

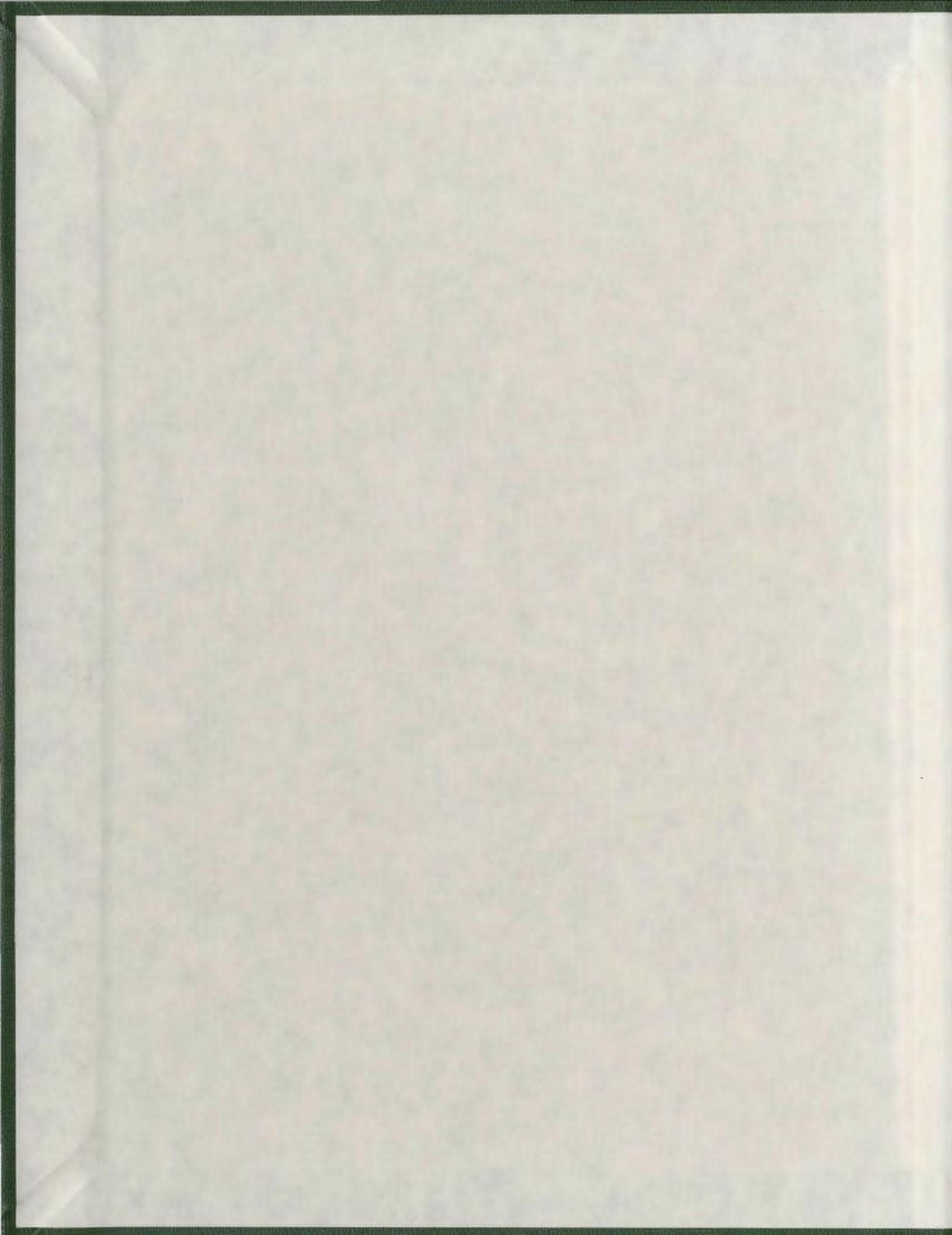
**THE EVALUATION OF TOYS:
Selection Criteria And Expert Consensus**

CENTRE FOR NEWFOUNDLAND STUDIES

**TOTAL OF 10 PAGES ONLY
MAY BE XEROXED**

(Without Author's Permission)

JOYCE ELAINE ANNE (MURPHY) SMITH



CC123



THE EVALUATION OF TOYS: SELECTION CRITERIA
AND EXPERT CONSENSUS

by

© Joyce Elaine Anne (Murphy) Smith

A Thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Education

Department of Educational Psychology
Memorial University of Newfoundland

Spring 1983

St. John's

Newfoundland

ABSTRACT

This research examined, in an educational and developmental framework, expert consensus on toy evaluation criteria. The work was aimed at discovering a) whether or not experts could concur about the "overall goodness" of toys, and b) whether particular characteristics of toys, namely their safety, durability, play value, educational-developmental value and interpersonal-social value were similarly viewed.

The study asked five main questions: 1) Do raters perceive significant differences among the toys? 2) Do Academics and Practitioners rate toys in similar ways? 3) Do individual experts agree in their ratings? 4) Is there agreement among the two groups of raters on the ratings of specific toys? and 5) Are there significant correlations between a) the overall goodness score and the summated variables score and b) the educational developmental value and play value scores?

To answer these questions, six academics and six practitioners were asked to rate 62 pictorially presented toys on the six criteria listed above. Ratings were done on a 7-point Likert-type scale. From these ratings, six criterion variable scores plus one summative score were obtained.

While it was found that experts, as a group, concur in their rating of a toy, it was also found that the rating of

a single expert may not necessarily be consistent with that of other experts. Post hoc analysis determined that for most variables, four expert's opinions are needed for a reliable rating.

Results also suggest that the raters appeared to rate according to two possible factors, one reflecting the toy's construction and the other its use. Other significant trends are also noted.

The findings of the study are discussed relative to Schmidt's (1973) views of development and education. Some limitations of the study and suggestions for future research are offered.

It is suggested that the Toy Evaluation Instrument used in this study may be useful in further attempts to explore the potential of toys as educational aids. It is also suggested that much of the value of the work may be in its practical nature and in its implications for parent training.

ACKNOWLEDGEMENTS

This study was able to be initially pursued and finally completed because of a number of important people in my life, some of whom are mentioned here.

To my father, Jack Murphy and my deceased mother, Mary (Beecher) Murphy, my first and foremost educators.

To my husband, Ken, my life force and stronghold who never once doubted my abilities, freeing me to learn and continue on ...

To our son, Michael Beecher, who by his very being, became my soul inspiration for this study.

To my learned advisor, Gary Jeffery, for a thoroughly enjoyable and enlightening learning experience. Thanks for sharing Schmidt and your world of toys with me.

To a wise educator, Bill Spain, whose logistics I admire and who is also my good friend.

To Robbie, John, Michael, Nancy and Kathy, the children, parents, relatives and friends of my DAYCARE who helped make what I was doing real.

To Mary Johnston, for typing this work and always excelling in her own.

To the expert raters who were interested enough to rate the many toys for this study.

To the School of Graduate Studies, Memorial University of Newfoundland, for providing the funding for my studies.

TABLE OF CONTENTS

Chapter	Page
I THE PLACE OF TOYS IN EDUCATION AND DEVELOPMENT .	1
Evaluation of Toys	3
II REVIEW OF THE LITERATURE AND STATEMENT OF HYPOTHESES	7
Toys - Education, Development and Research ...	7
Definition of a Toy	8
The Toy-Child Relationship	9
The Need for Careful Toy Selection	10
The Toy-Play Relationship	11
The Educational Uses of Toys: An Overview ...	14
Informal Uses: An Historical Perspective ..	14
The Formal Educational Uses of Toys	17
Therapeutic Uses of Toys	21
The Specifics of Toy Evaluation	22
Toy Lists	22
Criteria Lists and Evaluation Guides	25
Combined Lists and Selection Criteria	28
Representative Toy Evaluation Criteria	31
Age	31
Other Criteria	33
Goodness	33
Safety	34
Durability	35
Play Value	36
Educational-Developmental Value	37
Interpersonal-Social Value	39
III THE TOY EVALUATION INSTRUMENT	41
Background of the Instrument	41
Development of the Instrument	42
The Selection Criteria	42
Toy Selection	43
Rating Scale	44
Pilot Administration	46
The Final Form of the Instrument	46
IV METHODOLOGY	49
Selection of Raters	49
Procedure	49
Analysis	51
V RESULTS	56

Chapter	Page
VI DISCUSSION AND IMPLICATIONS	79
REFERENCES	89
APPENDIX 1. Toy Evaluation Handout	97
APPENDIX 2. Toys Used in the Study	105

LIST OF TABLES

Table	Page
1. A Summary of the Analysis of Variance Results for Criterion Variables 1-7 for Items (I)	58
2. Estimated Variances ($\hat{\sigma}^2$) of the Criterion Variable Ratings	59
3. A Summary of the Analysis of Variance Results for Criterion Variables 1-7 for Rater Category (C)	61
4. A Comparison of the Mean Ratings of the Criterion Variables 1-7 for the Twelve Raters Within the Two Rater Categories	62
5. A Summary of the Analysis of Variance Results for Criterion Variables 1-7 for Raters Within Rater Categories (R:C)	64
6. A Summary of the Analysis of Variance Results for Criterion Variables 1-7 for Category by Item (CI)	66
7. Generalizability of the Summated Ratings of Items When Raters, Raters Categories and Variables are Potential Error Sources	67
8. Matrix of the Mean Correlations of the Ratings of the Six Criterion Variables and the Summated Criterion Variables	77

LIST OF FIGURES

Figure		Page
1.	Generalizability Design for the Full Mixed Model and for the Reduced Model	54
2.	Estimated Generalizability Coefficients for Goodness Ratings	68
3.	Estimated Generalizability Coefficients for Safety Ratings	69
4.	Estimated Generalizability Coefficients for Durability Ratings	70
5.	Estimated Generalizability Coefficients for Play Value Ratings	71
6.	Estimated Generalizability Coefficients for Educational-Developmental Ratings	72
7.	Estimated Generalizability Coefficients for Interpersonal-Social Ratings	73
8.	Estimated Generalizability Coefficients for Summated Ratings	74

CHAPTER I

THE PLACE OF TOYS IN EDUCATION AND DEVELOPMENT

To fully understand a child's development it is necessary to understand that context within which the development takes place. Authors such as Werner and Kaplan (1963) and Schmidt (1973), in their discussions of child development, support this position and suggest that there is an interdependence of development and education. Education for these authors (especially Schmidt) can be considered as consisting of those experiences offered or available to the child in his own human and cultural context.

A child's development reflects learning and this learning in turn reflects the child's interactions with other human beings, with things, and with the language and culture he experiences. Using these experiences, the child creates his own world of meaning (Schmidt, 1973).

This research is focused in a general way on two of the above sources of child experience. These are the interactions of the child with the adults who educate him, and those interactions between the child and the things (especially toys) offered to him.

Of the child-adult interactions, Schmidt (1973) says:

...into their spontaneous care for the child there enters almost imperceptibly, the specific educational concern for the child, the desire to let the child become a certain kind of person. Some spontaneous

2

tendencies of the child are encouraged, others are checked, forbidden or simply eliminated by paying no attention to them. It is not simply by interacting with the child but by the specific educational intent of many of the adult's actions in relation to the child that the child's potentialities as an individual person are actualized. (p. 38)

It is clear for Schmidt that the adult caregiver directs and influences the child's education and consequent development through the experiences he offers (Schmidt, 1973). For one seeking to intentionally direct this education, a key consideration must be the selection of the experiences to be offered. One area of those experiences in which the adult caregiver potentially has a measure of direct control, is that of the selection of play materials which the child is offered or allowed to use (Jeffery, 1977).

Historically, play materials have been given little direct attention (Fraser, 1972; White, 1971). Currently, however, they are being more fully explored. This interest not surprisingly would appear to be increasing because of such obvious factors as increasing toy cost and variety (CTMA, 1978).

A more subtle and perhaps more relevant reason for this recently increased interest, would appear to relate to the potential educational and developmental value of play materials (Fowler, 1980a, 1980b; Gehlbach, 1980; Hutt, 1976; Jeffery, 1977, 1980; Olson, 1970; Schmidt, 1973).

Jeffery (1977), for example, suggests that it is very likely that toys do influence behaviours and further that

they potentially educate. This view is supported in the writings of several authors, such as, Montessori (1948), Berlieter and Englemann (1966), Allen (1968), Buist and Schulman (1969), Olson (1970), Kephart (1971), Zimmerman and Calovini (1971), Stone and Church (1973), Vandenhazel (1976), Kesner and Sunal (1980), and Fowler (1980a). It is on the basis of such concurring views that decisions about the need for careful research into the nature and impact of toys can be made (Dañce & Larson, 1976).

This study seeks to further the understanding of the educational and developmental impact of toys by assessing one aspect of this question, namely, the feasibility of toy evaluation. By addressing the question of toy selection, one can provide a starting point for further studies on the developmental and educational impact of these materials.

Evaluation of Toys

This study will specifically focus on the evaluation of toys.

There has been an increased general interest in the quality and suitability of toys in the past few years (Jeffery, 1977; Kesner & Sunal, 1980; Zimmerman & Calovini, 1971). This interest is reflected in the rapidly growing number of toy-lists available to consumers and in the proliferation of play material evaluation guides (CTTC, 1981-1982; Fowler, 1980a, 1980b; Jeffery, 1977; Wehman, 1979).

It becomes evident, when reviewing the lists and selection guides, that there are both a great diversity of toys to choose from, and an apparent large number of different selection criteria being offered and used. It can also be noted that not all of the authors of these guides and lists appear to have equally well developed rationale for their offerings.

Not all of the authors of toy lists, for example, state or demonstrate how they use their criteria in making their choices. While some authors appear to focus primarily on safety and construction features, others pay more attention to useability and the match of the toy to the child's developmental level. As well, while some lists simply name toys, leaving one to assume that the objects listed all presumably meet some unstated criteria, others offer lists which categorize toys under broad headings such as, age and/or probable type of play supposedly encouraged by the toy.

The criteria guides which have been perused tend to range from being quite general to highly specific. The larger number tend to be of the former nature. Whether or not an age range has been specified in lists and guides, most of them appear to be aimed at a particular age or age group (Allen, 1968; Canadian Toy Testing Council, 1981-1982).

When listing or evaluating toys a broad question relates to the global concept of the overall "goodness" of the toys (CTTC, 1981-1982; Zimmerman & Calovini, 1971).

The very title of one toy list, Good Toys (CTTC, 1980-1981),

reflects an assumption that toys chosen to be on this list, like others, meet some general criteria of goodness. Virtually no research exists on this or any of the above questions.

This research seeks to begin to shed some light on these questions by studying both the overall goodness of toys and also the specific criteria used when evaluating toys. The study also seeks to discover if there is, at least in the opinions of experts, consensus about how to apply these criteria.

This study specifically asks if experts share a general concept of toy "goodness." If the existence of such a shared concept can be demonstrated, a possible basis for the initiation of a consumer toy selection training program may have been found.

In addition to focusing on overall goodness, this study will look at a number of the more specific selection criteria found in the literature on toy evaluation and assess them relative to overall goodness. It will also seek to find out if expert consensus on these particular aspects of toys exists.

Another dimension of this study relates to the homogeneity of views held by expert raters. This study compares the perspectives of those working directly with children (i.e., preschool teachers and consultants) and also those less directly involved with children (i.e., academics specializing in early childhood education and child development). The study uses these two categorical groups of expert

raters to rate pictorially presented toys on a 7-point scale in terms of both goodness and five additional criteria, namely safety, durability, play value, educational-developmental value, and interpersonal-social value.

As an immediate goal, it is hoped that this research will provide useful information on toy evaluation. At a more general level, it is hoped that by enhancing our understanding of these bases, a foundation upon which better and more useful selection criteria and toy lists can be developed, will be established.

CHAPTER II

REVIEW OF THE LITERATURE AND STATEMENT OF HYPOTHESES

Each of the components which make up the framework of this study are discussed briefly below. This chapter will first deal with that literature relating to the educational and developmental nature of toys, and second to the topic of toy evaluation. This second topic will be dealt with initially in a general and then in a more specific manner.

Toys--Education, Development, and Research

Toys might be considered as being tools or instruments which can have an influential developmental force in the child's life. This view has been both directly and indirectly suggested by many authors including Montessori (1948), Berlieter (1966), Allen (1968), Olson (1970), Kephart (1971), Stone and Church (1973), Vandenhazel (1976), and Kesner (1980). Toys can have a potential developmental impact in the sense they may offer opportunities and materials from which the child can learn. Such educational opportunities are offered the child through both the nature and possible uses of the toy and through the particular kinds of interactions between the child and the adult which might be fostered by the toy (Herron, 1972; Levenstein, 1971; Nimnicht & Brown, 1972). Schmidt (1973) maintains that the

education offered in the context of both these interactions, has a central place in the conceptualization of child development. These sentiments are basic to the nature and thrust of this thesis and are echoed in the following words by Jerome Bruner: "Man is shaped by the tools and instruments he comes to use and neither the mind nor the hand alone can amount to much" (Jeffery, 1977, p. 1).

Definition of a Toy

Much confusion exists as to what objects are appropriately labelled toys. Definitions of toys abound (Caplan & Caplan, 1974). The general description which is used in this study is one offered by Jeffery (1977). He proposes a definition compatible with that offered by Olson (1970) who has explored child development through toy research. This definition states that:

A toy is any object or material, whether commercially designed for the purpose or not, which a child is either offered or allowed to use and which the child willingly uses.

