

Evaluation of Math and Science Summer Enrichment Programs in Saudi Arabia

Abdullah M. Aljughaiman
The National Research Center for Giftedness and Creativity
King Faisal University

Abstract

This paper presents the findings from a study aimed at identifying the impact of summer enrichment programs on students' thinking skills, research skills, and personal and social skills, and mastery of academic content. Another aim of the study was to determine the extent to which these programs meet quality standards for summer enrichment programs in Saudi Arabia. A quasi-experimental design was employed, with a pre- and posttest of the experimental group. Only one experimental group was used, because it was difficult to find a control group with the characteristics of the gifted students forming the study sample. Summer enrichment programs constituted the independent variable, and thinking skills; research skills; personal and social skills; mastery of academic content; and the standards known as Gifted Education Programming Standards, developed by Aljughaiman and colleagues (2009), were the dependent variables. The pretest was administered two days before the beginning of the program, during the preliminary phase, and the posttest was applied two days before completion of the program. The evaluation covered three programs, and the study sample consisted of 47 male students and 41 female students from the first year of middle school (average age, 12–13 years). Students were selected on the basis of their participation in the target programs of study (held in the summer of 2009). Thirty lecturers and trainers, and 20 members of the programs' executive teams participated in the administration of the pre- and posttests, which were questionnaires designed to evaluate the summer enrichment programs. The results indicate that the programs had a statistically significant effect on students' thinking skills, research skills, personal and social skills, and mastery of academic content.

Evaluation of Math and Science Summer Enrichment Programs in Saudi Arabia

The process of evaluating programs is considered to be a form of systematic inquiry to produce information. The process helps inform important judgments concerning a certain program, documents both the need for conducting the program and its impact on the participants, and suggests a developmental course for the program. The results of studies in the area of professional development are used to ensure that any institution which seeks to guarantee the continuity of its programs will have indicators that prove the success of those programs (Guskey, 2000).

Therefore, program planners are keenly aware of the high importance of the evaluation process in the success of educational programs, since the aim of that process not only is restricted to presenting information regarding the programs, but also extends to determining the appropriate pathways for their development (Royse, Thyer, & Padgett, 2010). Program evaluation is regarded as a fundamental means of providing program planners and decision makers with the information they need to make appropriate decisions about programs. The process can also be used to reveal to what extent a program's aims have been achieved, the manner in which those aims are achieved, and the obstacles that impede progress toward the aims. Program evaluation can be viewed as a central tool in the development process of any program, helping to determine whether the program should be supported, changed, or halted.

Evaluating gifted programs has long been ignored in the literature, even though it represents one of the basic components in the design of gifted programs (Purcell & Eckert, 2006). This gap in the literature has a negative effect on the practices that can be used in evaluating gifted programs (Avery & Van Tassel-Baska, 1997; Callahan & Reis, 2004).

The evaluation of summer enrichment programs has become an issue of high priority in the Kingdom of Saudi Arabia, related as such evaluation is to the development of those programs. The recognition of the importance of evaluation has resulted in inquiries by many key decision- and policymakers into the benefits of summer enrichment programs, their tangible outcomes, and the potential positive effects on the knowledge, skills, and personality of the gifted students participating in them. In addition, the extent to which these programs meet quality standards of enrichment programs is in need of examination.

Despite the fact that summer enrichment programs in Saudi Arabia are evaluated yearly by both external and internal evaluators (including evaluators from the Ministry of Education and the King Abdul Aziz and His Companions Foundation for the Gifted), there is a paucity of published studies that deal with the effects of these programs on participating students' thinking skills, research skills, personal and social skills, and mastery of academic content.

Given that the aims of summer enrichment programs in Saudi Arabia are various, numerous, and at times overlapping, educators and researchers emphasize the need for translating those aims into standards against which to measure the quality of those programs and judge their reliability and validity for the target subject (Alsorour, 2003; King Abdul-Aziz and His Companions Foundation for the Gifted, 2008; Wahba, 2007). As a result, there have been several serious scientific attempts to translate these aims into standards and criteria for determining the quality of summer enrichment programs for gifted students. The standards and criteria that are found could then help in identifying

the gap between the expected and the actual activities. Among the most prominent of these standards and criteria are those provided by the National Association for Gifted Children (NAGC, 2000).

The NAGC standards for evaluating the quality of gifted programs include standards for the identification of gifted students, standards for program design, standards for designing curriculum and for the selection of teaching strategies, standards for managing and organizing a program, standards for meeting the social and emotional needs of gifted children, standards for professional development, and standards for program evaluation (NAGC, 2000).

In light of the push toward the development of quality standards, Aljughaiman et al. (2009) have developed the Gifted Education Programming Standards. The development of these standards was based on a series of workshops organized in various cities in Saudi Arabia. The workshops were both jointly organized by and featured a variety of specialists in enrichment programs, along with reviews of the NAGC standards and a UNESCO report (UNESCO, 2005).

A review of studies that evaluated gifted programs in the Arab environment (Aljughaiman et al., 2009) or in foreign environments (Delcourt, Cornell, & Goldberg, 2007; Goldring, 1990; Kulik & Kulik, 1987, 1991) reveals that most of these studies have focused mainly on the impact of said programs on traditional variables, such as academic achievement and the attitudes of students and participants toward the programs. What the studies have largely ignored are the impacts on the development of the thinking and personal skills of the students involved. Other studies have focused on evaluating gifted programs in terms of the preparation, planning, difficulties, and constraints their supporters faced during the implementation phase, but do not discuss the quality standards of enrichment programs to verify the quality of services provided. Thus, most of the studies reviewed focused more on ways to improve the programs and the participants' views about them, rather than reaching conclusions about the extent of the programs' success.

