to the hypothesis that it is very important for young boys to
differentiate themselves from feminine activities. Rosenberg and Sutton-Smith
compared the game preferences of male and female children across a 30-year
span and found that the roles of boys and girls did not converge. While
the play preferences of the girls had expanded to include many activities that
had previously been considered exclusively masculine, the play preferences of
the boys had contracted.

Stein, Pohly and Mueller (1971) examined the effect of the sex label
of a task on the performance of 6th grade children. They found that the boys
expected to do better, believed it was more important for them to do better,
spent more time working on a task and liked the task better, when it was
defined as a masculine one. The girls expected to do poorer and believed it
was less important for them to do well when the task was defined as masculine, but spent the same amount of time on the task and liked it equally well regardless of whether it was labelled feminine, masculine, or neutral.

Similarly, Montemayor (1974) found that boys performed significantly better with a toy requiring motor skills and liked it better when it was identified as a toy for boys than for girls. Girls likewise performed better and liked the toy more when it was identified as one for girls than for boys.

There is also evidence that boys and girls prefer the toys that are identified as appropriate to their sex (Liebert, McCall and Hanratty, 1971). Indeed, the selection of toys or activities is the basis of some measures of sex identification in young children (Brown, 1956; Fein, Johnson, Kosson, Stork and Wasserman, 1975; Nadelman, 1974).

These studies indicate that boys prefer masculine toys, games and tasks, spend more time on the tasks, and do better on games when they are labelled masculine. The studies provide only some partial support for the hypothesis that it is more important for boys to separate themselves from feminine activities than for girls to separate themselves from masculine activities. This study was designed to test this hypothesis by extending this line of research to study the effects of sex label and sex identification on performance on a problem-solving task in young children. Nursery aged children were used in the study since it is during the ages of three to five that sex identity becomes crystallized.

The first hypothesis was that young boys would differentiate themselves from a task if it was identified as one appropriate for girls. Consequently, it was predicted that boys would perform best on a task if it was defined as appropriate for boys (boys do better condition), less well if the task was defined as appropriate for both boys and girls (both do well
condition), and most poorly if the task was defined as appropriate for girls (girls do better condition). Since girls do not have the same problems in differentiating themselves from the male role, such differentiation should be less important to them and different predictions were made for them. The second prediction was that the girls' performance would not differ on the task in the girls do better or both do well conditions, but that their performance would be poorer in the boys do better condition. Due to the sex-role standards (Kagan, 1964) in the conditions, which define the level of performance on the task as appropriate or inappropriate for the two sexes, both boys and girls were expected to perform most poorly on the task when it was defined as one appropriate for the opposite sex. However, due to the boys' need to differentiate themselves from the feminine role, it was further predicted that the difference between the boys' performance on the task when it was defined as appropriate (boys do better and both do well conditions) and their performance on the task when it was defined as inappropriate for them to do well (girls do better), would be greater than the equivalent difference between the girls' performances in the appropriate and inappropriate situations. The study also allowed a test of Lynn's prediction that males would surpass females in problem-solving skills.

The fifth and final prediction was that the boys' performance on the problem-solving task would be related positively to an assessment of their sex role. No prediction was made concerning the relation between problem-solving task performance and sex role for girls. Since various measures of sex roles have shown poor consistency among themselves, (Mischel, 1970) a variety of measures designed to assess sex-role identity, preference, and concepts in young children were used in the study. Their inter-relationships were also examined to ascertain their communality.
Method

Children from two nursery schools were individually assessed on a battery of tests in four or five sessions over a five month period.

1. The test battery included a measure of intelligence, the Peabody Picture Vocabulary Test (PPVT) (Dunn, 1939), which was used to match the subjects across conditions.

2. The Guilford Creativity Tests for Children; Making Objects (Guilford, 1971) was used as a measure of problem-solving ability, following a suggestion made by Guilford (note 1), who has characterized the tests as a measure of problem-solving involving divergent production (Guilford, 1971). The test, previously adapted for use with four year old children (Reis, 1976), was used as the task in the three conditions. In this test, the child is shown how to use a set of cardboard shapes to make pictures. He or she is then asked to make seven different pictures, for example, an ice-cream cone.