While Jeffery's definition is in keeping with the general thrust of this thesis, it does have two major limitations. Firstly, only after an object has been "used" can it be labelled a toy, and secondly, the definition offers little insight into what constitutes "use". Both of these concerns are outside of the specific scope of this study. Operationally, this study will not attempt to definitively define

"toy" but rather will accept the "definition" used by toy sellers. This research will restrict itself to studying those commercially available objects labelled and sold as "toys" by a major retailer. This definition was chosen as it was believed that it would more closely reflect the definition held by parents, the principal toy buyers.

The Toy-Child Relationship

To understand the developmental impact of toys it is necessary to look at the place of experience in development.

Even though the child himself is active in giving meaning to his gradually expanding world, he still has very little freedom to choose the nature of the materials and situations from which he derives these experiences. The child's experiences are basically limited to what he is offered by the adult as, what Schmidt (1973) would term, "educator".

The parent, daycare worker, or caregiver selects experiences to offer a child. The bases for these selections are many and include the offerer's history, values, and philosophy. They also include factors such as family size, income and geographic location. Other influential factors include the physical constraints imposed by the size and nature of the available play and storage areas and the presence of other people be they children or adults. It is not appropriate to deal with these factors here.

While it is appreciated that a child's development is

influenced by that child's experiences, the modification and control of these experiences is not an easy task. One source of child experience over which some degree of control might be relatively easily gained involves the play materials offered the child (Rheingold & Cook, 1975). It is these materials in general and their evaluation in particular to which this study is specifically addressed.

In the selection of play materials it is assumed that all toys can offer a possible learning experience to a given child (Quillitch, 1974). Even though toys may be potentially educational, they are not equally educational (Jeffery, 1977; Quillitch, 1974). The educational worth of a given toy to a particular child will depend on a number of factors, including the educator's goals for that child. In most instances, these goals will take into account such factors as the readiness of the child to assimilate the information available through the toy's use (i.e., information not too difficult or not already possessed by the child). The toy must also suggest, through its design and nature, potential uses to the child (Jeffery, 1977). Once these goals have been established, the adult caregivers can then actually select the toys (Fowler, 1980b).

The Need for Careful Toy Selection

Because the educational experiences offered through a toy are multiple, it is necessary that the "entire" toy be taken into account during selection. For instance, a toy

that encourages ego and drive and builds playfulness with ideas but that limits the child's opportunities for fantasy and encourages destructiveness is useless (Caplan & Caplan, 1974; Tudor-Hart, 1970).

The importance of needing to know what represents a good toy (i.e., one that offers desired experiences) becomes evident. A significant question that arises thus becomes what are representative toy criteria that will discriminate between good and poor toys?

The importance of toy selection criteria is suggested by the many toy evaluation guides and toy lists available to assist consumers in toy selection. It is difficult to know which, if any, to follow. The present study began with a thorough literature review (see p. 7) to discover which criteria are most frequently presented and therefore which, at least in the opinions of their authors, are most important.

The Toy-Play Relationship

As mentioned above, the research on toy selection criteria is scanty. One area in which some empirical study has been conducted, however, is that of the relationship of toys to play. Any impact toys may have can only occur if they are played with by the child.

Although some authors consider play to have no direct worth or purpose (Garvey, 1977; Sylva, Bruner & Genova, 1976),

others such as Piaget and Inhelder (1967), Herron and Sutton-Smith (1971), Ellis and Scholtz (1978) agree that young children's access to play activities is critical to optimal growth and development.

Gehlbach (1980) also agrees that play is important. He points out, however, that the definitional problems of play alone are immense and that the understanding of play as a learning activity is even more limited. Through a cybernetic model of play, he attempts to conceptualize it as a kind of behavioral system. In doing so, he provides an approach to understanding play that allows for empirical and conceptual access to variables which control the conduct of play experiences. These variables include the player, the plaything and the environment in which the two interact (Gehlbach, 1980). These variables, in his opinion, can be manipulated by a researcher, or a teacher.

Much of the available research supporting toy use falls either within this broad area or within the realm of ecological psychology which studies the details and consequences of environmental factors (Eibl-Eibesfeldt, 1970; Sylva, Bruner & Genova, 1976; and Vandenberg, 1978). Such work provides empirical justification for the toy approach offered in this study. For further discussion on play, its definitions and developmental role, the reader is directed to Millar (1968), Caplan and Caplan (1974), Piaget (1962), and Garvey (1977).

As mentioned above, some research has been done in the

various areas of play and the variables that effect it. Further studies in toy-play research explore the specific relationship between toys and the environment (Rheingold & Cook, 1975). That social play can be influenced by play materials has been demonstrated by Quilitch and Risley (1972; 1973), Poling (1976) and McDonald (1979). The relationship of toys to aggressive play (Turner & Goldsmith, 1976) and to imaginative play (Pulaski, 1973; Singer, 1973) have also been empirically studied. The teaching ability of toys has been studied by David Olson (1970) when he looks at how toys can teach the concept of diagonality to their users. These research issues, as they particularly relate to learning in the play situation, are further discussed below.

While several child developmentalists allude to toys (Bruner, 1963; Montessori, 1948; Olson, 1970; Piaget, 1962; Schmidt, 1973; Werner & Kaplan, 1963) and while most research in this area supports toy use indirectly, virtually no research has been found that carefully evaluates toys.

Dance and Larson's model (1976) for systematic theory building suggests that an empirical approach to research only becomes a useful tool after much observation has been done. They suggest that only after such observation can the decision be made to view something as what they term "a phenomenon". The existing amount of written materials suggesting a possible relationship between the nature of a toy and its educational-developmental impact may be

considered as being representative of a phenomenon. This research is aimed at providing a first step in building a theory linking conscious toy selection and desired developmental outcomes in the child.

The Educational Uses of Toys: An Overview

To appreciate the place of toys in development and education, it is useful to consider first the traditional or informal uses of toys and secondly their more current and formal uses. To understand the various uses toys have, it is necessary to look at the reasons for offering them to children.

This section will first look at the early reasons for toy giving from an historical perspective and then at the more current reasons. The three subsections below are based on the assumption that, where identifiable reasons exist for using toys, the actual offering of specific toys must reflect these reasons.

Informal Uses: An Historical Perspective

Historically, toys were seen to be valued by many different culture groups primarily as a means of keeping a child quiet or amused. This has been a very long tradition, for, as Gwen White points out in her book Antique Toys and Their Background (1971), "the early Christians were persecuted because they refused to worship the images of the Roman

Emperor and while hiding they would make toys to amuse the children and keep them quiet" (p. 10). White (1971) further adds that the amusement value today of throwing pebbles into the water or using sticks to float or mud to scoop was prevalent 4,000 years ago. Amongst the oldest toys found, she points out, are what appear to be carved stone balls, made by "stone age peoples".

Toys eventually began to reflect life skills preparation and included such things as tools for building, weapons for fighting, and infant dolls that could be dressed and cared for. These relics, along with other toys such as games, have been found in family graves in various parts of Europe and America. In the period of about 3000 B.C., when peoples were settling along the Nile, Tigris and Euphrates Rivers, toys that were similar to the wagons and pull-carts of the day increased in numbers (White, 1971). Antonia Fraser, in her book entitled A History of Toys (1972) points out that the children of the Greeks and Romans had many different toys made of a variety of materials including bone, wood, leather, and clay which were probably intended for a variety of income groups.

By 1490, more complex moving toys could be found such as "toy windmills with sails, tigers with wagging heads, falcons beating their wings and miniature churches with ringing bells" (White, 1971, p. 12). White goes on to add that in 16th century Europe, it was fashionable to have toy collections typically housed in elaborately carved

cabinets (Jeffery, 1977).

Both during and after the Renaissance, miniaturized toys representing realism flourished, with the word "toy" being used by the craft guilds to describe any tiny copy of the real thing (Jeffery, 1977).

Not only did toys chronicle what had gone on, but they could also predict what was to come (Encyclopædia Britannica, 1967; Huizinga, 1964). It has been pointed out, for example, that the Incas, a civilization which did not use the wheel in their everyday working life, had discovered it and were using it on their children's toys. Another example of toys being tried out on children is found in Europe and England in the period preceding the Industrial Revolution. These toys made considerable use of gears, levers, and miscellaneous mechanical devices.

Although the different kinds of toys and their uses was increasing, it was not until the nineteenth century, when the technique of stamping from sheet-metal was discovered, that various toys were turned out by the thousands (White, 1971). With the advent of this and the other modern techniques, the toy industry flourished into the big business that it is today. The next section discusses this era and its implications for toy use in the more formal educational and therapeutic domains.

The Formal Educational Uses of Toys

While toys have functioned in an informally educational way for a very long time, the beginnings of their being formally used as 'teaching materials' would appear to date back to the seventeenth and eighteenth centuries (Jeffery, 1982b).

Some of the early educators who contributed to laying the foundation for much of our current useage of educational toys are Pestalozzi, Froebel, Edgeware, and Montessori. Pestalozzi (1746-1827) proposed that education should entail direct experience and proceed from learning the concrete to the abstract. Both of these guidelines make obvious the use of learning materials which might, in one sense of the word, be considered as toys. He was also the first to propose the now widely accepted notion that learning occurred in both cummulative and sequential steps. This point, too, makes reasonable the use of perhaps more complex or differential materials, again including toys (Bluhm, 1971a).

Froebel (1782-1852), like Pestalozzi, was an advocate of education for young children. In a discussion of his work, Bluhm (1971b) points out that Froebel developed a series of what he called "gifts" for children intended to direct them toward creative self-activity, self-regulation, and to provide for them the opportunity to manipulate objects. These gifts took the form of specially chosen

or designed objects which were offered to the child. It was assumed that the child, through his use of these objects, would learn the various lessons believed to be offered by or associated with the particular "gift". Examples of gifts included such things as the cube, the ball, the cylinder, and the subdivisions of the cube, that would, Froebel felt, stimulate the young child to grasp the unity of the universe (Blum, 1971b). Depending on one's definition of what is meant by a toy, similarities between 'toys' and 'gifts' can be noted.

Early jigsaw puzzles (i.e., in the later 18th century) may also have been used as educational materials to help children learn large organized bodies of facts. Puzzles are among a large collection of teaching toys collected at the Norwich Museum in England. Two of the earliest examples of educational jigsaws are the maps of Africa and America, published in London in 1772 (Quillitch, 1974).

Richard Lowell Edgeware was a noted educationalist of the early 1800's who also advocated the use of jigsaws and other educational toys. In a book published in 1798 entitled Practical Education, he even included a chapter on toys (Hannas, 1971). In this work he states that "we are disposed to think favorably of any mode which unites amusement with instruction" (Hannas, 1972, p. 72).

It is interesting to note that the Japanese during this time were also utilizing this very method of educating during their Boys' and Girls' festivals. At these events,

they displayed elaborate sets of dolls in order to explain the social hierarchy and history of their country (Fraser, 1966).

In the United States there is also historical evidence of the formal use of educational toys. The Hopi Indians, for example, sought to impart religious education to their children through what were called "Kachina dolls" during their ceremonial rituals (Fraser, 1972). Similarly, such games as Growth of a Century published in 1889, were designed to teach the names of the U.S. Presidents, while the Game of the American Story and Glory, published in 1846, was used to teach American history.

A major and more current figure in the use of educational materials/toys was Maria Montessori. Born in 1870, she was widely known for her materials and educational methods. She considered each child "to be a self-activated learner at work in a prepared environment of programmed materials" (Orem, 1971, p. 390). These programmed materials might be considered as being carefully chosen and offered toys.

In agreement with Montessori, Yawkey and Silvern (1977) also suggest that toys can be regarded as educational aids and sources of potential experiences and learning.

Present day educators as Radford (1972), Olson (1970), May (1977), and Fowler (1980b) have also been actively involved in demonstrating how toys can be used in the classroom as a part of the regular curriculum materials.

Robert Quilitch (1974), in an article called "How Educational are Educational Toys?", discredits the educational claims of the many toy manufacturers on the basis that no research exists to demonstrate that toys actually teach. Appealing to the concerned parent, Quilitch (1974) insists that these manufacturers are bullying them into buying expensive specialty toys. If educational implies teaching a specific skill, his point is well taken in that toy manufacturers, when pressed for such information, have fallen on educational generalizations (Quilitch, 1972). A large number of other educators, however, feel that toys do have an educational value, albeit not always one which can be measured in terms of readily specifiable lessons or clearly and immediately demonstrable skills (Allen, 1966; Berieter, 1966; Fowler, 1980; Gehlbach, 1980; Jeffery, 1977; Kephart, 1971; Olson, 1970; Vandenhazel, 1976; Zimmerman & Calovini, 1971).

As was previously suggested, there is a small amount of research which suggests that toys do influence behaviour (McDonald, 1979; Poling, 1976; Quilitch & Risley, 1972; Rheingold & Cook, 1975), and perhaps even do teach specific lessons (Hutt, 1976; Olson, 1970; Singer, 1973). The most powerful demonstration of this is found in the work of Olson (1970) who empirically showed that children who were allowed to free play with carefully designed equipment, did acquire a previously absent concept of "diagonality". Play with these toys apparently resulted in this learning and

perhaps even "taught" the previously decided upon lessons.

Therapeutic Uses of Toys

Closely related to the educational uses of toys is their therapeutic useage. This is especially evident when therapy is considered to represent the teaching/learning of new skills or the strengthening of weak existing skills.

Toys are being currently used therapeutically in many parts of the world including Europe, Britain and North America (Jeffery, 1982a).

Much of the increased use is associated with the rapid development of toy libraries in providing materials for the handicapped and disabled (Jeffery, 1982a).

Handicapped and developmentally delayed children who typically have had relatively little opportunity to experience mastery and success in many of their activities, can often get those positive experiences from toy play (Jeffery, 1980; Wehman, 1979).

Much has been written about toys and the handicapped, and the reader is directed to Thorum (1976), Guthrie (1979), Jeffery, (1982a), Wehman (1976; 1979), Seidschlaw (1976) and Lear (1977) for a more detailed discussion of the topic.

The Specifics of Toy Evaluation

For the consumer seeking help in selecting a toy, guidance is available from two different types of sources. One source is the "toy list", the other is the "toy evaluation guide". A "list" simply offers a number of supposedly pre-tested or evaluated and hence recommended objects. The "guide" offers a number of criteria which should be taken into account when considering a specific toy. In some instances, a combined list/guide is available.

Several toy lists can be found, such as those offered by Hartley (1968), Stephenson (1977), and James Galt and Company Limited (undated) to name only a few. Similarly, lists of selection criteria are also plentiful and include offerings by Edgington (1962), Allen (1968) and Jeffery (1977). In the literature there also can be found a number of publications which combine the above. Examples of combined lists are those offered by Zimmerman and Calovini (1971), Braga and Braga (1973), Wehman (1976), and Cataldo (1978).

Toy Lists

Toy lists are basically intended to help the consumer choose from the many toys available. Listed toys are typically sub-categorized by age and/or the type of play they would likely foster. It is significant to note that toy lists can be found which focus on materials for both normal as well as many specific groups of exceptional

children. While lists for exceptional children were not focused on, similarities between both groups of lists were found. These were basically related to the categorization of the toys. The differences noted in the lists for exceptional children were in reference to the specific statements made as to the uses and appropriateness of the toys for these children (Guthrie, 1979; Seidschlaw, 1976; Wehman, 1976). Lists can also be found which include toys one should not buy (Swartz, 1971).

Several lists containing toys for normal children, principally, were reviewed. These lists include: 1. "Choosing Toys for Children of All Ages" (American Toy Institute, 1962); 2. "Selective Lists for Giving" (Hartley, 1968); 3. "Toys for Preschool Children" (Stephenson, 1977); 4. "Good Toys and/or The Toy Report" (The Canadian Toy Testing Council 1980-1981, 1981-1982); and 5. "Choosing Good Toys for Young Children" (James Galt and Company (undated)). These lists are considered by the researcher as typical. Other lists available include those by Goodman and Gustavon (1976), Seidschlaw (1976), and Canada Safety Council (1975).

The most widely published and readily available toy lists are those presented by the Canadian Toy Testing Council (CTTC). Revised versions of this list are published annually. Originally, the lists were called 'Good Toys' (from 1952-1981) but are now entitled 'The Toy Report'. In the 1978 publication of this report, for example, more than 500 individual toys were analysed for play value, safety and durability by volunteer CTTC members. By 1982, 1000 toys

were tested and assessment criteria were increased to include function, design and packaging. The catalogue, which is available for purchase, presents each toy by name and offers the manufacturer's recommended ages, approximate price and some miscellaneous comments. The comments are brief and typically related to the size, weight, shape, color and use of the toy. There is no direct consumer input into the evaluation of specific toys. All listed toys, however, have been "tried out" (i.e., informally tested by the children of the volunteers who help out the council). The remaining lists mentioned above are all quite general in nature and very brief in length.

Stephenson (1977) basically lists toys according to four types of play, 1) motor play, 2) constructive play, 3) creative and expressive play, and 4) intellectual play, while James Galt and Company (undated), Hartley (1968) and the American Toy Institute (1962) included age groups as well. While the two latter authors utilized three broad age ranges such as preschoolers (i.e., 3-5 years), early schoolers (i.e., 6-8 years) and pre-teens (i.e., 9-12 years), James Galt and Company Limited listed toys according to eleven specific age groups.