In addition, Callahan (2004) asserts that there is a risk of losing whatever evidence has been found of the effects of these programs on the students involved if no attention is given to designing new methods to evaluate those effects. It is noticeable that most programs which regularly evaluate their own activities pay little attention to the results of research studies based on quasi-experimental design, because these programs often are not concerned with attaining outcomes that can be generalized to other programs.

In light of what has been mentioned thus far, the problem raised by the current study can be more clearly defined by answering the following questions:

1. What is the effect of summer enrichment programs on acquiring academic content?
2. What is the effect of summer enrichment programs on gifted students' thinking skills and research skills?
3. What is the effect of summer enrichment programs on the personal and social skills of gifted students?
4. To what extent are the Gifted Education Programming Standards for summer enrichment programs met in the sample of the study?
5. Are there any differences in the responses of the executive team members, lecturers and trainers, and students regarding the Gifted Education Programming Standards?

Significance of the Study

Effective planning for the evaluation of programs is essential because it helps improve and develop the programs, promotes trust in them, and increases the support they receive from both decision makers and the local community they serve. Thus, the process of evaluating programs generally aims to produce information that will assist in making important judgments about the program being evaluated (Callahan, 1986), with the aim of improving it as much as possible (Borland, 1989).

In addition, guaranteeing the continuity of support for summer enrichment programs from decisions makers and the community requires providing evidence of the beneficial effects of these programs on the students' characteristics, their knowledge, and their thinking skills. Hence, the significance of this study stems from the fact that it provides valuable information for all stakeholders on the effects of these programs on the gifted students who participate in them.

Furthermore, the significance of the study inheres in its reliance on the Gifted Education Programming Standards, which allow the researcher to identify the strengths and weaknesses of the different aspects of programs offered to gifted students (i.e., management and organization, teaching and learning activities, guidance and counseling activities, a learning environment that promotes creativity, and evaluation).

Summer Enrichment Programs

For the purposes of this study, summer enrichment programs are those programs held for gifted students under the supervision of the Ministry of Education in the Kingdom of Saudi Arabia in cooperation with a number of centers for the gifted. These programs aim to take care of participating students in an integrated manner with regard to their mental, emotional, social, and physical needs. Furthermore, summer enrichment programs require students to reside on campus for the duration of the program (Ministry of Education, 2004).

Scope of the Study

The current study is confined to the following programs and students:

- The summer enrichment programs administered in three major cities in the Kingdom of Saudi Arabia: Jeddah, Dammam, and Qassim.
- Male and female students in the first year of middle school (average age 12–13 years).

Methodology

The study employed a quasi-experimental design with a pre- and posttest of the experimental group. Only one experimental group was used, because it was difficult to find a control group with the same characteristics as those of the gifted students forming the sample of the study. Summer enrichment programs constituted the independent variable, and academic content, thinking skills, research skills, personal and social skills, and Gifted Education Programming Standards were the dependent variables. The pretest was administered two days before the beginning of the program, during the preliminary phase; the posttest was administered two days before completion of the program.

Participants

The study sample consisted of 47 male students and 41 female students from the first year of middle school (average age 12–13 years). Students were selected on the basis of their participation in the target programs of study (held in the summer of 2009). Thirty lecturers and trainers and 20 members of the programs' executive teams administered the

pre- and posttest questionnaires to the students. The tests were designed to evaluate the summer enrichment programs.

The study sample was chosen according to certain criteria set by the Directorate General for the Gifted in the Ministry of Education in the Kingdom of Saudi Arabia.

Students were eligible to participate if they received

- A score within the top 5% on the abilities test, a test conducted by the Ministry of Education and standardized for the Saudi Arabian environment.
- A general academic achievement score that lies between 90% and 100%.

Study Instruments

To achieve its aims, the study used three tools: academic content tests, a personal and social skills questionnaire, an enrichment program evaluation questionnaire.

Academic Content Tests. Members of the executive team from each program constructed three tests (a pretest, for preliminary evaluation; quizzes, administered during the program as a method of ongoing evaluation; and a summative posttest) to evaluate the students' learning of academic content. Tests were designed according to the following guidelines:

- A table of contents specified what the students were expected to learn in each unit within the enrichment program.
- Each test would include three cognitive levels: understanding and application, problem solving, and critical thinking.
- The questions would be distributed among the three cognitive levels according to specific percentages (25% understanding and application, 50% problem solving, and 25 % critical thinking).

A questionnaire was created to examine the development of students' thinking and research skills. The questionnaire was based on the *Classroom Observation Scales-Revised (COS-R)*, developed by the Center for Gifted Education at the College of William and Mary, in Williamsburg, Virginia, in the United States, with support from the U.S. Department of Education (Van Tassel-Baska et al., 2003). After a review of the literature addressing the characteristics of gifted students, a self-report questionnaire was designed that attempts to measure the extent of the students' thinking and research skills. The students' level of agreement with each item of the questionnaire was established with the use of a Likert five-point scale. Students choose one of the following five alternatives: strongly agree (5 points, the highest on the scale), agree (4 points), agree to some extent (3), disagree (2), and strongly disagree (1 point, the lowest on the scale).

The scale was submitted to five specialists from the National Research Center for Giftedness and Creativity at King Faisal University and King Abdulaziz University. The validity and reliability of the scale were tested with a sample of 160 male and female students from the fifth and sixth elementary grades.

To verify the validity of the scale's factor analysis and to determine the saturation of the items on the scale, sample responses on the two items of thinking and research skills were submitted to a confirmatory factor analysis that used the LISREL program (version 8.8). The results confirmed the factor structure of the scale, establishing both the validity of the factor analysis of the scale and the validity of its items on the scale. (The items were found to be highly saturated, without having to delete any of them.)

Table 1 shows the results of the confirmatory factor analysis.