In addition, five measures pertinent to gender identity were used.

1. These included the gender constancy questionnaire, which assesses gender identity, gender stability over time, and gender consistency across situations (Slaby & Frey, 1975).

2. A projective measure of the child's sex identity was used, the Draw a Person Test (DAP) (Goodenough and Harris, 1963).

3. A standard measure of sex-role orientation or preference, the Its Scale for Children, ITSC) (Brown, 1956) was also given to the children. In this test, the children usually are required to make choices for a supposedly neutral stick figure. However, there is some evidence suggesting that the stick figure is perceived as masculine and consequently, the test may not be a valid measure for
girls (McGhee, 1975). In this study the stick figure was replaced by a drawing of an androgynous face of a child. A pilot study had found that female children identified the face as that of a girl and male children, as that of a boy.

4. A sex-preference questionnaire, based upon a method devised by Kagan and Lemkin (1960), asked the children to specify the parental preferences and sex identity of both cartoon figures and themselves. Other items asked the children about the sex of the children with whom they habitually played.

5. Finally, the person from the moon technique based on Hartley (1960) was used to ascertain the relative broadness of the children's sex-role concepts for males and females. This method required that the children indicated what activities they considered appropriate for boys, girls, men, and women. Responses to the moon questions were separately scored by two raters, who scored the responses as to type; the scoring agreement for the two raters was 95.4%. Three scores were selected to assess the children's perceptions of the broadness and similarity of sex roles. These were the number of activities seen as appropriate for males, the number of activities seen as appropriate for females, and the number of activities seen as appropriate for both males and females.

The children were given the tests in the same order, the PPVT and DAP; the Making Objects task; the ITSC; the moon questions and sex-preference questionnaire; and the gender constancy questionnaire in the last session. All children received the PPVT and the Making Objects task, but not all the children received all of the sex-role measures, due to scheduling problems.
Almost equal numbers of girls and boys were randomly assigned to one of three conditions. After giving the regular instructions for the Guilford Making Objects task, the experimenter casually remarked that this was a task on which either boys, or girls, or both generally did well depending on whether the subject was in the boys do better, girls do better, or both do well condition.

Subjects

There were 27 boys and 29 girls in the study. Their ages ranged from 42 to 63 months of age, with a mean age of 52.83, s.d. = 4.86. A 2 x 3 ANOVA (sex x the three conditions) showed no significant differences in age among the six groups of children in the three conditions of the study. Nor were there any significant differences among the groups in the children's intelligence as assessed by the PPVT scores.

The analyses of the children's responses on the measures of sex identity, sex-role preference, and sex-role concepts revealed little variation among the three conditions. There was no significant difference in the proportions of boys and girls in the three conditions attaining the relatively mature third and fourth stages on the gender constancy questionnaire. Nor was there a significant difference between the proportions of girls and boys drawing a same sex figure on the DAP. So few children drew an opposite sex figure, that it was not necessary to test for variations between the conditions of the study. When scored to reflect preference for one's own sex-role, the two measures of sex-role preference did not differentiate among the six groups. When scored in the procedure conventional for the ITSC, that is with higher scores reflecting a masculine orientation, as expected, the boys obtained higher scores than the girls, \( U = 196.5, z = 2.74, p < .01 \) and \( F (1,48) = 82.87, p < .0001 \) on the ITSC and sex-preference questionnaire respectively.
The groups of subjects differed significantly on one of the three scores from the moon question. Boys cited significantly more activities as appropriate for males than did females, $F(1, 47) = 7.90, p < .005$. Girls in the boys do better and both do well conditions gave fewer responses than boys in the boys do better condition, Scheffé $p < .05$. There were no significant differences among the groups in the number of activities considered appropriate for females or the amount of activities seen as appropriate for both males and females.