In listing toys under the age and play groupings, only James Galt and Company Limited and the CTTC use specific name brands. Among the other lists proposed, only general types of toys (i.e., blocks, dolls, cars, etc.) are offered (American Toy Institute, 1962; Hartley, 1968; Stephenson,

1977).

In summary, the toy list approach to evaluating toys has an extremely variable format, and there seems to be little consensus about how things should be organized. Most authors organize at least according to age, but the actual age categories vary from list to list, some using very broad and others using more specific categories. Only a few categorize according to the potential use of the toy.

Criteria Lists and Evaluation Guides

Toy evaluation guides or selection criteria lists present criteria to be used by the consumer either when selecting toys in general (i.e., by category of age, sex) or for particular children. Criteria offered are typically grouped according to the type of play likely to be encouraged and by the child's age. As was the case with toy lists, there are also included such specific attributes as safety, aesthetic value, and interpersonal worth. Some criteria are more specifically outlined than others. A number of guides considered typical by the researcher are reviewed below. These include: (1) "Criteria for Selecting Gifts" (Allen, 1968); (2) "Which One Shall I Buy?" (Edgington, 1968); (3) "A Toy Evaluation Guide" (Jeffery, 1977); and (4) "Choosing the Right Toy: A Checklist" (Kesner & Sural, 1980). Other available selection criteria guides include those presented by Rustici (1972), Elliott (1972), Miner (1974), Ostfeld (1975), Callaghan (1979) and

Wittenberg (1979).

The first two guides mentioned above offer general points for consideration when toy buying. Allen (1968), for example, provides the following eight criteria: 1) suitability to the child, 2) balancing the diet of the toys, 3) physically strengthening, 4) intellectually challenging, 5) emotionally satisfying, 6) socially and culturally orienting, 7) structurally sound and, 8) worth their price. In relation to the last criteria, Allen (1968) suggests that even though quality is not always to be associated with expensive toys, in the long run if they remain popular over a length of time they may turn out to be the "thrifty buys".

The author maintains that although the eight general criteria presented above should assist anyone interested enough in trying to buy that "perfect" gift, it is by no means a complete set of criteria.

Edgington (1968) offers only general criteria in her guide. She does this because she claims that each toy has to be evaluated in terms of what is known about the individual child for whom it is intended. The general criteria she outlines include: 1) suitability of the toy to the age and interest level, 2) durability of the toy, 3) adaptability of the toy, 4) whether assistance is needed to help the child with the toy, and 5) feasibility of getting a refund on an unsatisfactory toy, if it is being purchased by mail-order.

Without telling the reader what to buy, Jeffery (1977) provides a very specific evaluation guide that points out what factors might be considered before a toy is purchased. Space is also provided on the guide to actually evaluate a toy as to whether it is quite acceptable, adequate, or unacceptable. The author provides seven general criteria with very specific criteria outlined under each. The general criteria are labelled (1) safety, (2) durability, (3) suitability, (4) challenging and likely to motivate (5) development (physical, sensory, intellectual, interpersonal, personality), (6) economics and (7) miscellaneous. An example of specific criteria for the general criteria (suitability) would include, among others, (1) Can the toy be easily gripped? (2) Can it be easily used and carried? (3) Is it likely to produce undue frustration?

Kesner and Suhail (1980) provide a checklist for evaluating toys using 15 specific criteria. Three major headings were used to structure the 15 criteria and include (1) Safety and Durability, (2) Educational Value and Promotion of Creativity, and (3) Suitability to the Child. A rating scale of 1-5 is used for evaluating each criterion. This is intended to permit the potential buyers to make fine discriminations between very similar toys. Under the major heading 'Educational Value and Promotion of Creativity' for example, five specific criteria were a) stimulates active involvement, b) requires thinking, c) exercises the imagination, d) has unstructured useage

and e) is versatile and has multiple uses.

Combined Lists and Selection Criteria

This section discusses those guides which combine both a toy list and criteria. Seven such guides were reviewed. One of those focused on materials for the handicapped. Included were (1) "Toys as Learning Materials for Preschool Children" (Zimmerman & Calovini, 1971); (2) "A Balanced Diet of Toys" (Illinois Montessori Society, 1972); (3) "Toys and Materials" (Braga & Braga, 1973); (4) "Selection of Play Materials for the Severely Handicapped: A Continuing Dilemma" (Wehman, 1976); (5) "Playing" (Yawkey & Yawkey, 1977); (6) "Practical Pointers" (American Alliance for Health, Physical Education, and Recreation, 1977); and (7) "Activity Organization to Enrich Infant-Toddler Programs" (Cataldo, 1978).

Other available combinations of lists and guides include "Choosing Toys for Children" (Toy Manufacturers of America, undated), "The Story Behind Toys and Things" (Kight, 1976), and "Criteria For Educational Toys for Pre-School Visually Impaired Children" (Guthrie, 1979).

Zimmerman and Calovini (1971) list what they feel to be the qualities of a good toy. They include attractiveness, inviting, good-construction, durability, safety, nontoxicity, ability to challenge, fun and stimulating to the curiosity and imagination. They suggest that developmental evaluations provide a useful and necessary basis for the selection of

toys, especially when these toys are for handicapped or exceptional children. Samples of toys are provided beside each identified need (based on the above evaluations) for each of the four age groups.

The 1972 pamphlet by the Illinois Montessori Society discusses how toys should be judged using three basic criteria: durability, realism and aesthetic acceptability. These criteria are not, however, specifically defined. A toy list is provided with each toy described and listed under one of the following: a) sensoral awareness, (b) language readiness, (c) movement, balance, concentration and independence, (d) numerical understanding, (e) toys in the Sciences, and (f) those in the arts.

Toys were listed by the generic names on both of the above guides.

Braga and Braga (1973) summarize the kinds of materials that can be used with children at different levels of development. The generically labelled toys are listed under each age group.

Nine criteria for selecting materials are offered by Braga and Braga (1973) which include:

- (1) they should provide enjoyment for the child;
- (2) they should be flexible in their usefulness for children of different ages and abilities;
- (3) they should be suggestive to the child of what to do with them;
- (4) they should require some action on the part of the child;
- (5) they should involve the child in the learning of some skill or concept;
- (6) they should be flexible, able to be used in a variety of ways;
- (7) they should not require much adult direction although they may be conducive to adult involvement;

(8) they should be well-made and sturdy enough to hold up under steady use with no sharp, jagged edges or parts which could break easily; (9) Finally, materials should be chosen with reference to the children's cultural background and interests ...
(pp. 106-107)

In an article about toys for the severely handicapped, Wehman (1976) emphasizes programming certain behavioral areas by developing play skills that are fostered by particular groups of play materials. He provides a taxonomy of play materials functionally separated from each other. His eight groups of toys include toys for: infants, development of strength, constructive and creative play, for dramatic and imitative play and for artistic development. Also included are toys that stimulate knowledge and aid in school activity, hobbies, and special interests. The reasons for the groupings plus a series of toy examples are provided under each of the above headings.

An article offered by the American Alliance for Health, Physical Education, and Recreation (1977) emphasizes the sensory and motor areas as key ones for game selection. Recommended games are listed by name, along with a description of each game and a statement of the developmental purpose for it.

Yawkey and Yawkey (1977) provide a similar set of criteria to Zimmerman and Calovini's (1971). To the Zimmerman list they add that one should 1. balance the child's toy collection, 2. consider the aesthetic qualities of toys and 3. be careful of making impulse purchases.

Cataldo (1978), in presenting his list, begins by stating the five areas in which the preschool in which he was involved, was divided. The divisions included the social interaction area; the physical mastery and problem solving area; the sensory stimulation area; and the language learning area. For each area, recommended activities, equipment (toys), and appropriate adult behaviours, were listed along with purposes described for both the equipment and area. For example, in the comfort niche, cuddly toys, teethers, and rockers were situated. Here the adult would smile, sing, and rock or bounce the baby to increase his/her self-esteem and awareness.

Representative Toy Evaluation Criteria

When one reviews the various toy lists and evaluation guides cited above, it becomes apparent that there is a considerable degree of overlap in the criteria used in these sources. This section will identify these criteria. It is significant to note that the criteria, herein identified, provide the bases for the criterion variables used in the Toy Evaluation Instrument, which was developed for use in this study (see Chapter Three).

Age

A frequently cited toy evaluation criterion relates to the age appropriateness of the object. This issue also reflects interest in the age-stage of development question.

Zimmerman and Calovini (1971) voice the frequently held opinion that toys should be chosen keeping in mind the developmental level of the child. Selecting toys using this criterion assumes that the age-related competencies, interest of the child, and stage of development are well defined.

The prevailing stage model which has been presented by Piaget (1962), Hartley (1968), Stephenson (1977) and the CTTC (1978-1979 and 1980-1981) describes four general stages of play mentioned earlier, namely motor play, constructive play, creative and expressive play, and intellectual play. This model assumes that a child naturally moves in keeping with his biological development through these stages. The provision of appropriate toys at the optimal time, it is believed, will facilitate the child's movement through the various stages. While this view in part, was widely held, research emphasizing individual differences points out that not all children will progress through a given stage at a given age (Quillitch, 1972; Smilansky, 1968). For this reason, too strong an emphasis on age appropriateness would likely decrease the efficiency of the rating process.

While age may frequently be overemphasized, it cannot be discounted. As was pointed out above, there is both some justification for age-related toy useage and also for stage sequential useage. For these reasons, ignoring this facet of toy evaluation is inappropriate. It was decided that when operationalizing the study, age must be taken into

account but that it would not be a criterion variable.

Other Criteria

Six commonly cited (albeit somewhat differentially labelled) toy evaluation criteria can be identified in the literature. These are: overall goodness, safety, durability, play value, educational-developmental value and interpersonal-social value. Each of these criteria is individually discussed below.

Goodness. Criterion one, goodness, refers to the overall quality of the toy. Zimmerman and Calovini (1971) and the CTTC (1980-1981) refer to this variable when they discuss what constitutes a good toy. The very title of the CTTC report, entitled Good Toys 1980-1981, assumes that a toy which is a "good one" can be identified. It is a goal of this study to assess more specifically what experts might mean by this term (see pages 76-77).

No specific definition of what is encompassed by the term "goodness" can be found in the literature. In spite of this, authors of toy and evaluation criteria lists, as suggested above, continue to operate as if such a definition can be stated. One can only speculate as to what such a definition might be. Research perhaps entailing initially interviewing experts and toy users is needed to help establish a suitable definition.

In this study, to help differentiate what is meant by

this term, it was decided that a useful starting point might be the identification of those most commonly used toy evaluation criteria, the assumption here being that in using these criteria one was, in fact, delimiting key aspects of toy "goodness". When this was done, five criteria (each individually discussed in detail below) were identified.

Safety. Criterion variables numbers two and three (safety and durability) refer to the structural qualities of a toy. In most of the guides and lists reviewed, some mention was made of safety, hence this variable was included. In this regard, it is significant to note that the Committee on Consumer Policy in a report entitled Safety Requirements for Toys (1975), points out that no toy can be completely safe and that:

It is up to the parent to insure that a toy does not fall into the hands of a child for whom it is clearly unsuitable, whether by reason of age and mental or physical characteristics of the child concerned or other circumstances. (p. 15)

For this reason, ratings of safety appear to be a basic necessity. This point is further emphasized by Kesner and Sunal (1980) who state that although equal emphasis should be placed on each of the selection criteria, a toy of questionable safety should be immediately eliminated from consideration for purchase.

The Canadian Toy Manufacturer's Association (1978) adds

that Canada has pioneered the way in toy safety and that there are few, if any, unsafe toys on the market in Canada today.

A definition was thus discovered for this study based on the above comments and the reviews by Zimmerman and Calovini (1971), Jeffery (1977), Yawkey and Yawkey (1977), Kesner and Sunal (1980), and the CTTC (1980-1981 and 1981-1982). In this study, a safe toy is one that is free from sharp edges, does not injure the child if fallen on, has parts that cannot be swallowed, is nontoxic, nonflammable, washable and hygienic and also will not conduct electricity or give electric shock (see Appendix 1 for this and other criterion variable definitions).

Durability. The importance of toy durability has been discussed in several reviews (Allen, 1968; Braga & Braga, 1973; CTTC, 1980-1981; Edgington, 1968; Jeffery, 1977; Kesner & Sunal, 1980; Yawkey & Yawkey, 1977; Zimmerman & Calovini, 1977).

A durable toy is one that is well-constructed and finished and will still have potential usefulness if parts are lost or broken (see Appendix 1, p. 97).

If a toy is not durable, it is more easily broken and this increases the danger that the child may harm himself. In a discussion of the ease with which some toys can be broken, Caplan and Caplan (1974) suggest that toys easily broken are liable to be frustrating to the child and may

promote a sense of wrong doing. To this point, Tudor-Hart (1970) adds that easily broken toys are not only wasteful but encourage destructiveness in the child user.

Durability can be considered as being related to safety. A toy may not be durable due to poor construction. This often results in parts breaking off which could very likely render the toy unsafe. While there is a relationship between the two, durability is viewed by the researcher as an entity on its own, and was therefore selected for the study.

Play Value. A fourth variable considered relative to toy evaluation was play value. A toy with high play value is one that has a range of alternate uses requiring skills the child is currently or soon likely to master. A high play value toy offers opportunity for varied experiences, does not require one to perform an excessive number of repetitive activities, and encourages the involvement of the child (see Appendix 1, p. 97).

Although this criterion has been entitled play value by the researcher, it has not always been given this label. In many articles, what here is referred to as play value appears to be discussed in terms of a variety of criteria. Play value would appear to be what Jeffery (1977) and Zimmerman and Calovini (1971) refer to when they ask if the toy is challenging. Yawkey and Yawkey (1977), Miner (1974), and the CTMA (1980) similarly refer to a toy's

versatility. Olson (1970) asks about attractiveness.

One must caution here of overgeneralizing, as a term like attractiveness also carries a connotation of attraction as differentiated from use. While something may be attractive, it may not have a high utility.

Play value can be considered as being related to educational value. In most instances, a toy can only have play value if it is played with. The question thus becomes, why is a toy played with. Jeffery (1982b) proposes that a toy will only be played with if the child is learning from it. This author compares the two variables, play value and educational-developmental value, to determine the relationship between them. Alternative explorations on the reasons for toy play can be found among the traditional play theories (Berlyne, 1960; Ellis, 1973; Ellis & Scholtz, 1978; Hutt, 1966; Papert, 1980).

Jeffery (1980) also discusses the economics of play value. He advises the wise toy buyer to estimate the duration of play which is likely to occur with a given toy. By comparing the duration of play relative to toy cost, an estimate of the cost per hour of the play with the toy can be calculated. Interestingly, the Toy Report (CTTC, 1981-1982) couples play value and economical value, as well, when it asks, "would you buy the toy?" (p. v).

Educational-developmental. The fifth criterion assessed in this study is the educational-developmental value of a toy.

This criterion is considered to be different from educational value in the narrow sense of the term (i.e., referring to specific types of particular content taught in school).

A toy with high educational-developmental value is one that fosters such mental abilities as grouping, color recognition, pattern recognition and memory skills, general knowledge, and/or physical abilities such as fine and gross motor co-ordination and general muscle development (see Appendix 1, p. 97). Although there is some dispute as to whether or not a toy can be proven to demonstrate the above skills, there would seem to be a widely held consensus view that a relationship does exist. It is not the intent of this thesis to pursue this issue in detail. For further reading in this area the reader is directed to works by Olson (1970), Quillitch, Christopherson and Risley (1972), Hutt (1976), and Singer (1973).

The Educational-Developmental definition used does not mention particular skills such as numerical understanding, language readiness and sensoral awareness. These are ones listed by the Illinois Montessori Society (1972). These specific skill groups are considered as particular mental abilities of the type defined above.

Other specific mental and physical skills not detailed here could be included under the educational-developmental criterion as well. Kesner and Sunal (1980), for example, include criteria such as "unstructured usage" and "versatility" under their educational construct. For this study

these are considered as aspects of play value. Here again, one can appreciate the overlap between play value and educational value and the need for a clearer understanding of these terms.

Interpersonal-Social. A toy that offers interpersonal-social experiences is one that is likely to involve the child in activities with other people (i.e., child-child interaction, adult-child interaction) and/or to encourage the development of socialization skills (i.e., dramatic play, role playing and/or co-operative, assertive and competitive play) in the child (see Appendix 1, p. 97).

Some authors (i.e., Allen, 1968; Cataldo, 1978; Jeffery, 1977) identify social interaction and interpersonal orientation as significant criteria with which to evaluate toys. Other authors reviewed refer to this sixth criterion in terms of the type of play the toy encourages. These authors discuss topics like dramatic play (Hartley, 1968), "let's pretend play" (Stephenson, 1977), and social play (American Toy Institute, 1962). In this study these types of play are considered to be related to the interpersonal-social criteria used.

Much has been written about the realistic toy. For the purposes of this study realism has also been incorporated into the interpersonal-social variable. While Smilansky (1968) and Jeffery (1977) discuss the restricted range of possible uses of the realistic toy, Singer (1973) points out

that encouragement of make-believe play and imitation of parents through role play occurs through such realistic toy play.

While some research has been done on the influence of toys on interpersonal relationships (i.e., Caplan & Caplan, 1974; Smilansky, 1968), a detailed discussion of this topic is beyond the scope of this study. This research does, however, provide a basis for including this criterion as one of the variables in the study.

The six toy evaluation criterion previously discussed, provided bases upon which the Toy Evaluation Instrument used in this study was developed (see Chapter Three).