Table 1. Confirmatory factor analysis of thinking and research skills scale

	Item	Factors		
		Factor Loading	SE of Loading	t-value
1	Generation of new diverse ideas	51.0	10.0	5.25*
2	Creative problem solving	47.0	10.0	4.71*
3	Ability to learn new things quickly and proficiently	45.0	10.0	4.51*
4	Drawing results and reaching generalizations	69.0	09.0	7.47*
5	Proper use of project tools	38.0	10.0	3.76*
6	Gathering information and evidence to verify hypotheses	49.0	10.0	5.03*
7	Making links between ideas and summarizing them	55.0	10.0	5.68*
8	Categorization, organization, and evaluation of information	59.0	10.0	6.22*
9	Precision in recording notes	76.0	09.0	8.51*
10	Good use of diverse sources of knowledge	39.0	10.0	3.87*
Chi square (χ^2)		234.50		

Note: Fit indexes: $0 < RMSEA < 1$, best fit ($RMSEA = 0$); $0 < GFI < 1$, best fit ($GFI = 1$); $0 < AGFI < 1$, best fit ($AGFI = 1$); $0 < NFI < 1$, best fit ($NFI = 1$); * $P \leq 0.01$

It is evident from Table 1 that ($\chi^2 = 234.50$; $P < 0.01$). The values of the relevant indexes were as follows: root-mean-square error of approximation ($RMSEA$) = 0.23; goodness-of-fit index (GFI) = 0.70; adjusted goodness-of-fit index ($AGFI$) = 0.78; and normed fit index (NFI) = 0.81. These values show that the proposed model for the scale of thinking and research skills fits the data and is thereby factor valid.

The reliability coefficient of the scale was calculated with the use of Cronbach's alpha, whose value was 0.83, a statistically significant and acceptable value. The values of the path coefficients ranged from 0.38 to 0.76. These coefficients indicate that the 10 items of the scale together constitute one coherent theme. Figure 1 illustrates the factor structure of the thinking and research skills scale.

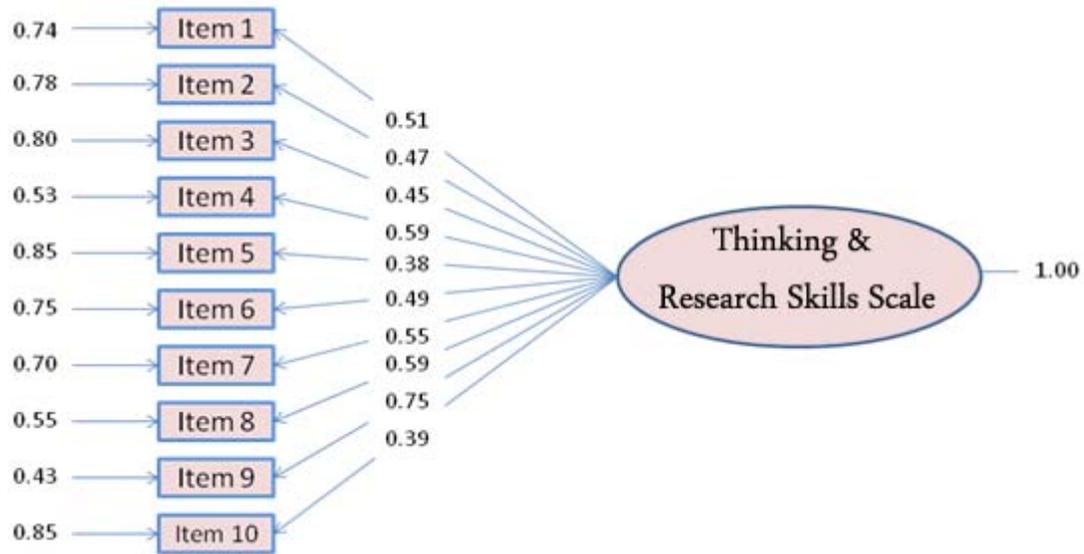


Figure 1. Factorial structure of the thinking and research skills scale *Personal and social skills questionnaire*. A personal and social skills questionnaire was constructed consisting of 12 items derived from a variety of scales found in the literature pertinent to the personal and social characteristics of gifted students. The score for each subject on each item was again determined with the use of a Likert five-point scale: strongly agree (5 points), agree (4), agree to some extent (3), disagree (2), and strongly disagree (1). Because there are overlapping relationships between the personal and social components of the scale, the subject's total score was used in place of scores on individual items to measure the subjects' personal and social skills.

The questionnaire was submitted to five specialists in giftedness, measurement, and evaluation. After the scale was put into its final form, a pilot study was conducted to identify the most appropriate factor structure by using an exploratory factor analysis on the 12 items of the scale. The analysis, in which 160 male and female students participated, showed that the items were saturated on one factor and that two items did not demonstrate sufficient saturation. The two items were amended, and a confirmatory factor analysis using LISREL (version 8.8) was conducted to verify the validity of the factor structure of the scale and the saturation of the items on the scale. The analysis confirmed the factor structure. Table 2 shows the results of the confirmatory factor analysis.