**Results**

The ANOVA on the Making Objects task scores indicates significant effects for the variables of conditions $F(1, 50) = 8.07, p < .01$ and the interaction between sex of subject and condition, $F(2, 50) = 10.59, p < .005$ (Table 1). The Duncan test reveals that in accordance with the first prediction, boys in the girls do better condition perform significantly poorer than all the other groups (Table 2). The other groups do not differ significantly among themselves. The boys' scores are higher in the boys do better condition than in the both do well condition, but the difference is insignificant. These results then provide some support for the hypothesis that boys differentiate themselves from a task when it is identified as one appropriate for girls.

Counter to the second prediction that the girls would perform best in the girls do better condition, the girls' scores are lowest in that condition, although not significantly so. It had been predicted that the difference in the boys' scores between the boys do better and both do well conditions and the girls do better condition would be greater than the difference between the equivalent scores for girls. This prediction is supported in as much as the boys' scores are significantly higher in the
first two conditions than in the last, while there are no significant differences among the girls' scores.

Lynn's hypothesis that boys have greater problem-solving skills is not supported by this data. The ANOVA on the task scores reveals no significant effects for the variable of sex. Furthermore, with the exclusion of the children in the task inappropriate conditions, the boys and girls do not differ.

The fifth prediction receives some support from the boys' data. The sex-role measures positively associated with the boys' Making Objects task scores are the sex-preference questionnaire and DAP performance. The more the boys prefer the same-sex parent and same-sex companions, the better are their scores on the problem-solving task, \( r (25) = .42, p < .025 \) (One-tailed test). Similarly, the 22 boys who drew a male figure on the DAP scored significantly higher on the problem-solving task, \( r (25) = 2.22, p < .025 \) (One-tailed test), than the five boys who drew a feminine figure.

In addition, the PPVT scores and the sex-preference questionnaire scores are significantly associated, so that the brighter the boys, the more they prefer the male sex, \( r (25) = .40, p < .05 \). Finally, the more intelligent the boys, the better their scores on the problem-solving task, \( r (25) = .62, p < .01 \).

None of the within condition correlation coefficients attain significance; significant coefficients are attained only for the total sample of boys.

There are no significant associations between the girls' scores on the Making Objects task and any of the sex-related measures. This lack of association among the girls' measures obtains both for the total group of girls and for the three groups individually in the conditions of the study.

The intercorrelations among the sex-related measures are shown in Tables 3 and 4. The tables also show the associations between the ages of
the children and their scores on the measures. The Kendall rank-order correlation was used with the ITSC and the gender constancy questionnaire, while the Pearson product-moment correlation was used with the other measures.

The age of the girls correlates significantly only with a score for the moon question. The older the girls, the fewer similar activities they perceive for males and females, $r (25) = -.38$, $p < .05$. The age of the boys correlates significantly with stages on the gender constancy questionnaire, $t = .23$, $z = 1.66$, $p < .05$ (One-tailed test).

Three of the five sex-role measures show some consistency for girls. The gender constancy questionnaire significantly associates with the ITSC. The more mature the stage of gender constancy the girls attain, the more feminine their orientation on the ITSC, $t = -.27$, $z = -2.00$, $p = .05$. The gender constancy questionnaire also correlates significantly with one of the moon question scores. The higher the stage of gender constancy reached by the girls, the greater the number of activities they see as appropriate for boys, $t = .36$, $z = 2.63$, $p < .01$.

The ITSC correlates significantly with the number of similar activities the girls perceive for males and females in the moon question. The more masculine the girls' scores on the ITSC, the greater the number of similar activities they cite for males and females, $t = .29$, $z = 2.12$, $p < .05$.

There are no significant differences between the girls who drew a female figure and those who drew a male figure on the DAP in age or on PPVT score, gender constancy stages, ITSC score, sex-preference questionnaire score, or the three moon question scores.

Four of the five sex-role measures show some consistency for boys. The gender constancy questionnaire shows significant associations with the
moon question scores. The higher the stage of gender constancy reached, the greater the number of activities for males, \( t = .35, z = 2.55, p < .05 \), and the greater the number of similar activities for males and females, \( t = .30, z = 2.21, p < .05 \), are given by the boys.