CHAPTER III

THE TOY EVALUATION INSTRUMENT

An instrument labelled the Toy Evaluation Instrument (TEI) was specifically developed for use in this study. This instrument and its development are discussed below.

Background of the Instrument

When looking for an instrument to assess the level of consensus among experts about toy selection, it quickly became apparent that finding a readily available instrument with an established reliability and validity was unlikely (Buros, 1978; Jeffery, 1980; Kesner & Sunal, 1980).

While literature pertaining to toy lists and selection criteria was plentiful, experimental research in this area, including that taking into account rater consensus, was very limited (Allen, 1968; Cataldo, 1978; Glass, 1978; Hartley, 1968; Jeffery, 1980; Olson, 1970; Stephenson, 1977; Wehman, 1979; Zimmerman & Calovini, 1971).

In an effort to overcome this difficulty, it was necessary to develop an instrument. The instrument that was developed (namely, the TEI) was based principally on work on toys carried out by Kinseil-Rainey (1972), Wehman (1979), and Jeffery (1980), and on more general literature on rating scale construction reviewed by Kerlinger (1974).

The criterion variables used in Jeffery's Toy Evaluation Guide (1977) are typical of those used in much of the literature written in this area. Many of these criteria were used in this study (see Chapter II for a detailed discussion of this and related literature).

Kerlinger (1974) makes the case for utilizing the rating scale for evaluating items quite explicitly while Kinseil-Rainey (1972) demonstrates the utility of using a popular catalogue for toy selection rather than the more typically used toy lists such as provided by Good Toys (1980-1981). For further discussion, see Kerlinger (1974), Kinseil-Rainey (1972), Glass (1978) and Jeffery (1980).

Development of the Instrument

The instrument used in this study consisted of three components. The first contained the selection criteria for rating the toys, the second included the 62 toys to be rated and the third component was the scale on which the toys were to be rated.

The Selection Criteria

A first step in the development of the instrument was a thorough literature review which sought to discover those criteria most frequently used for toy selection (Cataldo, 1978; Jeffery, 1980; Kesner, 1980; Stephenson, 1977; Wehman, 1976). Toy lists and selection criteria guides were also

studied as these, while not always listing criteria, were assumed to reflect the same. What emerged from this search were the six specific criterion variables which were used in the instrument. Five reflect actual toy attributes while the sixth was a more general rating of overall goodness. The five specific criteria were safety, durability, play value, educational-developmental value, and interpersonal-social value. A full description of these criteria can be found in Appendix 1.

Toy Selection

It has been estimated that there are 20,000 new toys distributed yearly (CTMA, 1978). Given this large number, it was difficult to decide on what would be a representative sample. It was decided that a useful source would be a catalogue such as offered by a major retailer of the big supply companies (i.e., Galt Toys (1982), Louise Kool (1982) or Primary School Supplies Limited (1982)). Several of these catalogues are not readily available to the average consumer and therefore, it was decided to find a more publicly accessible supplier. The Sears Christmas Wishbook Catalogue (1981) seemed to meet this criteria. This has a very large distribution (i.e., Newfoundland 23,000 copies) (Sears Management, 1981). It was felt that this would provide a fair representation of toys that would be readily available for purchase or might be found in the typical Canadian home.

The Sears Christmas Wishbook Catalogue (1981) has 63 pages devoted to toys. An initial screening of the approximate 720 toys was conducted to select clear, colored photographs that depicted a single toy in each picture. From this group, the actual toys for the study were chosen based on age-appropriateness for the 1½-3 year age group, inclusion of a catalogue description, and singular representation of one type of toy.

Duplication of more than one type of toy occurred only when there was a marked difference between two of the same type of toy. The particular age-range was chosen based on the manufacturer's recommendations and on the judgement of the researcher. The researcher's decision about an age-range took into account the safety of the toy and the probability of the child being able to use the object based on such qualities as size, weight, and function of the toy. In this way, it was hoped to eliminate age-appropriateness as a criterion in judging the toys.

Considering the above guidelines, 62 pictorially presented toys were selected for the study (see Appendix 2 for the List of Toys).

Rating Scale

The measuring device used was the equal-appearing interval scale (Kerlinger, 1974). This technique is based on the principle that items, in this case, toys, can be differentially rated or scaled. Each item (toy) is assigned

a scale value by the rater and the strength of the rating indicates the degree to which the rater feels the particular item is described by the scale (Kerlinger, 1974).

Experts rated the items (toys) along a continuum ranging from very poor (a numerical value of 1) to very good (a numerical value of 7). The 7-point scale was chosen over others because it was felt that the response variance would be far greater for a 7-point scale than that likely to be found for a 2, 3, or 5-point scale (Kerlinger, 1974).

T. O. McGuire (1973) discusses some requisites for scales, a number of which are appropriate to this study. These include the following: 1) A scale must be representative of the attitudes in the domain; (2) They must be well-defined for the population of interest; (3) They must be appropriate to the concepts in the domain; (4) They must have polar opposites; and (5) They must be clear in relation to the sample of subjects.

There is no way of knowing, says McGuire (1973), whether complete coverage has been accomplished. Ultimately, he suggests that scale selection is a matter of investigator judgement.

For further information on the nature and use of rating scales, see T. O. McGuire (1973), Kerlinger (1974) and Edwards (1951).

Pilot Administration

Prior to administering the instrument to the expert sample, the final form was carefully reviewed by the researcher and a colleague knowledgeable in the area of children's toys. The purpose of this procedure was to help establish the representativeness of toys that might be offered to 1½ - 3, year olds.

The review entailed rating all the toys using the five criterion variables selected for study. Based on this review, it was decided that two revisions be made. Firstly, a sixth undefined criterion variable was included, entitled "overall goodness". It would provide a basis against which to measure the relationship between the other five criteria. Secondly, the definition of the interpersonal-social variable was revised for a clearer understanding of this criterion. Hence, the final form, as administered, is described in the next section.

The Final Form of the Instrument

Each rater was asked to rate the 62 pictorially presented toys using the six criterion variables on a 7-point rating scale.

The pictures of the toys were individually mounted on three inch by four inch index cards with tape. Beside each picture was taped the description of the toy from the catalogue. It included the size, color, function, and material the toy was made of. Five criteria were defined for the raters on the handout sheet as outlined below:

1. A Safe toy is one that is free from sharp edges, doesn't hurt when fallen on, has parts that cannot be swallowed, is nontoxic, nonflammable, is washable and hygienic, will not conduct electricity or give electrical shock.
2. A Durable toy is one that is well constructed and finished, will still have usefulness if parts are lost or broken, is not likely to have parts breaking off.
3. A toy with Play Value is one that has a range of alternative uses, requires skills the child is currently mastering or soon likely to, offers opportunity for varied experiences, does not require performing an excessive number of repetitive activities, encourages involvement of the child.
4. An Educational-Developmental toy is one that fosters such mental abilities as grouping, colour recognition and memory skills, general knowledge and/or physical abilities such as fine- and gross-motor co-ordination, balancing skills, eye-hand co-ordination, and general muscle development.
5. An Interpersonal-Social toy is one that involves other people (i.e., child-child interaction, parent-child interaction) and/or encourages the development of socialization skills. (i.e., dramatic play, role playing, and/or cooperative, assertive and competitive play.)

The sixth criterion variable "overall goodness" was undefined. This was done so as to assess the relative influence of the remaining five criteria (i.e., safety, durability, play value, educational-developmental value, and interpersonal-social value) on an expert's rating of "goodness" and also to find out if other possible criteria might be used by the experts when making their evaluations. By summing variables 2 through 6 (i.e., safety through interpersonal-social) and comparing this summated score with the goodness score, the researcher hoped to determine the relation, if any, between

goodness and these other criteria.

The rating scale was diagrammatically presented on the top page of each rating sheet for the raters (see Appendix 1 for a copy of the Rating Scale). Each expert was asked to rate the 62 toys twice.

As they appeared on the Directions and Rating Sheets, the following instructions were given to each rater:

Step 1

Directions

- A. Please rate each of the toys on your perception of their overall appropriateness or goodness for children, in general, between the ages of $1\frac{1}{2}$ and 3 years.
- B. Rate each toy individually on a 7-point scale such that one (1) represents a very low rating (i.e., very inappropriate, very poor), while seven (7) represents a very high rating (i.e., very appropriate, very good).
- C. Please record the number of the toy being rated.

Step 2

Directions

- A. Please rate the toys again, considering their appropriateness for the same age bracket of $1\frac{1}{2}$ to 3-year-olds.
- B. This time, rate each toy on the 7-point scale, keeping in mind the 5 definitions.

The complete directions and answer sheets can be found in Appendix 1. Items (toys) can be reconstructed by reference to the toy list in Appendix 2 and the Sears Christmas Wish-Book, 1981. The methodology followed in the study will be discussed in the next chapter.

CHAPTER IV

METHODOLOGY

Selection of Raters

As many experts as could feasibly be contacted were asked to volunteer as raters for the study.

The raters used in this study were professionals whose focus of expertise was children. Those included early childhood development educators, primary preschool curriculum specialists, educational and child psychologists, and day care/nursery school practitioners. In total, 12 experts were identified to take part in the study.

The experts were divided into two categorical groups. Half of these 12 raters were from an academic setting, while the remaining six were practitioners working directly with children and child services (i.e., nursery school and school board settings). Of the academics, five held Doctorate Degrees in the related area and one was at the Master's level. Among the practitioners, five held Diplomas in Primary Education and were Directors of Daycare Schools, while one held a Doctoral Degree and was working as a Primary Consultant within the Department of Education.

Procedure

Each of the twelve raters individually completed 62 toy

ratings twice. First, the expert was asked to rate each toy on the "goodness" variable. Immediately following this, the expert was asked to rate each toy on the safety, durability, play value, educational-developmental and interpersonal-social variables. Each toy was completely rated before going on to the next toy.

The researcher met with each expert for a verbal explanation of the rating task. The information given to each rater was as follows:

We are attempting to test the effectiveness of different ways of training parents in the selection of good-quality toys. In order to accomplish this, it is necessary that we obtain information from experts as to what are quality toys. We would also like to know what the relative contribution of different aspects of these toys are.

To this end we are seeking your assistance.

Now, please read the following directions outlined on pages 1 and 3 of the Toy Evaluation Handout Sheets. If you have any questions related to these directions, please ask them now. (See Appendix 1 for TEI directions).

At this time, each expert was required to rate one toy on the spaces provided in the Toy Evaluation Handout Sheets using the 7-point scale. This toy was randomly selected from the set by each rater.

Afterwards, experts were given the remaining toy cards and asked to complete the ratings on their own. Each expert was asked to record all responses on the Handout Sheet. Each rater was allowed as much time as required to complete the task.

For each rater, the 62 toy cards were randomized in

order to overcome any possible autocorrelational effects among the ratings.

The TEI yielded seven scores per toy. These were the six criterion variable scores plus the one summative score (i.e., the sum of the scores of safety, durability, play value, educational-developmental value, and interpersonal-social values).

Analysis

The first four research questions were examined using standard analysis of variance techniques and by applying Generalizability Theory based on Cardinet, Tourneur, and Allal, (1976; 1981). The fifth question was answered by computing correlations between variables and estimating average correlations for the raters.

Generalizability theory offers a more comprehensive and coherent framework than classical psychometric theory for the study of educational and psychological measurement. Traditionally, reliability has required that the conditions of measurement be equivalent. Generalizability theory claims that each observation belongs to a multitude of possible sets of observation (Cardinet et al., 1976). The authors of this theory propose that when research focuses on the conditions of the measurement rather than on the persons measured, it becomes necessary to transpose the dimensions of the measurement design so as to differentiate conditions

while generalizing over persons (Cardinet et al., 1981). In order to do this, they propose new concepts including the notions of facet of differentiation and facet of generalization as complementary aspects of measurement design. The facet of generalization, say Cardinet et al. (1976), is any source of variation which affects the measures taken of the objects under study. A set of objects or characteristics which are to be compared constitutes what the authors call a facet of differentiation in a study.

In their recent article, Cardinet et al. (1981) discuss some further procedures in relation to the treatment of designs in which the objects of measurement have one or more of the following characteristics, appropriate to this study:

- a) they are constituted by other factors than persons
- b) they are defined by the crossing or nesting of several factors which may be fixed or random, and
- c) they may be defined in several alternative ways.

The objects of measurement in this study are the toys (items), a random variable and the facet of differentiation. This facet is random because the toys selected for the study were considered representative of those available to the consumer.

The facets of generalization were, first, the variables used by the raters to rate a toy. This facet was fixed because it was assumed that of all the possible ways that a toy might be rated, the ones selected were the only ones of

real interest. It will be shown later, how, when the rating is focused on a particular characteristic of a toy, this facet is excluded from the model (see Figure 1 for the reduced model). The second facet of generalization was the raters themselves. This was considered a random variable. The third facet of generalization was the rater category. This was considered a fixed variable because of all the possible ways of categorizing raters, all "expert" raters could be classified into one or the other of the two categories.

The design permitted the computation of generalizability coefficients extending the classical definition of reliability, which has been described as the ratio of true variance to observed variance (Cardinet et al., 1976). In computing generalizability coefficients, the numerator estimates the true variance and contains the components of variance corresponding to the facet of differentiation while the denominator contains the expected variance of the observed scores that pertain to the facet of differentiation or to the interaction of the facets of differentiation and generalization (Cardinet et al., 1981).

Statistical analysis were conducted using the SPSS (1975) and SAS (1979) programs.

The full model, then, could be described in standard terminology as raters within rater categories crossed with variables crossed with toys. This is depicted in Figure 1 and yields eleven sources of variance. This model explored

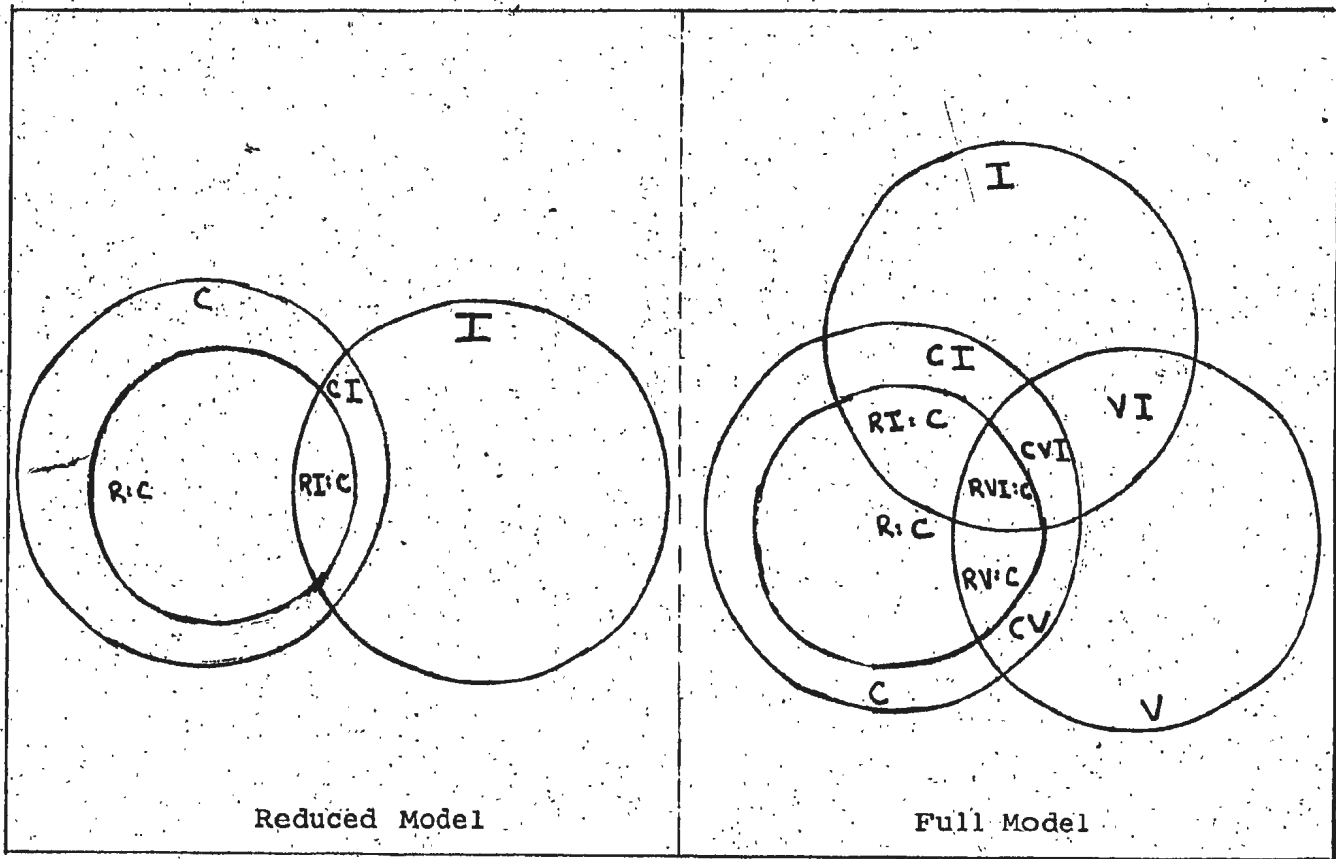


Figure 1. Generalizability Design for the Full Mixed Model and for the Reduced Model

1. C represents two rater categories, academic and practitioner which is a fixed facet.
2. V represents the six toy characteristics or variables being rated, a fixed facet.
3. I represents each toy or item being rated, a random facet.
4. R:C represents six raters within each rater category, a random facet.

the reliability of the summated ratings, which are the average of all the ratings on the several variables rated. It is important because it helps to describe the overall or global view taken by raters of a toy.