Table 2. Confirmatory factor analysis of the personal and social skills questionnaire

	Items	Factors		
		Factor Loading	SE of Loading	t- value
1	Takes responsibility for assigned tasks	50.0	10.0	5.18*
2	Enjoys collaborative work	81.0	09.0	9.55*
3	Is able to try new things	33.0	10.0	3.24*
4	Completes work even under pressure	0.37	10.0	3.73*

5	Is accurate and disciplined in performance of assigned tasks	77.0	11.0	6.99*
6	Is committed to task	48.0	10.0	4.91*
7	Shows motivation for research and exploration	60.0	09.0	6.44*
8	Defends his or her views in various ways	64.0	10.0	6.63*
9	Expresses personal opinions on different issues without being influenced by the views of others	40.0	10.0	4.03*
10	Researches information independently	45.0	10.0	4.60*
11	Produces new ideas	40.0	10.0	4.06*
12	Strives for perfection	76.0	09.0	8.76*
	Chi square (χ^2)	40.241		

Note: Fit indexes: $0 < RMSEA < 1$, best fit ($RMSEA = 0$); $0 < GFI < 1$, best fit ($GFI = 1$); $0 < AGFI < 1$, best fit ($AGFI = 1$); $0 < NFI < 1$, best fit ($NFI = 1$); * $P \leq 0.01$.

It is evident from Table 2 that ($\chi^2 = 241.40$). The values of the relevant indexes were as follows: root-mean-square error of approximation ($RMSEA$) = 0.18; goodness-of-fit index (GFI) = 0.73; adjusted goodness-of-fit index ($AGFI$) = 0.69; and normed fit index (NFI) = 0.77. These values demonstrate that the proposed model for the scale of personality and social skills fits the data and is thereby factor valid.

The reliability coefficient of the scale was calculated with the use of Cronbach's alpha, whose value was 0.78, a statistically significant and acceptable value. The values of the path coefficients are ranged from 33.0 to 77.0. These coefficients indicate that the 10 items of the scale together constitute one coherent theme. Figure 2 illustrates the factor structure of the personal and social skills scale.

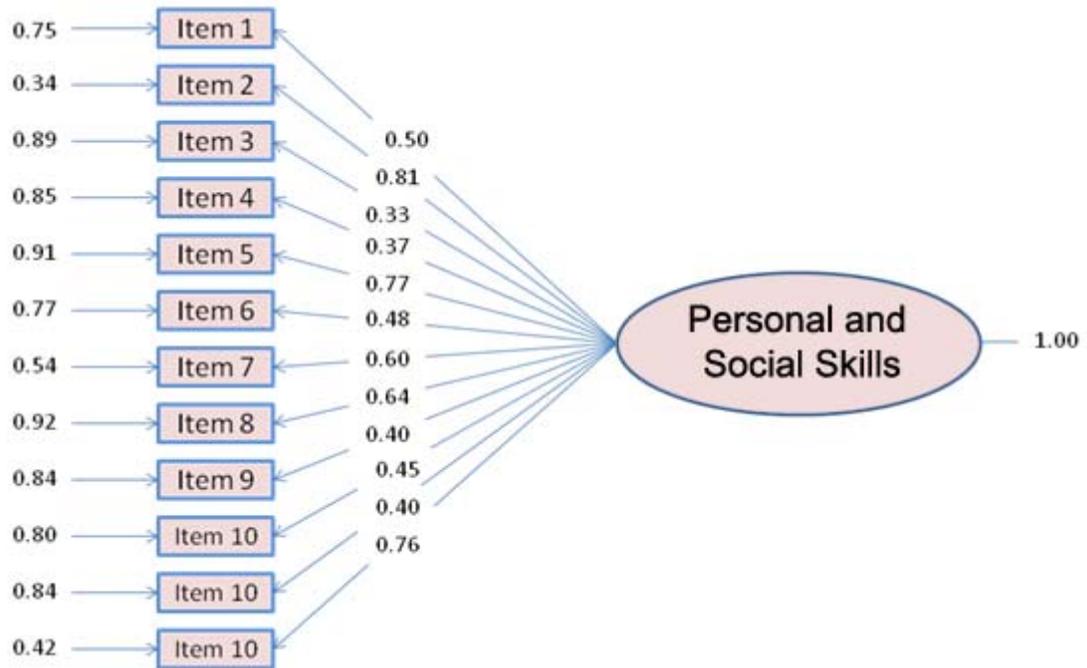


Figure 2. Factorial structure of personal and social skills scale

Enrichment program evaluation questionnaire. The enrichment program evaluation questionnaire was developed in light of the list of criteria set forth by Aljughaiman et al. (2009) to evaluate the quality of enrichment programs. The criteria include the following areas: management and organization, teaching and learning activities, guidance and counseling activities, evaluation, and a learning environment that promotes creativity. These parameters were partially derived from the criteria for the evaluation of enrichment programs adopted by the National Association of Gifted Children (United States), along with efforts to identify the needs and aspirations of the stakeholders of the enrichment programs in the Kingdom of Saudi Arabia. This process of deriving the aforementioned criteria was undertaken through the organization of a series of workshops in various cities with an array of specialists in the field of gifted education.

Three questionnaires were derived from the list of criteria for the evaluation of enrichment programs: a questionnaire for assessing the abilities of the lecturers and trainers of the programs, a questionnaire for the students' evaluation of the programs, and a questionnaire for parents' evaluation of the programs.

A confirmatory factor analysis using LISREL (version 8.8) was conducted to verify the validity of the factor structure of the questionnaires and the saturation of the items in each questionnaire. The analysis confirmed the single-factor structure for each questionnaire and also confirmed that the items of each questionnaire were highly saturated and statistically significant. No items were deleted. This result established the validity of the factor structure, as well as of the items included in each of the three questionnaires:

1. *The questionnaire for evaluating the executive team of the programs (self-evaluation).* Analysis of the results showed that ($\chi^2 = 132.15$; $P < 0.01$). The values of the relevant indexes were as follows: root-mean-square error of approximation (RMSEA) = 0.14; goodness-of-fit index (GFI) = 0.80; adjusted