The boys' responses on the sex-preference questionnaire are significantly associated with their ITSC scores. The more the boys prefer male figures, the more masculine their scores on the ITSC, \( t = .29, z = 2.04, p < .05 \).

There are no significant differences between the boys who drew a male figure and those who drew a female figure on the DAP in age, or on PPVT score, gender constancy stages, ITSC score, sex-preference questionnaire score or the three moon question scores.

**Discussion**

The predictions for the boys' performance were largely supported by the data. The results of this study, therefore, support the hypothesis that boys differentiate themselves from feminine activities, and that it is more important for boys to separate themselves from opposite sex activities than for girls to do so. Such differentiation occurs, even if it means that the boys do poorly on a problem-solving task. However, the boys' performance was poorer only when the task was defined as one which girls do well and by implication, boys therefore, should not do well. This suggests that young boys do not object to performing tasks that are suitable for both sexes, but only those that are labelled exclusively feminine.

The girls' performance, contrary to the study's prediction, was not significantly affected by the experimenter's comments. In fact, their poorest performance was on the task identified as one that girls did well. This finding is contrary to Montamayor's finding (1974) that girls performed better with
a toy when it was defined as being for girls. The children in the Montemayor
study were on the average approximately 27 months older than the children in
this study. This suggests that sex-role standards are not yet as powerful for
the young girls as they have become for older girls.

The influence of sex-role standards has been especially demonstrated
in the problem-solving performance of older subjects, where the usual pattern
is one of better problem-solving by males (Hoffman and Maier, 1961). With
older subjects, a masculine sex-role identification has been associated with
better problem-solving skills for both males and females (Milton, 1957).

The better task scores achieved by the young boys in this study who
preferred masculine company and who drew masculine figures is in agreement
with the results obtained with older subjects and lends support to Lynn's
hypothesis that the development of masculine sex identity facilitates the
development of problem-solving skills in boys. In this study, however, the
results do not support Lynn's hypothesis that due to the nature of their sex-
identification process, boys become better problem-solvers than girls, nor do
the results support the previous finding based on older subjects, that mas- cus- 
line sex identity is associated with better problem-solving by females as well
as males. The equality of the performance of boys and girls at this early age
emphasizes the importance of socialization processes in producing the usual
pattern of poorer performance by older females.

The results of this study also emphasize the limitations of the sex-
role measures. It is noteworthy that contrary to the usual expectations that
sex-role preferences, concepts and identity become more definite as young
children grow older, only two out of a possible 12 associations between the
children's age and test scores were significant. In addition, the pattern of
associations among the sex-role measures indicates that these tests certainly
are not interchangeable. Since the tests attempt to assess the children's
Problem-solving performance

concepts of sex roles, sex-role preference, and gender identity from different approaches very high associations among the tests, therefore, should not necessarily occur. However, what associations exist are generally of such moderate levels that comparisons of studies using these different measures must be done very cautiously.

Finally, it is worthwhile to point out that if the masculine need to reject feminine activities stems from the almost exclusive role of the mother as the main child caretaker, then one way to reduce this masculine need is to introduce more males into child rearing activities. Greater contact with male models, for example, nursery teachers, fathers, baby sitters, etc. should provide the boy with opportunities to learn his sex role more simply and more directly and should help to alleviate the masculine need to emphasize the definition of oneself as apart from girls and women.


Montemayor, R. Children's performance in a game and their attraction to it as a function of sex-typed labels. Child Development, 1974, 45, 152-156.