It was important also to separately explore the reliability of the ratings on each of the variables which were rated. In this case, the variables do not enter the model as a source of variance. Rather, a reduced model seen in Figure 1 can be described for each of the variables. In standard terminology, this model is raters within rater categories crossed with toys, and yields five sources of variance.

Mean square estimates were found using procedures described in Glass and Stanley (1970). F ratios, and where appropriate, Quasi F ratios, were computed using procedures described in Glass and Stanley (1970) and Winer (1962).

Variance component estimates, universe score variances, and error variances were computed using procedures described by Cardinet et al. (1976; 1981).

In order to answer the question about correlations between variables, correlations between variables were computed for each ~~rater~~, giving 12 sets of correlation coefficients. Average correlations for the 12 raters were computed using Fisher's B Transformation.

The level of significance used for rejection of the null hypothesis was $\alpha = .05$.

CHAPTER V

RESULTS

To gain information useful in answering the specific research questions posed in the study, the toy evaluation instrument (TEI) was used. The administration of this instrument yielded seven scores for each of the 62 toys rated. Six of the seven scores were derived from the criterion variable ratings for the toys (i.e., 1. overall goodness, 2. safety, 3. durability, 4. play value, 5. educational-developmental value and interpersonal-social value). A seventh score was the summation of criterion variables two through six above, to provide an overall rating of each toy. To answer the specific questions posed, three main groups of variables or facets (see Cardinet et al., 1981) were analysed. They were: 1. Toys or items (I), the facet of differentiation which is a random effect, 2. Rater Category (C) a facet of generalization which is a fixed effect, and 3. Raters Within Rater Category (R:C), also a facet of differentiation that is random.

The results of the various analyses will be presented in relation to the specific research questions which they are seeking to answer.

Question 1. Are there significant differences in the ratings of individual toys?

The results of the analysis as applies to this question are found in Table 1. Raters did perceive toys as being significantly different from one another on each of the six criterion variables measured by the TEI. Raters were able to identify toys (items) which rate high and others which rate low and also could discriminate among them. Experts rated toys differently in terms of variable one, goodness; variable two, safety; variable three, durability; variable four, play value; variable five, educational-developmental value; and variable six, interpersonal-social value.

The average ratings of the toys were such that some were judged to be more safe, durable, have more play value, educational-developmental value and interpersonal-social value than others.

Table 1 also presents the results of the analysis of variance for the summated rating of variables safety, durability, play value, educational-developmental value and interpersonal-social value. This summated score suggests that there was a generalized or global difference in the mean ratings of toys.

Toys (items) were the facet of differentiation. Table 2 shows the variance estimates which were calculated. The variance for toys (items) represents the true or universe score variance. Differences in the range of ratings for safety and durability were smaller than for the variables goodness, play value, educational-developmental, and interpersonal-social values. The experts' broadest range of

Table 1

A Summary of the Analysis of Variance Results
for Criterion Variables 1-7 for Items (I)

Criterion Variable	MS (Between)	MS (Error)	df	F ¹
Goodness	21.966441	1.6215	61;610	<u>13.547</u>
Safety	9.933721	1.2095	61;610	<u>8.213</u>
Durability	14.695245	1.5737	61;610	<u>9.338</u>
Play Value	20.887097	1.0950	61;610	<u>19.075</u>
Educational- Developmental	25.531993	1.3651	61;610	<u>18.703</u>
Interpersonal- Social	17.167305	1.5132	61;610	<u>11.345</u>
Summated (Global)	86.111	4.905	61;610	<u>20.168</u>

¹All F ratios significant at or beyond a .05 level are underlined.

Table 2
 Estimated Variances ($\hat{\sigma}^2$) of Criterion Variable Ratings

	C	I	R:C	CI	RI:C
Goodness	-0.0241647 ¹	1.6954109	.1201569	.0504083	1.6215098
Safety	-0.1832681	.7270205	1.257192	0.0646968	1.2094747
Durability	-0.1000558	1.0934603	0.573321	0.0439626	1.573722
Play Value	0.007523	1.6493414	0.2498118	0.1039152	1.09500
Educational- Developmental	0.0529673	2.0139065	.2682179	.0395058	1.3651155
Interpersonal- Social	^p 0.5093939	1.3045096	.3989071	0.0650684	1.5131897
Summated (Global)	-0.0044	0.56834	0.23622	0.01533	0.56667

¹See Gronbach et al. (1972) for a discussion of negative variances

ratings was for the educational-developmental value and the lowest range of ratings was for the safety variable.

Question 2. Do academics and practitioners rate toys significantly differently?

In the full mixed model the analysis of variance results (see Table 3) indicate that no generalized difference was found between the two groups. Specifically, however, the ratings on the interpersonal-social variable were significantly different for the two groups (see Table 3).

In comparing the variance estimates for rater category (see Table 2), the variances among variables were small and negative in the cases of goodness, safety, and durability. Cronbach, Gleser, Nanda, Rajaratnam (1972) suggest that when negative variances are obtained, the variances may be set to zero.

An examination of mean ratings of criterion variables (see Table 4) suggests that practitioners tend to assign higher interpersonal-social scores to toys than do academics. Although not significantly different, practitioners also tended to rate goodness, play value, and educational-developmental values higher than academics. The durability variable was rated the same by both academics and practitioners.

Table 3

A Summary of the Analysis of Variance Results for
Criterion Variables 1-7 for Rater Category (C)

Criterion Variable	MS (NUM)	MS (DENOM)	df	F ¹
Goodness	2.0099	10.9951	26.03;14.58	.153
Safety	12.5772	80.7530	1.22;10.41	.156
Durability	1.7363	18.3017	98.82;11.01	.095
Play Value	21.0905	16.5833	1.11;12.18	1.076
Educational- Developmental	39.3006	19.5968	1.07;11.84	2.005
Interpersonal- Social	217.6434	216.13037634	1.01;11.49	<u>7.732</u>
Summated (Global)	107.819	153.715	1.50;15.78	0.688

¹All quasi-F ratios significant at or beyond a 0.05 level are underlined

The significance of differences of rater categories was tested using the quasi-F procedure described by Winer (1962, pp. 199-200). $F = \frac{MS (C) + MS (RI:C)}{MS (R:C) + MS (CI)}$ - numerator

MS (R:C) + MS (CI) - denominator

Table 4

A Comparison of the Mean Ratings of the Seven
Criterion Variables for the Twelve Raters
Within the Two Rater Categories

		Goodness	Safety	Durability	Play Value	Educational- Developmental	Interpersonal -Social	Summated
Academic								
	1	4.532	5.968	4.823	4.339	3.774	3.516	22.419
	4	4.710	6.000	5.048	4.597	4.645	3.855	24.145
	5	4.194	4.016	3.790	4.065	3.790	3.694	19.355
	8	3.516	2.597	3.145	2.823	2.774	2.355	13.694
	11	4.339	4.645	3.952	3.710	3.694	2.177	18.177
	12	3.807	5.758	3.952	4.258	3.516	3.355	20.839
Group Mean		4.183	4.831	4.119	3.965	3.699	3.159	19.772
Range		3.5-4.7	2.6-6.0	3.0-5.0	2.8-4.5	2.8-4.6	2.0-3.9	18.0-24.0
Practitioner								
	2	4.371	4.726	4.532	3.807	4.581	5.065	22.710
	3	4.548	3.258	2.839	4.145	3.645	3.855	17.742
	6	4.468	5.677	5.403	4.758	4.229	3.403	23.468
	7	4.145	5.016	4.129	4.613	4.032	4.403	22.194
	9	4.113	4.823	3.871	4.419	4.774	4.581	22.468
	10	3.726	4.000	4.113	4.016	3.645	4.113	19.887
Group Mean		4.229	4.583	4.148	4.293	4.151	4.237	21.411
Range		3.7-4.5	3.0-5.7	2.8-5.0	3.8-4.7	3.6-4.7	3.0-5.0	17.7-23.0

Question 3. Do individual experts agree on toy ratings using the TEI?

The results of the analysis as it applies to this first part of the question are found in Table 5. Little consensus was found among the individual experts in their ratings of the six criterion variables assessed in the study. These differences are illustrated by the significant differences found in each of the comparisons.

Overall, experts did not agree on their ratings of the toys for the summated criterion variables. This general difference was significant as suggested by the analysis of variance table for the Full Mixed Model (see Table 5). The estimated variance for Raters Within Rater Category, as seen in Table 2, also points out that raters differ most on the safety and durability variables, while their range of ratings for goodness, play value, educational-developmental and interpersonal-social values is somewhat lower. These differences in variance suggest differences in rater bias with raters having the smallest differences in bias on their ratings for goodness and the greatest interrater bias on the ratings of the safety variable. Within categories of raters, the safety ratings have the greatest range of mean ratings. This is supported by a comparison of the mean ratings for each rater (see Table 4).

Table 5

A Summary of the Analysis of Variance Results
for Criterion Variables 1-7 for
Raters Categories (R:C)

Criterion Variable	MS (Between)	MS (Error)	df	F ¹
Goodness	9.071237	1.6215	10;610	<u>5.594</u>
Safety	79.155376	1.2095	10;610	<u>65.446</u>
Durability	37.119237	1.5737	10;610	<u>23.587</u>
Play Value	16.583333	1.0950	10;610	<u>15.145</u>
Educational- Developmental	17.994624	1.3651	10;610	<u>13.182</u>
Interpersonal -Social	26.245430	1.5132	10;610	<u>17.344</u>
Summated (Global)	109.625	19.753	10.16,80.75	<u>5.550</u> ^{*2}

¹ All F or quasi-F ratios at or beyond a .05 level are underlined.

² F* represents a quasi F ratio.

Question 4. Do academics and practitioners rate individual toys significantly differently?

Academics were significantly different from practitioners in their ratings of particular toys on two variables, namely safety and play value (see Table 6). While the difference among ratings for safety and play value does exist, it is small, as can be seen in Table 2. As well, no overall significant difference was found among the summated ratings for the two groups of raters (see Table 6).

To assess the generalizability of the summated ratings of items (toys) when rater categories (C), raters (R:C) and variables (V) are potential error sources, generalizability coefficients were calculated for each of the criterion variables for both individual raters and the group of raters (see Table 7). The reliability of the rating of one toy with respect to the rating of another toy (ρ_s^2) and also the reliability of the placement of a toy on a scale of 1-7 (ρ_Δ^2) were the two basic concerns of the questions posed. The group of raters overall were found to be reliable

($\rho_\Delta^2 = .920$; $\rho_s^2 = .890$) with criterion variable reliabilities ranging from .878 for safety to a reliability of .948 for play value for ρ_s^2 and .780 for safety to a reliability of .937 for educational-developmental value for ρ_Δ^2 . The reliabilities of ratings for a single rater overall, however, were low for both ρ_Δ^2 and ρ_s^2 (i.e., .414 and .510). The rating reliabilities were particularly low for three of the criterion variables. These ratings ranged from

Table 6
 A Summary of the Analysis of Variance Results for
 Criterion Variables 1-7 for Category by Item (CI)

Criterion Variable	MS (Between)	MS (Error)	df	F ¹
Goodness	1.9239598	1.6215	61;610	1.187
Safety	1.5976566	1.2095	61;610	<u>1.321</u>
Durability	1.8374977	1.5737	61;610	1.168
Play Value	1.718491	1.0950	61;610	<u>1.569</u>
Educational- Developmental	1.6021505	1.3651	61;610	1.174
Interpersonal- Social	1.9036003	1.5132	61;610	1.258
Summated (Global)	5,374	4,905	61,610	1.259

¹All F ratios at or beyond a .05 level are underlined.

Table 7

Generalizability of the Summated Ratings of Items
When Raters, Rater Categories, and Variables
and Potential Error Sources¹

Criterion Variable	Group of Raters		A Single Rater	
	P_s^2	P_Δ^2	P_s^2	P_Δ^2
Goodness	.926	.921	.511	.493
Safety	.878	.780	.375	.228
Durability	.893	.859	.410	.337
Play Value	.948	.936	.601	.551
Educational- Developmental	.947	.937	.596	.552
Interpersonal- Social	.912	.891	.463	.410
Summated (Global)	.920	.890	.510	.414

¹The Generalizability Coefficients were calculated following procedures described by Cardinet et al. (1981)

$$\text{i.e. } P_\Delta^2 = \frac{\sigma^2_I}{\sigma^2_I + (1/12) \sigma^2_{R:C} + 1/12 \sigma^2_{RI:C}}$$

$$P_s^2 = \frac{\sigma^2_I}{\sigma^2_I + (1/12) \sigma^2_{RI:C}}$$

$\rho_s^2 = .375$, $\rho_\Delta^2 = .228$ for safety to $\rho_s^2 = .410$, $\rho_\Delta^2 = .337$ for durability and to $\rho_s^2 = .463$, $\rho_\Delta^2 = .410$ for inter-personal-social value (see Table 7).

From the preceding discussions, it would appear that judgements about toys by only one expert are not likely to be an acceptable basis for toy selection. However, using twelve experts, as was done in this study, appears to produce quite reliable ratings.

Using procedures described in Cardinet et al. (1976, 1981), generalizability estimates were found for different numbers of raters. Figures 2 to 8 show the relationships between the numbers of raters and estimated generalizability for each of the variables rated. It can be seen that, with the exception of the variables, safety and durability, generalizability estimates exceed .80 when four raters are used. These figures can be used to estimate the actual sample sizes of raters required to achieve desired levels of generalizability. Optimum levels of reliability are determined on the bases of researcher judgement and on the situation in which the rating takes place.

Question 5a. Is there a significant correlation between the overall goodness score and the summated variables score?

b. Is there a significant correlation between the play value score and the educational-developmental score?

The mean intercorrelations of the six criterion variables (goodness, safety, durability, play value, educational-