- goodness-of-fit index (AGFI) = 0.78;), and normed fit index (NFI) = 0.69. These values demonstrate that the questionnaire developed for evaluating the executive team of the program fits the data and is factor valid. The value of Cronbach's alpha was 0.77, a statistically significant and acceptable value.
2. *The questionnaire for evaluating the lecturers and trainers of the programs.* Analysis of the results showed that ($\chi^2 = 97.33$; $P < 0.01$). The values of the relevant indexes were as follows: root-mean-square error of approximation (RMSEA) = 0.17; goodness-of-fit index (GFI) = 0.76; adjusted goodness-of-fit index (AGFI) = 0.82; and normed fit index (NFI) = 0.69. These values demonstrate that the questionnaire developed for evaluating the lecturers and trainers of the programs fits the data and is factor valid. The value of Cronbach's alpha was 0.68, a statistically significant and acceptable value.
 3. *The questionnaire for the student's evaluation of the programs.* Analysis of the results showed that chi square (χ^2) was equal to 112.78, a value that is not statistically significant. The values of the relevant indexes were as follows: root-mean-square error of approximation (RMSEA) = 0.11; goodness-of-fit index (GFI) = 0.81; adjusted goodness-of-fit index (AGFI) = 0.83; and normed fit index (NFI) = 0.76. These values demonstrate that the questionnaire developed for students' evaluation of the programs fits the data and is factor valid. The value of Cronbach's alpha was 0.70, a statistically significant and acceptable value.

Data Analyses

To answer the questions posed in the study, some statistical analyses were performed with the use of the Statistical Package for Social Sciences (SPSS) program (version 16.0). The *t*-test for related samples was used to answer the first, second, and third research questions. Averages and the standard deviation were used to answer the fourth question, and one-way analysis of variance (ANOVA) was used to answer the fifth question. In addition, LISREL (version 8.8) was used to ensure the validity of the factor structure of the thinking skills, research skills, and personal and social skills scales. To ensure the strength of the experimental effect size of the independent variable (the program), effect size was calculated with the equation:

$$\text{Effect Size} = \frac{t \times \sqrt{2(1-R)}}{N} \quad (\text{Aron \& Aron, 1995}).$$

where *R* denotes the correlation coefficient between the two applications of the pre- and posttest, *N* designates the size of the study sample, and *t* indicates the value of *T*.

Results

The effect of the program on the acquisition of academic content. In order to answer the first research question and verify the program's effect on the acquisition of academic content, a *t*-test for related samples was used to measure the differences between the average scores of students in the two applications of the pre- and posttest for academic content. Table 3 displays the results of the *t*-test.

Table 3. Results of the *t*-test on the differences between the mean scores of the two applications (pre- and posttest) on the acquisition of the academic content test ($n = 88$)

	<i>M</i>	SD	df	<i>t</i> -value
Pre test	2.94	1.08	87	24.83*
Post test	6.76	1.08		

* $P \leq 0.001$.

It is evident from the table that the value of t is 83.24, a statistically significant value at the level of significance $P \leq 0.001$. It follows that there are differences in the acquisition of the academic content test between the average of the pretest ($M = 2.94$) and the average of the post test ($M = 6.76$) and these differences are in favor of the posttest.

These results indicate the success of the summer enrichment programs in bringing about positive effects on the academic achievement of the students who participated in them.

The effect of the program on the students' thinking and research skills. In order to answer the second research question and verify the program's effect on the thinking and research skills of participating students, a t -test for related samples was employed to measure the differences between the averages of students' scores in the two applications (pre- and posttest) of the thinking and research skills scale. Table 4 shows the results of this t -test.

Table 4. Results of t -test of the differences between the mean scores of the two applications (pre- and posttest) of thinking and research skills scale.

	<i>M</i>	SD	df	<i>t</i> -value
Pretest	27.11	6.90	87	18.46*
Posttest	38.84	5.61		

* $P \leq 0.01$.

According to Table 4, the value of t is 18.46, a statistically significant value at the level of significance $P \leq 0.01$. It follows that there are differences in the thinking and research skills scale between the average of the pretest ($M = 27.11$) and the average of the posttest ($M = 38.84$) and these differences are in favor of the posttest.

These results indicate the success of the summer enrichment programs in bringing about positive effects on the thinking and research skills of the students who participated in them.

The effect of the programs on students' personal and social skills. In an attempt to answer the third research question and verify the program's effect on the personal and social skills of participating students, a t -test for related samples was utilized to measure

the differences between the averages of students' scores in the two applications (pre- and posttest) of the personal and social skills scale. Table 5 shows the results of the *t*-test.

Table 5. Results of *t*-test on the differences between the mean scores of the two applications (pre- and posttest) of the personal and social skills scale.

	<i>M</i>	SD	df	<i>t</i> -value
Pretest	26.44	4.37	87	26.05*
Posttest	40.91	6.57		

$P \leq 0.001$

From Table 5, the value of *t* is seen to be 26.05, a statistically significant value at the level of significance $P \leq 0.001$. It follows that there are differences in the personal and social skills scale between the average of the pretest ($M = 26.44$) and the average of the posttest ($M = 40.91$) and these differences are in favor of the posttest.

These results indicate the success of the summer enrichment programs in bringing about positive effects on the personal and social skills of the students who participated in them.

The extent to which enrichment programs meet program evaluation standards. To answer the fourth research question, about the extent to which the quality standards of the summer enrichment programs are met, the average and the standard deviation of the responses of the executive team members, lecturers and trainers, and students on the questionnaires were calculated. Table 6 shows the results of these calculations.