Reference Note

Table 1

Analysis of Variance for Task Scores

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (A)</td>
<td>1</td>
<td>16.66</td>
<td>1.76</td>
</tr>
<tr>
<td>Conditions (B)</td>
<td>2</td>
<td>76.02</td>
<td>8.07**</td>
</tr>
<tr>
<td>A x B</td>
<td>2</td>
<td>99.75</td>
<td>10.59***</td>
</tr>
<tr>
<td>Error</td>
<td>50</td>
<td>9.42</td>
<td></td>
</tr>
</tbody>
</table>

**p < .01

***p < .005

Table 2

Mean Task Scores for Boys and Girls in the Three Conditions

<table>
<thead>
<tr>
<th></th>
<th>Conditions</th>
<th>Boys</th>
<th></th>
<th>Conditions</th>
<th>Boys</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Boys do Better</td>
<td>Both do</td>
<td>Girls do</td>
<td>Boys do Better</td>
<td>Both do</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Well</td>
<td>Better</td>
<td></td>
<td>Well</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>18.72</td>
<td>17.30</td>
<td>15.90</td>
<td>17.85</td>
<td>17.75</td>
</tr>
<tr>
<td>s.d.</td>
<td></td>
<td>3.39</td>
<td>2.56</td>
<td>3.41</td>
<td>4.05</td>
<td>2.06</td>
</tr>
</tbody>
</table>

*The underlined means do not differ significantly from each other, Duncan test, p < .05.
Table 3

Correlation Coefficients for Associations among the Sex-Role Measures and with Age for Girls

<table>
<thead>
<tr>
<th>Correlation Coefficients</th>
<th>ITSC Scores</th>
<th>Moon Question Scores Male, Female, Similar Activities</th>
<th>Sex-Preference Questionnaire Scores</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender Constancy Stages</td>
<td>t =</td>
<td>-.27*</td>
<td>.36* .12 -.04</td>
<td>.04</td>
</tr>
<tr>
<td>ITSC Scores</td>
<td>t =</td>
<td>-.01 -.03 .29*</td>
<td>-.20</td>
<td>-.22</td>
</tr>
<tr>
<td>Moon Question Scores</td>
<td>r =</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male Activities</td>
<td>r =</td>
<td>.06</td>
<td></td>
<td>-.11</td>
</tr>
<tr>
<td>Female Activities</td>
<td>r =</td>
<td>.17</td>
<td></td>
<td>-.21</td>
</tr>
<tr>
<td>Similar Activities</td>
<td>r =</td>
<td>-.16</td>
<td>.38*</td>
<td></td>
</tr>
<tr>
<td>Sex-Preference Questionnaire Scores</td>
<td>r =</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

n = 27

* p < .05

All p values are based on two-tailed tests.
Table 4
Correlation Coefficients for Associations among the Sex-Role Measures and with Age for boys

<table>
<thead>
<tr>
<th>Correlation Coefficients</th>
<th>ITSC Scores</th>
<th>Moon Question Scores</th>
<th>Sex-Preference Questionnaire Scores</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender Constancy Stages</td>
<td>$t = -$0.1</td>
<td>$r = 0.35^{*}$</td>
<td>$r = 0.30^{*}$</td>
<td>$r = -0.13$</td>
</tr>
<tr>
<td>ITSC Scores</td>
<td>$t = 0.01$</td>
<td>$r = -0.21$</td>
<td>$r = -0.15$</td>
<td>$r = 0.29^{*}$</td>
</tr>
<tr>
<td>Moon Question Scores</td>
<td></td>
<td></td>
<td></td>
<td>$r = 0.05$</td>
</tr>
<tr>
<td>Male Activities</td>
<td>$r = 0.02$</td>
<td></td>
<td></td>
<td>$r = 0.10$</td>
</tr>
<tr>
<td>Female Activities</td>
<td>$r = 0.01$</td>
<td></td>
<td></td>
<td>$r = 0.22$</td>
</tr>
<tr>
<td>Similar Activities</td>
<td>$r = -0.20$</td>
<td></td>
<td></td>
<td>$r = -0.31$</td>
</tr>
<tr>
<td>Sex-Preference Questionnaire Scores</td>
<td>$r = 0.27$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$n = 27$

$*p < .05$

All p values are based on two-tailed tests, except the p value for the coefficient between gender constancy stages and age.