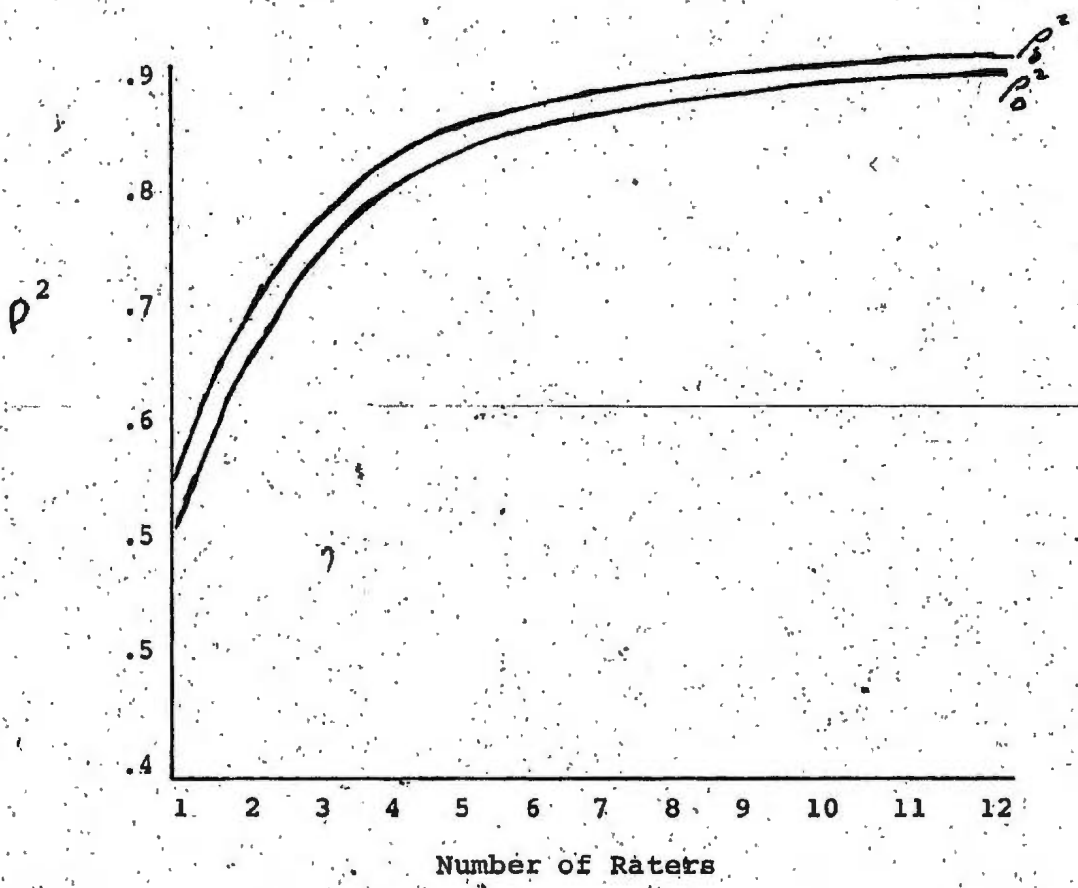


Figure 2. Estimated Generalizability Coefficients for Goodness Ratings

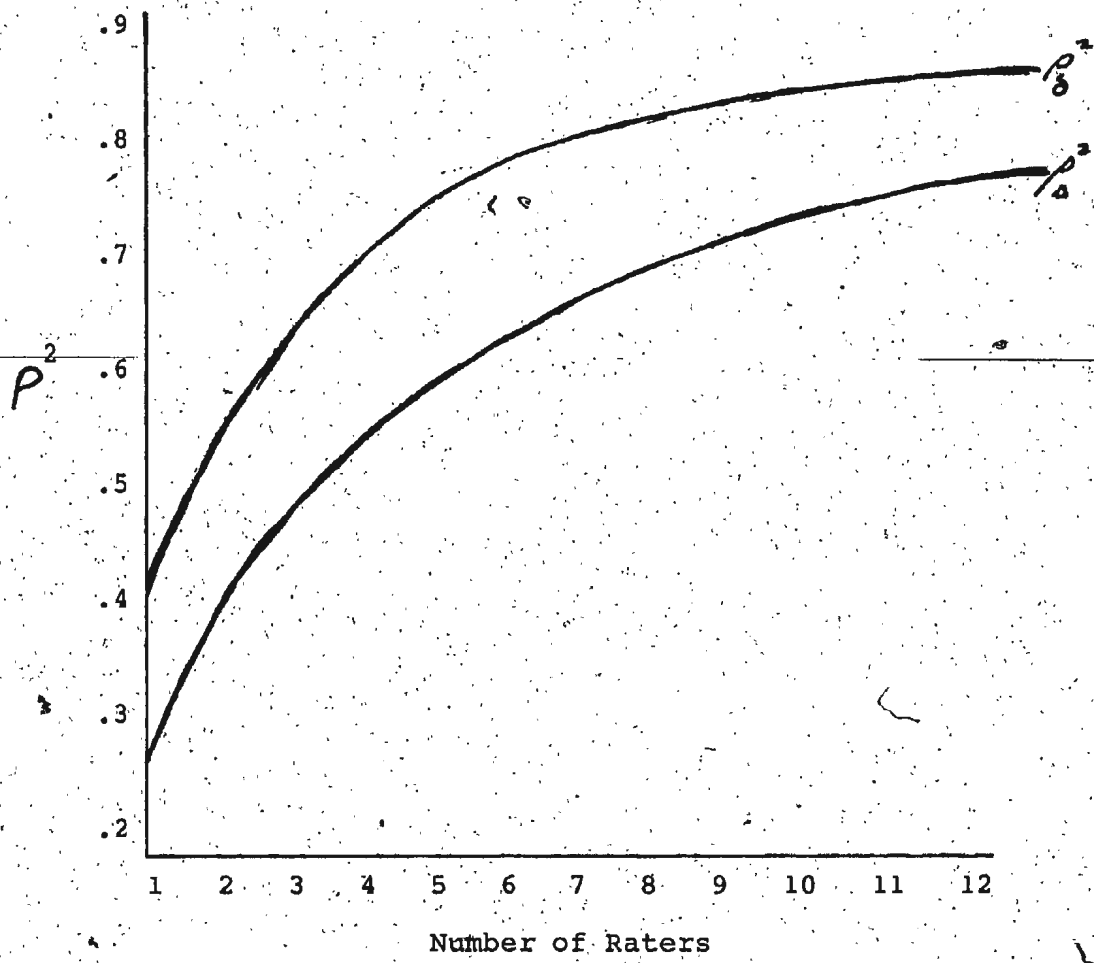


Figure 3. Estimated Generalizability Coefficients for Safety Ratings

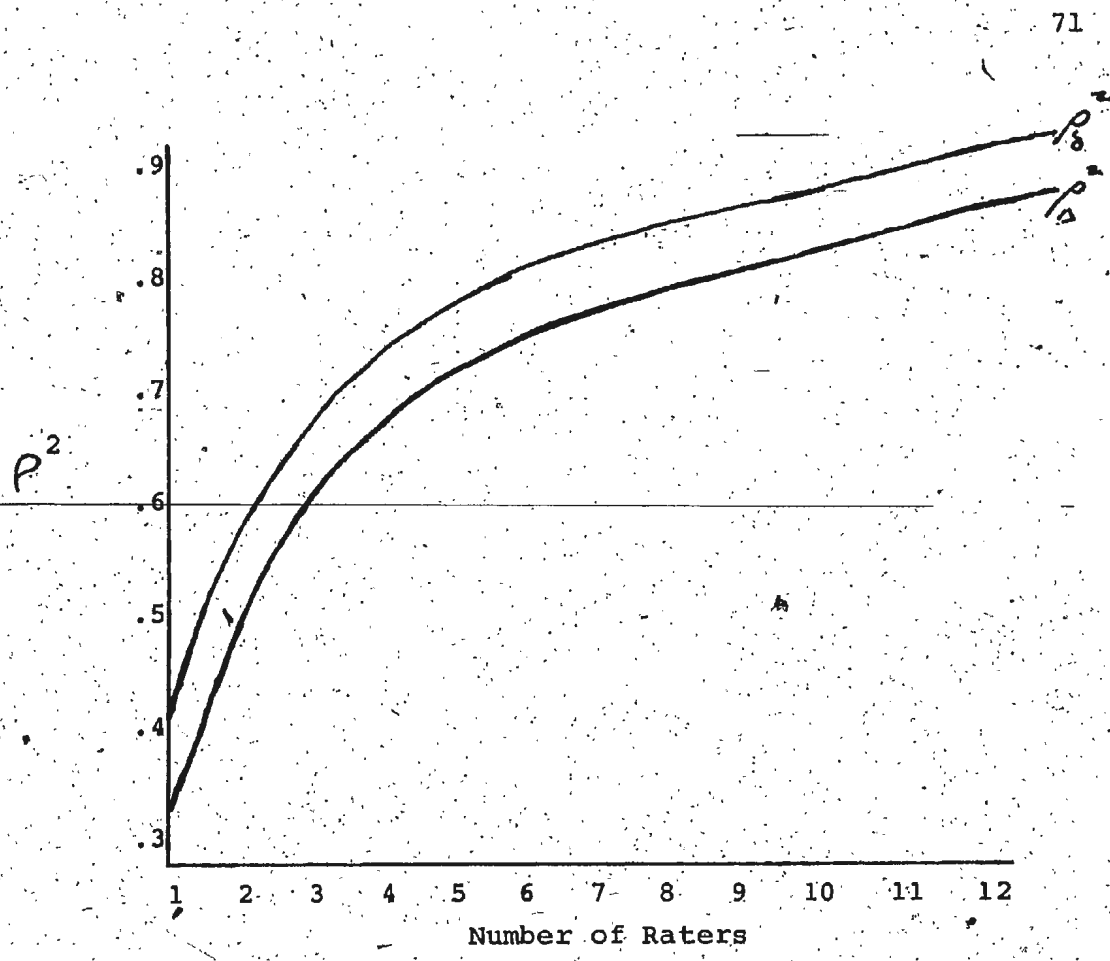


Figure 4. Estimated Generalizability Coefficients for Durability Ratings

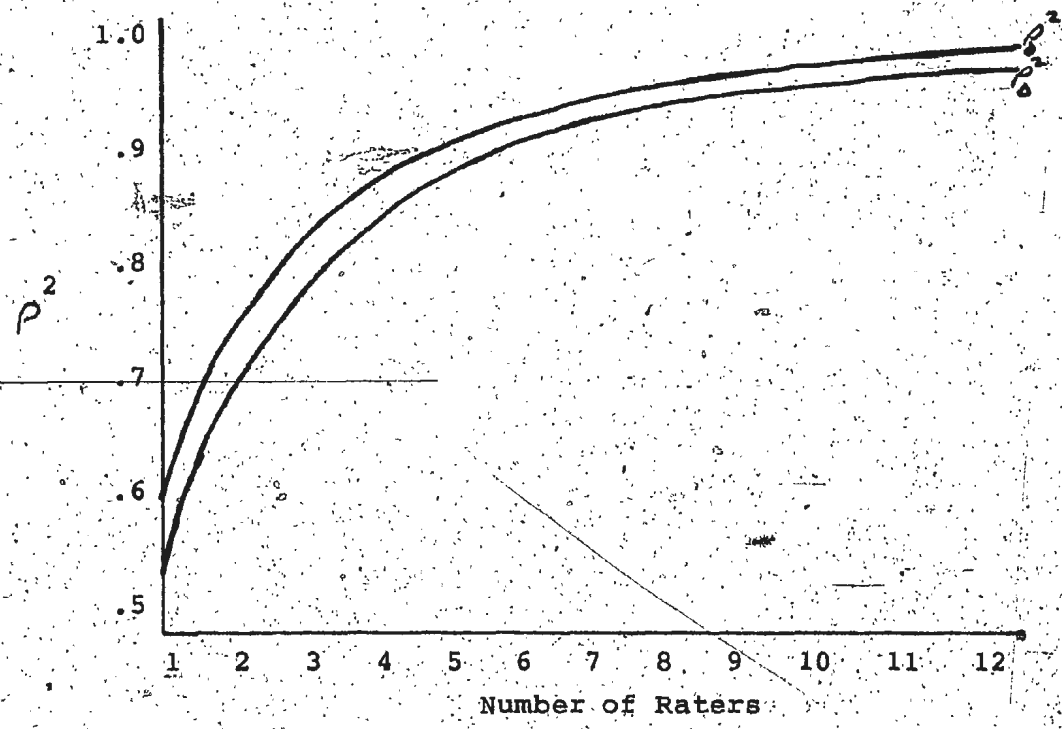


Figure 5. Estimated Generalizability Coefficients for Play Value Ratings

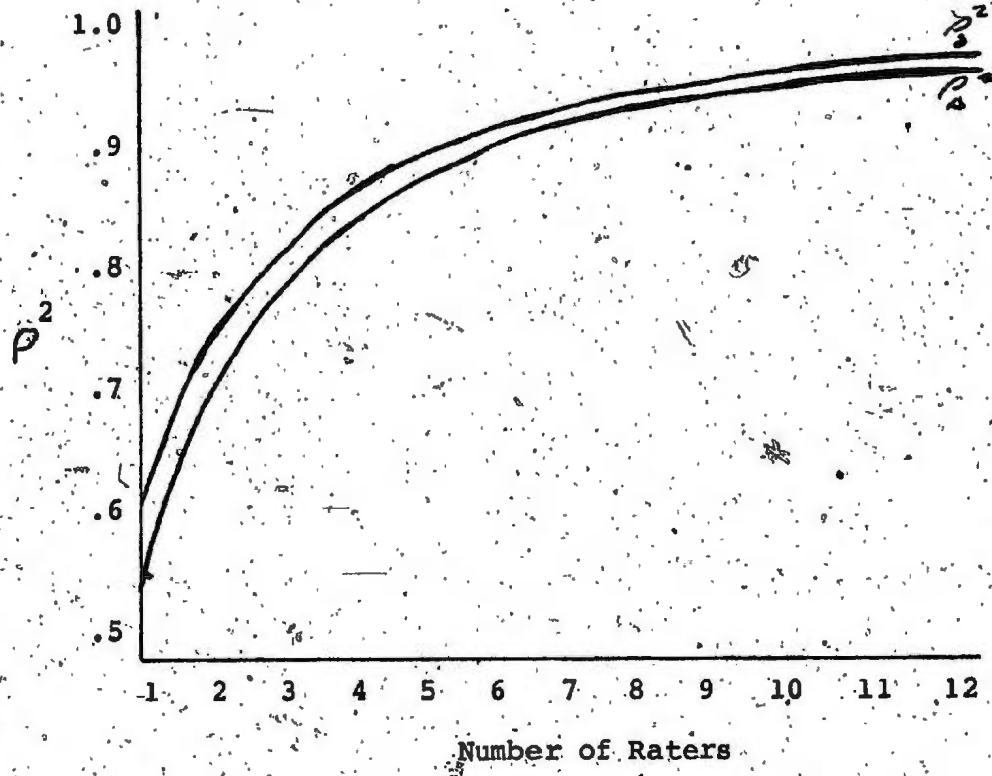


Figure 6. Estimated Generalizability Coefficients for Educational-Developmental Ratings

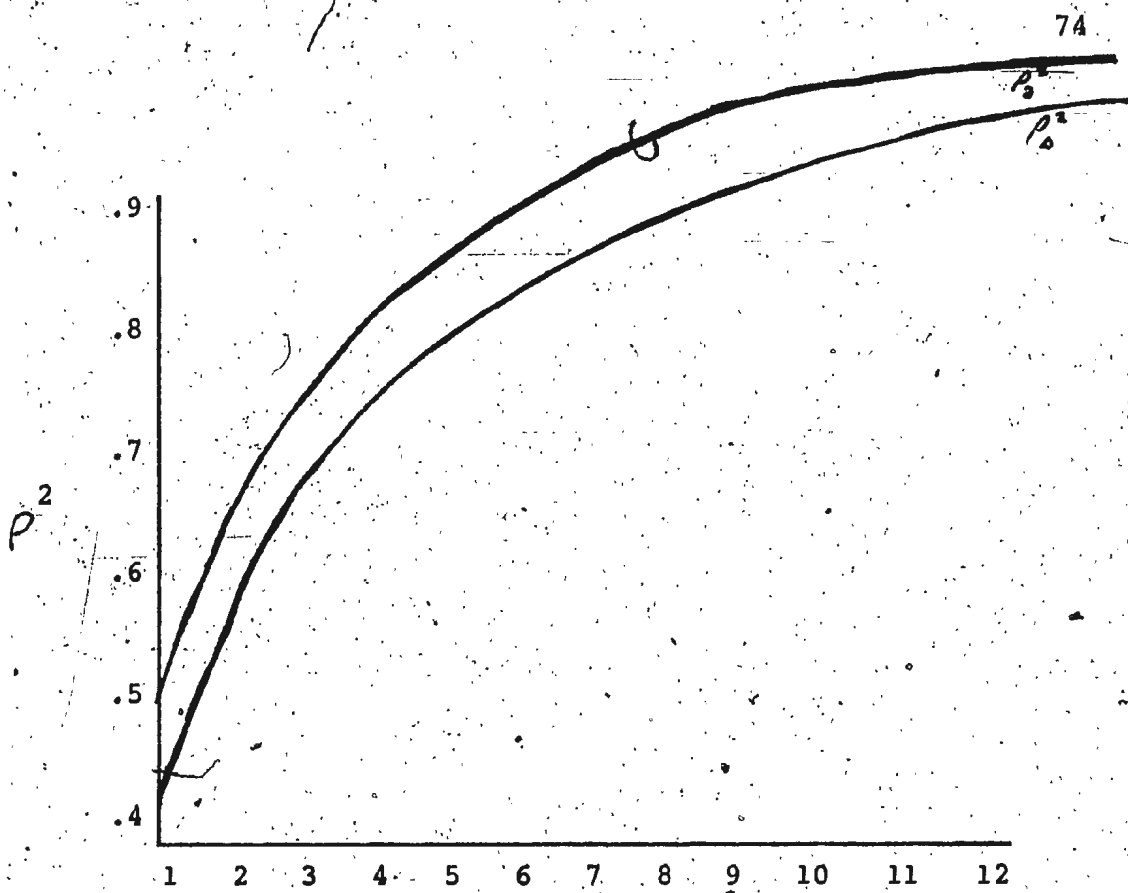


Figure 7. Estimated Generalizability Coefficients for Interpersonal-Social Ratings

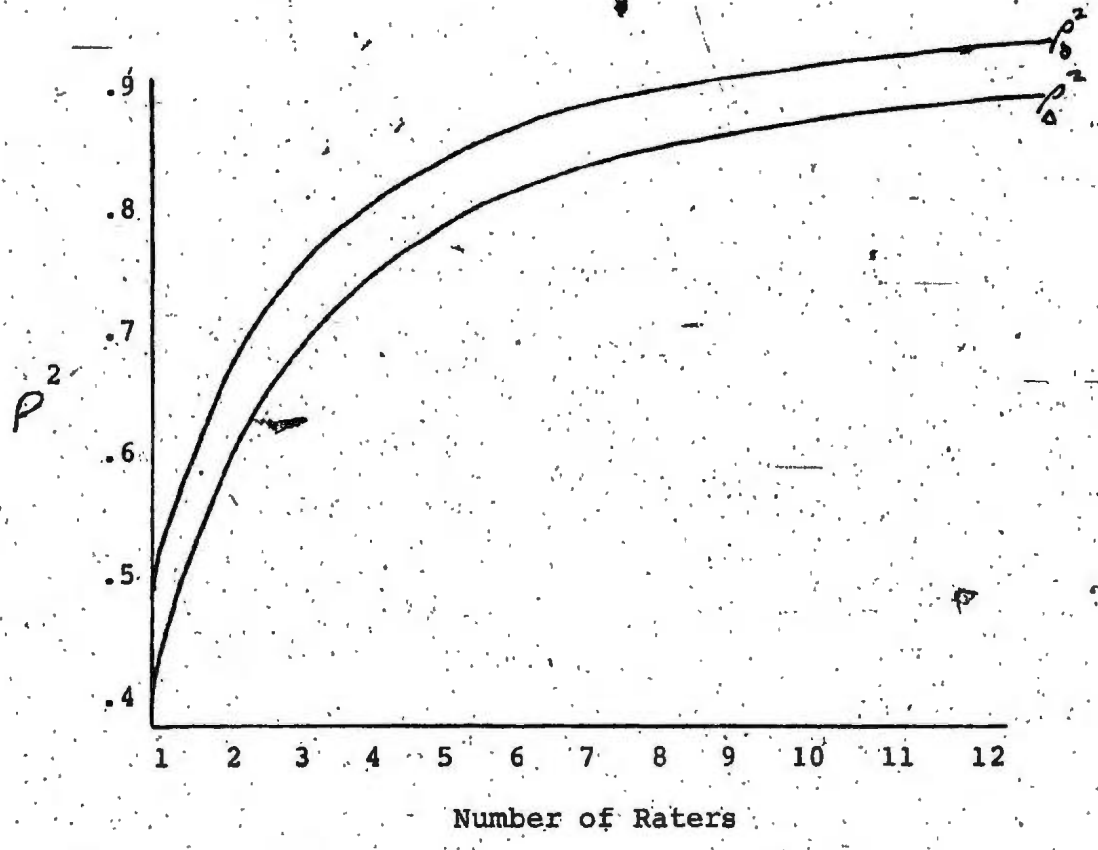


Figure 8. Estimated Generalizability Coefficient for Summated Ratings

developmental and interpersonal-social) and the summated score derived from the five criterion variables scores (i.e., safety, durability, play value, educational-developmental value and interpersonal-social value) are shown in Table 8.