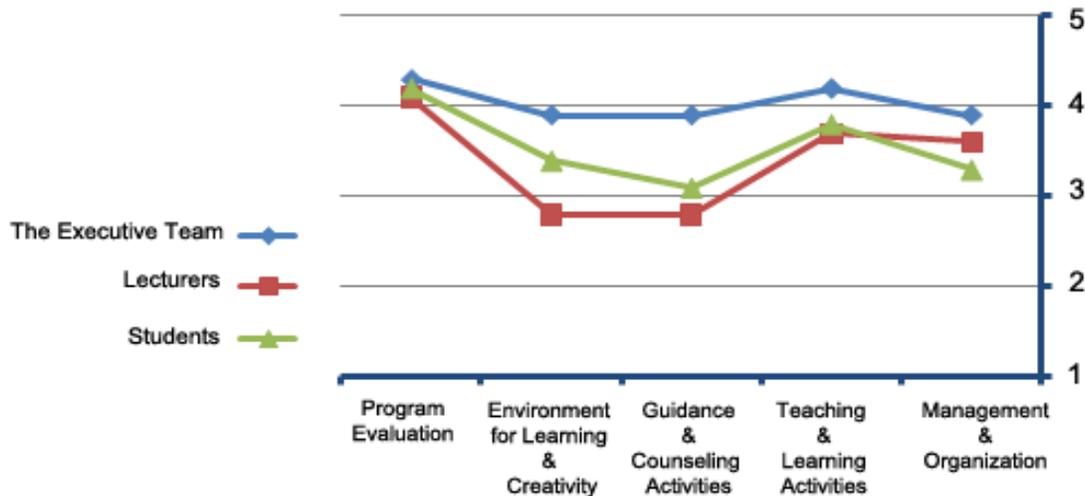
Table 6. Averages and standard deviations of the responses of the executive team members, lecturers and trainers, and students on the questionnaires evaluating summer enrichment programs.

Areas	Executive Team		Lecturers and Trainers		Students	
	<i>M</i>	SD	<i>M</i>	SD	<i>M</i>	SD
Management and organization	3.9	0.64	3.6	0.94	3.3	0.92
Teaching and learning activities	4.2	0.68	3.7	0.64	3.8	0.70
Guidance and counseling activities	3.9	0.88	2.8	0.86	3.1	0.77
Learning environment that promotes creativity	3.9	0.75	2.8	0.79	3.4	0.90
Evaluation	4.3	0.70	4.1	0.68	4.2	0.64

Table 6 shows that the average responses of the executive team, lecturers and trainers, and students in the area of management and organization meet the standards for

this area at a medium level of confidence (average of executive team = 3.9, average of lecturers and trainers = 3.6, average of students = 3.3) whereas in the area of teaching and learning activities, the responses of the executive team members meet the standards at an above-average level of confidence (average of executive team = 4.2) and the responses of the lecturers and trainers and of the students meet the standards to a medium level of confidence (average of lecturers and trainers = 3.3, average of students = 3.8). With regard to the area of guidance and counseling activities, the executive team's and students' responses meet the standards of at a medium level of confidence (average of executive team = 3.9, average of students = 3.1); by contrast, the lecturers' and trainers' responses exhibit a less-than-average level of confidence (average of lecturers and trainers = 2.8). In the area of a learning environment that promotes creativity, the average responses of the executive team and the students meet the standards at a medium level of confidence (average of executive team = 3.9, average of students = 3.4) whereas the lecturers' and trainers' responses exhibit a less-than-average level of confidence (average of the lecturers and trainers = 2.8). Finally, in the field of program evaluation, the responses of the executive team, lecturers and trainers, and students meet the standards at an above-average level of confidence (average of executive team = 4.3, average of lecturers and trainers = 4.1, average of students = 4.2).

Figure 3 shows the averages of the responses of the executive team members, lecturers and trainers, and students on the questionnaire aimed at evaluating summer enrichment programs.



The differences in the responses of the executive team members, lecturers and trainers, and students on the Gifted Education Programming Standards. In order to examine the fifth research question, about whether there are statistically significant differences in the responses of executive team members, lecturers and trainers, and students about the extent to which the Gifted Education Programming Standards are met, the levels of confidence calculated were converted to adjusted norm scores (i.e., *t*-scores). Adjustment was necessary because of the differences in the number of items on each questionnaire and the differences in the total value among the three questionnaires that were employed. A one-way ANOVA was conducted in which the SPSS program was used to study the

differences in the responses of the executive team members, lecturers and trainers, and students on the standards of the summer enrichment programs. Table 7 shows the results of the ANOVA.

Table 7. Results of analysis of variance of the responses of the executive team members, lecturers and trainers, and students on the questionnaire about the Gifted Education Programming Standards.

Areas	Sum of Squares	df	Mean square	<i>F</i>
Management and organization	1840.01	2	920.01	7.440**
Teaching and learning activities	28.41	2	14.20	0.141
Guidance and counseling activities	5.52	2	2.76	0.027
Learning environment that promotes creativity	0.71	2	0.36	0.004
Evaluation	0.69	2	0.34	0.003

$P \leq 0.01$.

Table 7 shows statistically significant differences in the responses of the executive team members, lecturers and trainers, and students on the dimension of management and organization ($F = 7.440$, $P \leq 0.01$). To identify the source of these differences, Sheffé's test for post hoc analysis was employed. Comparisons showed that the differences in the responses of the executive team members and of the lecturers and trainers were in favor of the executive team members. There were no statistically significant differences among the executive team members, lecturers and trainers, and students on the dimensions of teaching and learning activities ($F = 0.141$), guidance and counseling activities ($F = 0.027$), learning environment that promotes creativity ($F = 0.004$), and evaluation ($F = 0.003$).

The results of the fifth research question show convergence among the responses of the executive team members, lecturers and trainers, and students as far as the extent to which the Gifted Education Programming Standards were met in the measured dimensions (teaching and learning activities, guidance and counseling activities, a learning environment that promotes creativity, and evaluation). The results also indicate that there are differences between the executive team members, on the one hand, and the lecturers and trainers, on the other, in the dimension of management and organization.