To answer research question 5 a), the overall goodness score and the summated variables score were correlated. These variables were found to correlate significantly ($r = .81$). This correlation is higher than those between the overall goodness score and each of the individual variable scores (see Table 8). It was also higher than those correlations found between the summated score and all but two of individual criteria variable scores. These were play value ($r = .89$) and educational-developmental value ($r = .86$).

Given the size of the correlation between goodness and the summated variables score, it appears possible to estimate one score from the other.

To answer research question 5 b), the play value and educational-developmental scores were correlated. This correlation was found to be significant ($r = .84$). This was the highest correlation found between any two of the single criterion variable scores. This relationship suggests that a toy which has a high educational worth is also rated as having a high play value.

The lowest correlations found were between educational-developmental and safety scores ($r = .40$) and interpersonal-social and safety scores ($r = .29$).

Table 8

Matrix of the Mean Correlations of the Ratings of
the Six Criterion Variables and the
Summated Criterion Variable

	Safety	Durability	Play Value	Educational- Developmental	Interpersonal -Social	Summated
Goodness	.521	.603	.800	.770	.540	.811
Safety		.709	.440	.404	.288	.689
Durability			.586	.518	.387	.786
Play Value				.837	.703	.894
Educational- Developmental					.648	.862
Interpersonal- Social						.766

This suggests that raters tend to view safety and durability as one entity. There is a high correlation between these two variables and a lower correlation between safety and durability and all other variables.

As well, play value, educational-developmental, and interpersonal-social values appear to be a unitary factor. The data in Table 8 suggests that raters tended to rate these variables in a similar manner.

CHAPTER VI

DISCUSSION AND IMPLICATIONS

This study focused on one aspect of the relationship between a child's early development and education, and toys. More specifically, the study sought to discover whether experts in developmental and early childhood education could concur on (1) the "overall goodness" of toys and (2) on specific characteristics of toys, namely their safety, durability, play value, educational-developmental value and interpersonal-social value. The study can be seen as a response to the discovery that the topic of toy evaluation has been relatively unresearched.

Operationally, the study examined the overall or global views on toy goodness held by 12 expert raters. It explored also the reliability of expert ratings on each of the six aspects of toys listed above.

Five questions were posed by the researcher. Firstly, the study sought to find whether raters could discriminate between different toys and if so, did they concur on which toys rated high and low on the above mentioned scales. Questions two, three, and four of the study dealt with potential error variance due to the sampling of the conditions of the rating process. These questions dealt with possible bias due to rater category, differences between individual raters within rater category, and differences in

ratings of specific toys by rater category. The final question of the study dealt with the degree to which raters perceived sameness between the variables rated.

Overall, reliable consensus ratings were obtained for the specific criterion variables. It was also found that a single expert's ratings were quite unreliable. This suggests that lack of agreement between individual raters about the adequacy of toys might be expected. Post hoc analysis suggested that ratings by four experts would be the minimum considered appropriate for reliably rating any given toy. This was found to be the case on all criterion variables, except safety and durability. These latter two variables would appear to require 11-12 raters before they could be considered reliable.

Generally, raters as a group were able to discriminate among the 62 toys. They were able to reliably agree about toys that rate high and low on all variables. Basically, the experts perceived a broad range of differences in the 62 toys on all variables except safety. On this variable, more homogeneous ratings were obtained.

Academics and practitioners as a group generally agreed in their toy ratings. When rating particular toys, however, the Academics and Practitioners rated two toy characteristics significantly differently. These were play value and safety. Individual experts tended to have consistent general biases in their ratings.

One research question looked at the educational-

developmental and play value relationship. These variables were found to be highly correlated. As well, the summated and goodness scores were significantly correlated. This correlation was not as high as those found between goodness and play value, and goodness and educational-developmental value. Raters also appeared to rate according to what might be two major factors (namely a "construction factor" and also a "use factor"). There appeared to be a degree of independence of these two factors. These possible factors are discussed at greater length below.

The results demonstrated that even though a general difference was found between the 62 toys, the range of mean toy ratings for safety was restricted. This may have been for a variety of reasons. Perhaps the toys used were homogeneous on this characteristic. Another explanation might be that raters believed that already approved toys are generally safe. A third reason might be that in the one and a half to three year age range covered, a given toy might be thought unsafe early in the range but safe later in the range, or vice versa. This would lead to ambiguity on the part of raters, hence a greater homogeneity and lower reliability. This is an example of a well-known phenomenon of regression towards the mean.

For the two rater groups, ratings on the interpersonal-social variable were significantly different, with the practitioners' ratings of this characteristic of toys being more extreme. A possible explanation of this is that practitioners

may focus more on interpersonal-social values of toys than do academics because they may have both more opportunity to see toys used in this way and may be relatively less concerned about the child's learning and education. Sharing, getting along together, and caring about one another are perhaps more important to the practitioner than the academic.

Practitioners also tended to perceive a greater amount of difference in the play value and also in educational-developmental value than did academics. These differences, however, were not significant. While a greater awareness of the play value of toys might be more real for the practitioner, especially in these difficult economic times, one would not have expected them to rate educational-developmental value more extremely. It would seem that the academics, in their educational milieu, would be more aware of such concerns. An explanation for this possible trend is that an underlying common factor may form the basis for ratings of all three of these variables (i.e., play value, educational-developmental value and interpersonal-social value). This is supported by the intercorrelations of these variables.

Although there were significant differences in the toy ratings of individual experts on all variables, rater bias for safety ratings and to a lesser degree for durability ratings was consistently more evident than for the other variable ratings. Individual experts differed considerably in what they viewed as safe toys and durable toys. There was a greater amount of difference in these ratings than in

ratings of the other variables.

It was found that the academics and the practitioners also rated particular toys differently on the safety and play value variables. For example, academics and practitioners did not agree on the play value of a doll. In other instances, the toys the academics might have felt to be safe were not so perceived by the practitioners.

It appears that the safety variable may be the one most sensitive to the age of the child. Because the age level (i.e., one and a half to three years) is quite wide, there may be some confounding effects. Toys judged safe for a three year old may simply not be seen to be safe for a younger child. Academics and practitioners may also have differed in the way they responded to this factor for specific toys. A similar difficulty may also exist with the play value factor where age of the child, as it relates to the developmental level of the child, may effect how much a given toy may be played with.

Because educational-developmental and play value are significantly correlated, one might expect that the two groups of experts would also rate toys differently on the educational-developmental value. This, however, was not the case. The two groups may see the educational-developmental value being less effected by actual age of the child than are play value or safety. Even though these variables are highly correlated ($r = .84$), they are certainly not perfectly correlated. Experts saw decided differences

between a toy's play value and educational-developmental value. The reasons or bases for these differences need further study.

The patterns of the correlations calculated, suggest the possibility of the existence of the two factors, mentioned earlier, namely one being the quality of "construction factor" (as reflected by similar ratings on safety and durability) and the other being the "use factor" (as reflected by similar ratings on play value, educational-developmental, and interpersonal-social variables).

The overlap between safety and durability variables ($r = .71$) may relate to the quality of construction factor. As mentioned in Chapter II, a toy which is deemed poorly constructed (see Appendix I for full definition of a durable toy) may often be seen to be unsafe, as well. Even though there is an apparent significantly common element between safety and durability, raters also appeared to perceive differences (see Appendix 1 for definitions of each variable).

The high correlations between play value, educational-developmental value, and interpersonal-social value suggests the reflection of the possible common "use" factor mentioned earlier.

The most significant correlation is that between educational-developmental value and play value ($r = .84$). This statistic suggests that a toy that has play value is one that also has educational-developmental worth. One

might speculate that if the child is not learning from a toy, he/she, in the rater's opinion, is not likely to play with that toy. It also follows from the data that raters may feel that if toys are not played with, the child won't learn from them. The definitions of play value and educational-developmental value were intended to be sufficiently different in nature to preclude any significant correlation between them. Nonetheless, raters consistently associated these two toy variable characteristics.

This study used only toys listed by the manufacturer as suitable for the one and a half to three year olds. Due to this restriction and because the ratings were obtained from experts who had not devoted study specifically to the nature of the relationships of toys and play, one can only generalize findings to other age ranges and other types of raters with caution.

In generalizing from these results, one must also keep in mind that only pictorially presented toys were rated. A replication study using toys which are physically present is needed.

As a result of this study, other possibilities for research present themselves. Principally, the following areas could be explored. A replication of this study which looked at ratings offered by parents of one and a half to three year olds would shed light on how the perceptions of parents about toy characteristics relate to those of experts.

A knowledge of the relationship between the ratings of these groups could provide an assessment of the possible need for parent training and a possible basis for the content of such training.

This study has established that the TEI has the potential to be used as a measure of ability of experts to discriminate between potentially good and poor toys, both in terms of overall goodness and of specific toy characteristics. However, it is the only measure of consensus about toys presently available, that the researcher is aware of. It therefore may or may not reflect what one might find when toys are actually played with by children.

Field testing of actual toy use is needed. Additional studies assessing the relationship between toys and such areas as learning and interpersonal relations are also needed. The whole topic of the degree to which experts can predict these actual relationships is simply not known.

A cautionary note is needed here. While field testing of actual toys might be deemed desirable, such testing cannot realistically be expected to be the long term solution to the toy selection issue. There are simply too many new toys produced annually and too many changes in toy design, nature, etc. (CTMA, 1978) to ever consider field testing as a basis for toy selection. Effective and well-designed criteria and evaluation guidelines, similar to those used in this study, would appear to be a more practical solution to finding which toys are good ones. Based on this research

strategy, it might be possible to develop criteria sufficiently specific to allow valid individual expert ratings. An ultimate goal would be the establishment of criteria which would provide guidelines for use by individual parents.

An important question unanswered by this study relates to the ability of raters, both parents and experts, to discriminate between similar toys with respect to the various toy characteristics. In view of the results of this study, research looking at the play value, and safety variables of individual toys in particular are necessary.

The construct of "toy goodness" also needs very clear definition. This study showed that, given the definitions of the five specific toy variables, two underlying rating factors (namely the construction and use factors) seemed to emerge, both related to the "goodness" construct. This view that "goodness" is not unitary but rather reflects at least these two possible factors raises very practical issues for selectors of toys, particularly when selection must be made between similar toys. The actual nature of these factors and their possible relative weights need to be researched.

Another study suggested by this research would be one assessing the suitability of manufacturer's age recommendations. This need has also been raised by Quilitch (1972). It may be hypothesized that safety could be viewed by experts in the study to be seemingly sensitive to the age of the child. From this hypothesis, age may not necessarily

appear to play as large a role in the evaluation of other variables. One can therefore question age recommendations by manufacturers. Are these based on what appear to be the useability of the toy or simply on the safety factor? The relationship between age, safety, and useability could be further explored to help in finding ways to improve rating criteria.

Another issue raised by this study relates to the lack of specificity in the literature about the relationship between a child's age and developmental level (see Chapter II). It may be possible to hypothesize that raters do not have a consensus opinion about developmental competencies to associate with an age. If this is the case, it would lead to differences in ratings despite accurate application of the rating criteria.

In summary, it is important to mention that toy evaluation does not mean the standardization of toys. Although this study demonstrates that universally "good" toys exist, one must be cautious in giving advice about selecting such good toys for an individual child. The assumption that toys are viewed as categorically good does not necessarily mean that they are suitable for a particular child or that they reflect the interest and aspirations individual parents hold for their children (Schmidt, 1973).

REFERENCES

- Allen, A. Criteria for selecting gifts. Childhood Education, 1968, 45 (3), 118-121.
- American Alliance for Health, Physical Education and Recreation. Practical pointers. Washington, D.C.: Information and Research Utilization Center, 1977. (ERIC Document Reproduction Service No. ED 154 546)
- American Toy Institute. Choosing toys for children of all ages. New York: American Toy Institute, 1962.
- Bereiter, C., & Englemann, S. Teaching disadvantaged children in the preschool. Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Berlyne, D. Conflict, arousal, and curiosity. New York: McGraw-Hill, 1960.
- Bluhm, S. Frederick Froebel. In Dreighton, L. C. (ed.), The encyclopedia of education, 1971a, 4, 98-105.
- _____. Johann Pestalozzi. In Dreighton, L. C. (ed.), The encyclopedia of education, 1971b, 7, 87-92.
- Braga, J. D., & Braga, L. L. Child development and early childhood education. Chicago: Model Cities Chicago Committee on Urban Opportunities, 1973.
- Bruner, I. The process of education. New York: Vintage Books/Random House, 1963.
- Bruner, I. S., Jolly, A., & Sylvak. Play: Its role in development and evolution. New York: Penguin Books Ltd., 1976.
- Buist, C., & Schulman, J. Toys and games for educationally handicapped children. Springfield, Illinois: Charles C. Thomas, 1969.
- Buros, O. K. Mental measurement yearbook, volumes I-VIII, 1938-1978.
- Callaghan, A. Battling baby boredom. Daycare and Early Education, 1979, 6 (3), 48.
- Canada Safety Council. Selecting the right toy for the right child. Canada Safety Council, 1975.
- Canadian Toy Manufacturers Association. Mixing business and pleasure: An outline of the Canadian toy and recreation industry. Kleinburg, Ontario: Sampson Matthews Limited, 1978.

Canadian Toy Testing Council. Good toys. Ottawa, Ontario: Tyrell Press Ltd., 1980-1981.

The toy report. Ottawa, Ontario: Tyrell Press Ltd., 1981-1982.

Cardinet, L., Tourneur, Y., & Allal, L. The symmetry of generalizability theory: Applications to educational measurement. Journal of Educational Measurement, 1976, 13 (2), 119-135.

Extension of generalizability theory and its applications in educational measurement. Journal of Educational Measurement, 1981, 18 (4), 183-204.

Cataldo, C. Z. Activity organization to enrich infant toddler programs. Daycare and Early Education, 1978, 29-32.

Committee on Consumer Policy. Safety requirements for toys. Paris: Organization for Economic Co-operation and Development, 1975.

Cronbach, L., Gleser, G., Nanda, H., & Rajaratnam, N. The dependability of behavioral measurement: Theory of generalizability for scores and profiles. New York: John Wiley and Sons, Inc., 1972.

Dance, F. E., & Larson, C. E. The functions of human communication. New York: Holt, Rinehart, and Winston, 1976.

Edgington, R. Which one shall I buy? Academic Therapy, 1968, 3 (3), 191-193.

Edwards, A. Techniques of attitude scale construction. New York: Appleton, Century, Crofts, Inc., 1957.

Eibl-Eibesfeldt, I. Ethology: The biology of behavior. New York: Holt, Rinehart and Winston, 1970.

Elliot, D. How to select and evaluate materials. New York: Products Information Exchange Institute, 1971-1972 (ERIE Report No. 5).

Ellis, M. Why people play. Englewood Cliffs, N. J.: Prentice-Hall, 1973.

Ellis, M. J., & Scholtz, G. L. The activity and play of children. Englewood Cliffs, N. J.: Prentice Hall, 1978.

- Encyclopedia Britannica. Toys, Volume 22, 1967.
- Fowler, W. Infant and child care: A guide in education in group settings. Boston: Allyn and Bacon, 1980a.
- _____. Curriculum and assessment guides for infant and child care. Boston: Allyn and Bacon, Inc., 1980b.
- Fraser, A. A history of toys. London: Spring Books, 1972.
- Garvey, C. Play. Glasgow: Fontana/Open Books, 1977.
- Gehlbach, R. Instructional play: Some theoretical prerequisites to systematic research and development. Educational Psychologist, 1980, 15 (2), 112-134.
- Glass, G., & Stanley, J. Statistical methods in education and psychology. Englewood Cliffs, N. J.: Prentice-Hall, 1970.
- Glass, G. Standards and criteria. Journal of Educational Measurement, 1978, 15 (4), 237-260.
- Goodman, E., & Gustavon, F. Toys that care, non-sexist teaching aids, pre-school adult. New York: Toys That Care and Other Items, Inc., 1976. (ERIC Document Reproduction Service No. ED 131 848)
- Guthrie, S. Criteria for educational toys. Journal of Visual Impairment and Blindness, 1979, 73 (4), 144-146.
- Hannas, L. The English jigsaw puzzle. London: Wayland Publishers, London, 1972.
- Hartley, R. E. Selective lists for giving. Childhood Education, 1968, 45 (3), 122-127.
- Herron, R. E., & Sutton-Smith, B. Child's play. New York: Wiley, 1971.
- Herron, M. A toy can be more than a plaything. American Education, 1972, 8 (10), 21-24.
- Huizinga, J. Homo ludens: A study of the play element in culture. Boston: The Bacon Press, 1964.
- Hutt, C. Exploration and play in children. In Bruner, J. S., Jolly, A., & Silva, K. (eds.) Play its role in development and evolution. Middlesex, England: Penguin Books, 1976.
- James Galt and Company. Galt early stages. Cheshire, Cheshire: James Galt and Company Ltd., 1982.