Effect Size

The effect size was measured to ensure the effect of the enrichment program (independent variable) on the students' learning of academic content, thinking and research skills, and personal and social skills (dependent variables). Table 8 shows the values of the effect size on the latter three variables.

Table 8. Effect size (η^2) of the program on the significant variables.

Variable	η^2
Academic content	0.60

Thinking and research skills	0.42
Personal and social skills	0.60

Table 8 indicates that the values of the effect size (η^2) of the program on the variables examined ranged between 0.42 and 0.60. The value of the effect size of the program on the acquisition of the academic content was 0.60, on thinking and research skills was 0.42, and on personal and social skills 0.60.

The effect of the programs on acquisition of academic content. The results of the study presented here indicate that the enrichment programs had a statistically significant effect on the acquisition of academic content, at the level of $p \leq 0.01$ and with a total effect size of 0.60. This finding might be due to a high concentration on the academic area on the part of the instructors. The instruction strategies used in the summer programs capitalized on students' involvement in taking the responsibility for their own learning. Most of the learning experiences provided in the programs were based on a problem-solving approach to learning concepts.

This result should be taken with caution for four main reasons. First, a review of students' profiles indicates that the result is not unexpected, because the participating students are high academic achievers. Second, the pre- and posttests measuring academic content were developed and scored by each of the program's teachers. Hence, it is possible that teachers gave more attention to the topics covered in the tests, a factor that was then reflected in students' achievement scores. Third, the pretest depended more on the competencies of the program than on those of the students, so students achieved relatively low scores because they had few of the program's competencies at the outset. Thus, any acquisition of academic content by the students during the program would have an impact on the posttest. Fourth, the pre- and posttests have a narrow range of degrees (total score = 10), meaning that any improvement the students make between taking them can lead to significant differences. The upshot is that it is hard to draw a clear image about the biases of the teachers. The positive impact on the acquisition of academic content accords with the results of Delcourt, Loyd, Cornell, and Goldberg (1994), who examined the effects of a number of gifted programs (special schools, special classes, and pullout programs) on academic achievement. Their results demonstrated statistically significant differences between the mean scores of students who participated in the different enrichment programs and the mean scores of students who did not participate (in favor of the participating students).

The effect of the programs on thinking and research skills. The results of the study indicate that the enrichment programs had a statistically significant effect on the development of thinking and research skills, at the level of significance $p \leq 0.01$ and with a total effect size of 0.42. This finding can be interpreted in light of the gifted students' responsiveness to higher order, open-ended questions that encourage discovery and exploration, as well as motivate the students to think about topics in a more detailed, abstract manner. Furthermore, the result confirms the crucial role of the summer enrichment programs in the development of gifted students' thinking skills through the provision of specialized activities and exercises that strengthen and broaden the students' capabilities and skills. Another important factor that might have contributed to this result

is that most of the programs participating in the study focused on providing choices based on students' interests and encouraged students to be self-learners.

The effect of the programs on personal and social skills. The results of the study show that enrichment programs have a statistically significant effect on the development of personal and social skills, of at the level of significance $p \leq 0.01$ and with a total effect size of 0.60. One plausible interpretation of this finding is that the effects of enrichment programs on the personal and social skills of the participating students is a good indicator of the success of such programs in developing not only students' knowledge and thinking skills (as has been confirmed in the answers to the first and second research questions), but also the important personal and social dimensions that help gifted students realize their full potential. This finding corresponds to what is commonly asserted in models of giftedness: that appropriate education is a key factor in the development of gifted students and that their exposure to enrichment programs is essential to the maturation of their personal and social characteristics, such as self-confidence, independence, perseverance, and teamwork. The finding is consistent with those of Aljughaiman et al. (2009), which indicated a positive impact of school enrichment programs on personal and social skills.

To what extent are the quality criteria of enrichment programs achieved?

The results obtained from the responses of executive team members, lecturers and trainers, and students showed that the standards of management and organization were met to a satisfactory degree by all of those groups. The standards for teaching and learning activities were met to an above-average degree by the executive team members and to a medium degree by lecturers and trainers and by students. The standards of guidance and counseling activities and of providing a learning environment that promotes creativity were met to a medium degree by the executive team members and by the students and were not met (i.e., the degree of achievement was less than average) by the lecturers and trainers. Finally, the standards of program evaluation were met to an above-average degree by the executive team members, lecturers and trainers, and students.

The differences of opinion among the executive team members, lecturers and trainers, and students with regard to the aforesaid four criteria can be attributed to many factors. First, the Gifted Education Programming Standards cover a number of areas that are not covered by either the pre- and posttests or the questionnaires. In addition, when the researchers talk directly with stakeholders, the stakeholders tend to give more detailed explanations about what their hopes for and criticisms of the programs are; neither of these elements are reflected on the test. Second, the teachers, supervisors, administrators, and students typically interact with the Gifted Education Programming Standards throughout the duration of the program, and the researchers participate in an external review of the program utilizing the standards two to three times over the course of the program. As a result, this process yields more robust data about the program than is reflected in the tests and questionnaires.

Differences in the attitudes of the executive team members, lecturers, and students in the quality criteria of enrichment programs

The results of the study showed that there are statistically significant differences among the executive team members, lecturers and trainers, and students in the dimension of management and organization ($P \leq 0.01$) in favor of the executive team. This result can be explained by a lack of effective communication between the program's administrators and teachers. Teachers tend to focus on their teaching and their relationships with

students, while thinking that they do not receive adequate administrative support, whereas executive team members tend to view their own performance more positively. In contrast, the results found no statistically significant differences among the executive team members, lecturers and trainers, and students in the following dimensions: teaching and learning activities, guidance and counseling activities, learning environment that promotes creativity, and evaluation.

Counseling needs of students

The results indicate that the counseling needs of the students are not being met, and this shortcoming is not measured on the pre- or posttest, or the questionnaires. However, this result was observed by the researchers during visits to the programs and in interviews with the students and their teachers. The most probable explanation is that the program administration did not include any specialized counseling staff.

Recommendations

The results of the study point to some essential recommendations for officials, educators, decision makers, and those responsible for the planning and implementation of gifted programs. Among the most important of these recommendations are the following: (1) Teachers should be trained to integrate thinking skills, research skills, personality, and social skills into the curriculum and (2) more effective ways of evaluating the students' acquisition of the curriculum content need to be developed. It is essential to focus on quality standards in all aspects of enrichment programs for the gifted, from planning, organization, and the formulation of teaching and learning activities, to guidance and counseling activities, providing a learning environment that promotes creativity, and program implementation and evaluation. Further recommendations include providing training for the executive team members, lecturers and trainers, and students on methods of self-evaluation; encouraging parental involvement in all stages of planning, implementing, and evaluating enrichment programs; and employing professional counselors to respond to the special needs of students participating in the programs.

References

- Aljughaiman, Abdullah, Maajeeny, Osamah, Alfheid, Abdulrahman, Abu Nasser, Fathi, Abu Oof, Talat, Ayoub, Alaa, and Alhusain, Ibraheem. (2009). *Evaluation of Gifted Education Program in Public Schools in The Kingdom of Saudi Arabia*. Unpublished study housed at the Ministry of Education, Kingdom of Saudi Arabia, Riyadh.
- Alsorour, Nadia Hayel. (2003). *An Approach to Education of the Gifted and Talented*. Beirut: Dar Al Fikr for Printing and Publication.
- Aron, A., & Aron, E. N. (1995). Three suggestions for increased emphasis in the study of close relationships. *Journal of Social and Personal Relationships*, 12, 559–562.
- Avery, L., & Van Tassel-Baska, J. (1997). Making evaluation work: One school district's experience. *Gifted Child Quarterly*, 41(4), 124–132.
- Borland, J. H. (1989). *Planning and implementing programs for the gifted*. New York: Teachers College Press.
- Callahan, C.M. (1986). Asking the right questions: The central issue in evaluating programs for the gifted and talented. *Gifted Child Quarterly*, 30(1), 38–42.
- Callahan, C.M. (1992). *Evaluation lectures*. Helena, MT: Project EDGE, Montana Office of Public Instruction.

- Callahan, C. M. (2004). Asking the right questions: The central issue in evaluating programs for the gifted and talented. In C. M. Callahan (Ed.), *Program evaluation in gifted education*, pp. 1–12. Thousand Oaks, CA: Corwin.
- Davis, G., and Rimm, S. (2010). *Education of the gifted and talented*. Needham Heights, MA: Allyn and Bacon.
- Delcourt, M, Cornell, D, & Goldberg, M. (2007). Cognitive and affective learning outcomes of gifted elementary school students. *Gifted Child Quarterly*, 51, 359–381.
- Delcourt, M. A. B., Loyd, B. H., Cornell, D. G., & Goldberg, M. D. (1994). Evaluation of the effects of programming arrangements on student learning outcomes (Research Monograph 94108). Storrs, CT: The National Research Center on the Gifted and Talented, University of Connecticut.
- Goldring, E. B. (1990). Assessing the status of information on classroom organizational frameworks for gifted students. *Journal of Educational Research*, 83, 313–326.
- Guskey, T. R. (2000). *Evaluating Professional Development*. Thousand Oaks, CA : Corwin Press.
- King Abdul-Aziz and His Companions Foundation for the Gifted. (2008). "*Mawhiba*" *Summer Gifted Education Programs' Guide*. Unpublished document housed at the King Abdul-Aziz and His Companions Foundation for the Gifted, Riyadh.
- Kulik, J. A., & Kulik, C. C. (1987). Effects of ability grouping on student achievement. *Equity & Excellence*, 23, 22–30.
- Kulik, J. A., & Kulik, C. C. (1991). Ability grouping and gifted students. In N. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (pp. 178–196). Needham Heights, MA: Allyn & Bacon.
- National Association for Gifted Children (NAGC, 2000). *Pre-K–Grade 12 Gifted Program Standards*. Washington, DC : National Association for Gifted Children.
- Ministry of Education (MOE). (2004). *Evaluation school practices of the school enrichment programs*. Unpublished report. The Administration of Gifted Education, Ministry of Education, Saudi Arabia.
- Purcell, J. H., & Eckert, R. D. (2006). *Designing services and programs for high-ability learners: A guidebook for gifted education*. Thousand Oaks, CA: A Joint Publication of Corwin Press and the National Association for Gifted Children.
- Royse, D., Thyer, B., & Padgett, D. (2010). *Program Evaluation: An Introduction*. Belmont, CA : Wadsworth .
- UNESCO. (2005). *Education for All – The Quality Imperative*. Retrieved June, 8, 2010: <http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/efareport/reports/2005-quality/>
- Van Tassel-Baska, J., Avery, L., Struck, J., Feng, A., Bracken, B., Drummond, D., & Stambaugh, T.(2003). *The William and Mary Classroom Observation Scales—Revised*. Center for Gifted Education, The College of William and Mary, School of Education, Williamsburg, Virginia, United States.
- VanTassel-Baska, J., & Brown, E. (2009). An analysis of gifted education curriculum models. In F. Karnes & S. Bean (Eds.), *Methods and materials for teaching the gifted* (pp.75–108). Texas: Prufrock Press.

Wahba, Mohammed. (2007). *The Gifted and Talented: Identification Criteria and Their Education*. Alexandria: Dar Alwafa' for Printing and Publication World.