Choosing good toys for young children. Cheadle, Chesire: James Galt and Company, Limited, undated.

Jeffery, G. H. Toys: Their developmental role and evaluation. Paper presented at the Maritime Day-Care Association Conference, Halifax, N.S., 1977a.

Toy evaluation guide. Newfoundland: Department of Educational Psychology, 1977b.

Influencing children's recreation, play, and learning: The powers of toys. Paper presented at the Canadian Association for Physical Health, Education and Recreation, St. John's, Nfld., 1980.

The uses of toys with handicapped and disabled children. Paper presented at the Canadian Society for the Study of Education Conference, Ottawa, 1982a.

The power of toys. Book in preparation, 1982b.

Kephart, N. C. The slow learner in the classroom. Columbus, Ohio: Charles E. Merrill, 1971.

Kerlinger, F. Foundations of behavioral research. New York: Holt, Rinehart and Winston, 1974.

Kesner, J., & Sunal, C. Choosing the right toy: A checklist. Instructional Journal of Early Childhood, 1980, 12 (2), 76-80.

Right, J. The story behind toys 'n things. Daycare and Early Education, 1976, 3 (5), 15-17.

Kinseil-Rainey, L. Incidental learning of sex role characteristics. Unpublished master's thesis, Southern Illinois University, 1972.

Lear, R. Play helps: Toys and activities for handicapped children. London: Heinemann Health Books, 1977.

Levenstein, P. Symposium on parent-centered education: Learning through (and from) mothers. Childhood Education, 1971, 48 (3), 130-134.

Louise Kool and Sons. Today's Child. Toronto: Louise Kool and Sons Ltd., 1982.

May, D. Toying around in elementary science. Science and Children, 1977, 15 (2), 18-19.

- McDonald, R. Different play materials and the social interaction of disabled and nowdisabled preschool children. Unpublished master's thesis, Dalhousie University, 1979.
- McGuire, T. O. Semantic differential methodology for the structuring of attitudes. American Education Research Journal, 1973, 10 (4), 295-306.
- Millar, S. The psychology of play. New York: Penguin Books, 1968.
- Montessori, M. The discovery of the child. Adyar Madras, India: Kalakshatra Publications, 1948.
- Nie, N., Hull, C., Jenkins, J., Steinbrenner, K., & Bent, D. Statistical package for the social sciences. New York: McGraw-Hill Inc., 1975.
- Nimnicht, G., & Brown, E. The toy library: Parents and children learning with toys. Young Children, 1972, 28 (2), 110-116.
- Olson, D. R. Cognitive development: The child's acquisition of diagonality. New York: Academic Press, 1970.
- Orem, R. C. Montessori today. New York: G. P. Putnam's Sons, 1971.
- Ostfeld, B. We've been asked. Daycare and Early Education, 1975, 4, 30-31.
- Papert, S. Mind storms. New York: Basic Books, Inc., 1980.
- Piaget, J.— Play dreams and imitation in childhood. London: Routledge and Kegan Paul, 1962.
- Piaget, J., & Inhelder, B. The psychology of the child. New York: Basic Books, 1967.
- Poling, A. The effects of particular toys on the social play of preschool children. Mental Retardation Bulletin, 1976, 4 (2), 85-93.
- Primary School Supplies Limited. 1982 school catalogue. Ontario: Primary School Supplies Limited, 1982.
- Pulaski, M. A. Toys and imaginative play. In Singer, J. (ed.), The child's world of make-believe. New York: Academic Press, 1973.

- Quillitch, H. R., Christopherson, E. R., & Risley, T. The organization of group care environments: Toy evaluation. Paper presented at the American Psychological Association, Honolulu, 1972.
- Quillitch, H. R., & Risley, T. R. The effects of play materials on social play. Journal of Applied Behavioral Analysis, 1973, 6 (4), 573-578.
- Quillitch, H. R. How educational are educational toys? Washington, D. C.: Office of Education, 1974.
- Radford, D. Science models and toys. New York: MacDonal Educational, 1974.
- Rheingold, H. L., & Cook, K. V. The contents of boys and girls' rooms as an index of parents behaviour. Child Development, 1975, 46, 459-463.
- Rustici, I. Criteria for selecting realia. Journal of Early Childhood Education, 1972.
- Schmidt, W. H. O. Child development: The human cultural and educational context. New York: Harper Row, 1973.
- Seidschlaw, B. Play and learn with toys and games: A bibliography of toys and games that teach institutionalized children. Pierre, S. D.: South Dakota State Library Commission, 1977. (ERIC Document Reproduction Service No. ED 127 924)
- Simpson-Sears. Christmas wishbook catalogue. Halifax, N.S.: Simpson-Sears Ltd., 1981.
- Simpson-Sears Management. Personal communication, October 1981.
- Singer, I. L. The child's world of make believe. New York: Academic Press, 1973.
- Smilansky, S. The effects of sociodramatic play on disadvantaged pre-school children. New York: John Wiley and Sons Inc., 1968.
- Stephenson, A. Toys for preschool children. Prospects, 1977, 7 (4), 528-539.
- Stone, L., & Church, J. Childhood and adolescence: A psychology of the growing person. New York: Random House, 1973.
- Swartz, E. Toys that don't care. Boston: Gambit Inc., 1971.

- Sylva, K., Bruner, J., & Genova, P. The role of play in the problem-solving of children 3-5 years old. In Bruner, J., Jolly, A., & Sylva, K. (eds.), Play. New York: Basic Books, 1976.
- Thorun, A. Instructional materials for the handicapped. Salt Lake City: Olympus Publishing Company, 1976.
- Toy Manufacturers of America. Choosing toys for children. New York: Toy Manufacturers of America Inc., undated.
- Tudor-Hart, B. Toys, play and discipline in childhood. London: Routledge, 1970.
- Turner, C. W., & Goldsmith, D. Effects of toy guns and airplanes on children's antisocial free-play behaviour. Journal of Experimental and Child Psychology, 1976, 21 (2), 303-315.
- Vandenberg, B. Play and development from an ethological perspective. American Psychologist, 1978, 33 (8), 724-738.
- Vandenhazel, B. J. Toys as teaching aids. The Educational Courier, 1976, 2, 12-14.
- Werner, H., & Kaplan, B. Symbol formation. New York: Wiley, 1963.
- Wheman, P. Selection of play materials for the severely handicapped: A conditioning dilemma. Education and Training of Mentally Retarded, 1976, 11 (1), 46-50.
- Recreation programming for developmentally disabled persons. Baltimore: University Park Press, 1979.
- White, G. Antique toys and their background. New York: Arcó, 1971.
- Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Company, 1962.
- Wittenberg, H. Toy choosing guidelines. Child Focus, 1979, 10-17.
- Yawkey, T., & Silvern, S. Play and playing processes of the young child in early education programs: A Piagetian analyses. Pennsylvania State University: University Park Press, 1977. (ERIC Document Reproduction Service No. ED 144-717)
- Zimmerman, L., & Calovini, C. Toys as learning materials for preschool children. Exceptional Children, 1971, 37 (9), 642-654.

APPENDIX I
TOY EVALUATION HANDOUT

TOY EVALUATION HANDOUT

Dear Rater,

We are attempting to test the effectiveness of different ways of training parents in the selection of good-quality toys. In order to accomplish this, it is necessary that we obtain information from experts as to what are quality toys. We would also like to know what the relative contribution of different aspects of these toys are.

To this end, we are seeking your assistance.

Please read the following directions and rate the toys accordingly.

Thank you

Name _____

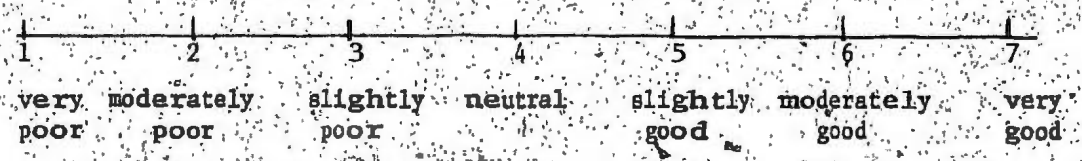
Expert's Toy Rating

You are asked to rate all of the 62 toys pictured, twice. First, rate them in terms of general or overall goodness, and secondly, in more specific ways.

Step 1

Directions

- A. Please rate each of the toys on your perception of their overall appropriateness or goodness for children, in general, between the ages of 1½ and 3 years.
- B. Rate each toy individually on a 7-point scale such that one (1) represents a very low rating (i.e., very inappropriate, very poor), while seven (7) represents a very high rating (i.e., very appropriate, very good)
- C. Please record the number of the toy being rated.



<u>Toy Number</u>	<u>Goodness</u>	<u>Toy Number</u>	<u>Goodness</u>
1.	_____	11.	_____
2.	_____	12.	_____
3.	_____	13.	_____
4.	_____	14.	_____
5.	_____	15.	_____
6.	_____	16.	_____
7.	_____	17.	_____
8.	_____	18.	_____
9.	_____	19.	_____
10.	_____	20.	_____

Name _____

Step 2

Directions

- A. Please rate the toys again, considering their appropriateness for the same age bracket of 1½ to 3-year-olds.
- B. This time, rate each toy on the 7-point scale, keeping in mind the 5 definitions listed below:
1. A Safe toy is one that is free from sharp edges, doesn't hurt when fallen on, has parts that cannot be swallowed, is non-toxic, nonflammable, is washable and hygienic, will not conduct electricity or give electrical shock.
 2. A Durable toy is one that is well constructed and finished, will still have usefulness if parts are lost or broken, is not likely to have parts breaking off.
 3. A toy with Play Value is one that has a range of alternative uses, requires skills the child is currently mastering or soon likely to, offers opportunity for varied experiences, does not require performing an excessive number of repetitive activities, encourages involvement of the child.
 4. An Educational-Developmental toy is one that fosters such mental abilities as grouping colour recognition, pattern recognition and memory skills, general knowledge and/or physical abilities such as fine-and gross-motor co-ordination, balancing skills, eye-hand co-ordination and general muscle development.
 5. An Interpersonal-Social toy is one that involves other people (i.e., child-child interaction, parent-child interaction) and/or encourages the development of socialization skills. (i.e., dramatic play, role playing, and/or cooperative, assertive and competitive play.)

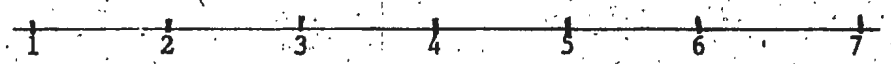
Name _____

1	2	3	4	5	6	7
very poor	moderately poor	slightly poor	neutral	slightly good	moderately good	very good

<u>Toy Number</u>	<u>Goodness</u>	<u>Toy Number</u>	<u>Goodness</u>
21.	_____	42.	_____
22.	_____	43.	_____
23.	_____	44.	_____
24.	_____	45.	_____
25.	_____	46.	_____
26.	_____	47.	_____
27.	_____	48.	_____
28.	_____	49.	_____
29.	_____	50.	_____
30.	_____	51.	_____
31.	_____	52.	_____
32.	_____	53.	_____
33.	_____	54.	_____
34.	_____	55.	_____
35.	_____	56.	_____
36.	_____	57.	_____
37.	_____	58.	_____
38.	_____	59.	_____
39.	_____	60.	_____
40.	_____	61.	_____
41.	_____	62.	_____

U

Expert Ratings



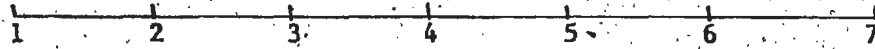
very moderately slightly neutral slightly moderately very
poor poor poor good good good

Educational - Interpersonal -
Developmental - Social -

<u>Toy Number</u>	<u>Safe</u>	<u>Durable</u>	<u>Play Value</u>	<u>Educational - Developmental Value</u>	<u>Interpersonal - Social Value</u>
1.	_____	_____	_____	_____	_____
2.	_____	_____	_____	_____	_____
3.	_____	_____	_____	_____	_____
4.	_____	_____	_____	_____	_____
5.	_____	_____	_____	_____	_____
6.	_____	_____	_____	_____	_____
7.	_____	_____	_____	_____	_____
8.	_____	_____	_____	_____	_____
9.	_____	_____	_____	_____	_____
10.	_____	_____	_____	_____	_____
11.	_____	_____	_____	_____	_____
12.	_____	_____	_____	_____	_____
13.	_____	_____	_____	_____	_____
14.	_____	_____	_____	_____	_____
15.	_____	_____	_____	_____	_____
16.	_____	_____	_____	_____	_____
17.	_____	_____	_____	_____	_____
18.	_____	_____	_____	_____	_____
19.	_____	_____	_____	_____	_____
20.	_____	_____	_____	_____	_____

Expert Ratings

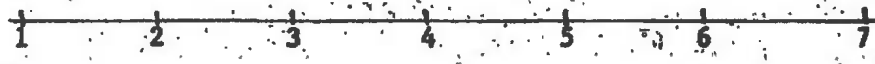
Name



very poor moderately poor slightly poor neutral slightly good moderately good very good

Toy Number	Safe	Durable	Play Value	Educational - Developmental Value	Interpersonal - Social Value
21.	_____	_____	_____	_____	_____
22.	_____	_____	_____	_____	_____
23.	_____	_____	_____	_____	_____
24.	_____	_____	_____	_____	_____
25.	_____	_____	_____	_____	_____
26.	_____	_____	_____	_____	_____
27.	_____	_____	_____	_____	_____
28.	_____	_____	_____	_____	_____
29.	_____	_____	_____	_____	_____
30.	_____	_____	_____	_____	_____
31.	_____	_____	_____	_____	_____
32.	_____	_____	_____	_____	_____
33.	_____	_____	_____	_____	_____
34.	_____	_____	_____	_____	_____
35.	_____	_____	_____	_____	_____
36.	_____	_____	_____	_____	_____
37.	_____	_____	_____	_____	_____
38.	_____	_____	_____	_____	_____
39.	_____	_____	_____	_____	_____
40.	_____	_____	_____	_____	_____

Expert Ratings



very poor moderately poor slightly poor neutral slightly good moderately good very good

Toy Number	Safe	Durable	Play Value	Educational -	Interpersonal -
				Developmental Value	Social Value
41.	_____	_____	_____	_____	_____
42.	_____	_____	_____	_____	_____
43.	_____	_____	_____	_____	_____
44.	_____	_____	_____	_____	_____
45.	_____	_____	_____	_____	_____
46.	_____	_____	_____	_____	_____
47.	_____	_____	_____	_____	_____
48.	_____	_____	_____	_____	_____
49.	_____	_____	_____	_____	_____
50.	_____	_____	_____	_____	_____
51.	_____	_____	_____	_____	_____
52.	_____	_____	_____	_____	_____
53.	_____	_____	_____	_____	_____
54.	_____	_____	_____	_____	_____
55.	_____	_____	_____	_____	_____
56.	_____	_____	_____	_____	_____
57.	_____	_____	_____	_____	_____
58.	_____	_____	_____	_____	_____
59.	_____	_____	_____	_____	_____
60.	_____	_____	_____	_____	_____
61.	_____	_____	_____	_____	_____
62.	_____	_____	_____	_____	_____

APPENDIX 2

TOYS USED IN THE STUDY

1. Cash register
2. 70-piece building set
3. Ride-on running shoe
4. Bert and Ernie dolls
5. Humming top
6. Walt Disney nesting blocks
7. 'Pooh' toy radio
8. Mickey Mouse talking phone
9. Rail runner musical train
10. Do-see-do musical centre
11. Chime ball
12. Wind-up alphabet train
13. Pooh guitar
14. Bath activity center
15. Building beakers
16. 10-piece tool kit
17. Animal coaster
18. Strawberry Shortcake snail cart
19. Winnie-the-Pooh rocker
20. Miss Piggy and Kermit cars
21. Jetliner
22. Pooh stacking blocks
23. Pooh says ...
24. Sesame Street swimmers
25. Yoda, the jedi master
26. Faireyland express
27. Pull-A-Tune express

28. Pooh "wooden blocks"
29. 3 swimmers
30. Mr. & Mrs. Potatoe Head
31. Snoopy the cook
32. Lacing board
33. ABC wood block set
34. Form fitter learning cube
35. Push-button phone
36. Big band play set
37. Peg bench
38. Pooh story tube
39. Matchbox "super fast"
40. Sesame Street counting scale
41. 24-piece kitchen set
42. Land speed racer
43. Musical crib toy
44. Musical roly poly clown
45. Gumball vending bank
46. Winnie-Pooh take-a-part loco
47. Bear shape sorter
48. Drumming panda bear
49. Pooh sit and drive
50. Pooh musical T.V.
51. Shape fitter truck
52. Pooh poppin top
53. Jack-in-the-box
54. Pooh train
55. Kooky cook's set
56. Baby Brenda
57. Tug boat and barge
58. Snoopy sno-cone machine
59. Magnetic play school
60. Snoopy copter-pull toy
61. Circus train
62. Frog-chorus piano

