

AGE DIFFERENCES IN THE PERCEPTION OF THE EMOTIONAL CONTENT OF NONFILTERED  
AND FILTERED VOCAL EXPRESSIONS

by

KENNETH W. McCLUSKEY

---

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL  
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF  
MASTER OF ARTS

---

DEPARTMENT OF PSYCHOLOGY

---

UNIVERSITY OF MANITOBA  
WINNIPEG, MANITOBA, CANADA

JUNE, 1974

AGE DIFFERENCES IN THE PERCEPTION OF THE EMOTIONAL CONTENT OF NONFILTERED  
AND FILTERED VOCAL EXPRESSIONS

by

Kenneth W. McCluskey

A dissertation submitted to the Faculty of Graduate Studies of  
the University of Manitoba in partial fulfillment of the requirements  
of the degree of

MASTER OF ARTS

© 1974

Permission has been granted to the LIBRARY OF THE UNIVERSITY OF MANITOBA to lend or sell copies of this dissertation, to the NATIONAL LIBRARY OF CANADA to microfilm this dissertation and to lend or sell copies of the film, and UNIVERSITY MICROFILMS to publish an abstract of this dissertation.

The author reserves other publication rights, and neither the dissertation nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

## PARTIAL COPYRIGHT LICENCE

In the interest of facilitating research by others at this institution and elsewhere, and in consideration of my having been accepted for enrolment as a graduate student at this institution, I hereby grant a licence to the University of Manitoba to make copies of any thesis or dissertation or substantial parts thereof, the copyright in which may vest in me, which I may submit to this institution in the future, provided that this licence is subject to the following conditions:

1. This licence shall continue for the full term of the copyright or for so long as may be legally permitted.
2. The Universal Copyright Notice shall appear on all copies made under the authority of this licence.
3. It shall also be permissible for this institution to submit the aforementioned thesis or dissertation to the National Library of Canada for microfilming and to make available or to authorize the National Library of Canada to make available microfilm copies.
4. This licence does not permit the sale of authorized copies at a profit, but does permit the collection by the institution or institutions concerned of charges covering actual costs.
5. All copies made under the authority of this licence shall bear a statement to the effect that the copy in question is being made available in this form by authority of the copyright owner solely for the purpose of private study and research and may not be copied or reproduced except as permitted by the copyright laws without written authority from the copyright owner.

In exceptional cases, where adequate cause can be shown to delay publication, the student may request the Dean of Graduate Studies to restrict access to any copies of a thesis or dissertation delivered to the University of Manitoba, for a period of up to three years after submission and the Dean, in consultation with the student's Advisor, shall determine for what period, if any, such access will be so restricted.

June 27/74  
(Date)

\_\_\_\_\_  
(Signature of Student)

## ABSTRACT

Two types of speech samples, semantic-controlled and semantic-free, were considered. Semantic-controlled vocal expressions were recorded by female actresses, who recited a standard paragraph while attempting to simulate the emotions of happiness, sadness, love, and anger. Semantic-free stimuli were subsequently prepared by using an electronic filter (which removed semantic content while leaving intact the tonal qualities of speech) to render the original samples unintelligible. The nonfiltered and filtered speech samples were then played to separate groups of boys ranging in age from six through eleven. It was found that children listening to the nonfiltered samples were able to identify the emotional content of speech significantly more accurately than those hearing the filtered vocal expressions. In both conditions, discrimination ability was a positive monotonic function of age. The negative emotions (anger and sadness) were judged significantly more accurately than the positive emotional states (happiness and love) by children in both conditions.

## ACKNOWLEDGEMENTS

I would like to extend sincere thanks to my academic advisor, Prof. R.R. Niemi, for his invaluable assistance in the planning and design of this study, as well as for his guidance with respect to the stylistic elements of thesis writing. Indeed, I feel that Prof. Niemi's method of supervision -- "directing without confining" -- has been, in large measure, responsible for my very survival in graduate school.

In a very real sense, I have had two advisors for this thesis. During Prof. Niemi's leave of absence, Dr. M.P. Janisse kindly stepped in as interim advisor. Many of the ideas implemented in this research were the direct result of one of Dr. Janisse's graduate courses. In addition, his constructive (and forthright) criticisms of the work in general, and his hints regarding how to write "scientifically," were much appreciated.

Considerable thanks are also extended to Dr. L. Brockman for reading the manuscript and offering many useful suggestions for improving the style and form of presentation.

As well, I am grateful to Mr. P. Kryszkewicz for designing and constructing the filtering device; to Mr. J.P. Smyrski for helping with the data analysis, and to Mr. G.L. Badali for binding numerous copies of the thesis.

At a more personal level, I have to thank my family: Jill (and Harold) for bearing with me (i.e., keeping absolutely silent) during the writing of this paper; Dad, for his frequent "emergency" trips to and from University (and various other points in the city) on my behalf; and Mom, for her help with typing, xeroxing, proofreading, etc., etc., etc. throughout my academic career. Certainly, had it not been for her encouragement (i.e., threats of

retribution), I may never have completed the present investigation. There is a great deal of truth to the long-standing anecdote in our household that she deserves the M.A. as much (or more) than I.

And finally, I want to thank my future wife Cristina for making all the work worthwhile. She prepared all the graphics presented in this thesis, along with several other diagrams for term papers. More important, however, was her moral support during our "period of suffering."

## TABLE OF CONTENTS

	<u>Page</u>
ABSTRACT .....	i
ACKNOWLEDGEMENTS .....	ii
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
INTRODUCTION .....	1
OVERVIEW .....	1
EMOTION .....	4
VOCAL COMMUNICATION OF EMOTION .....	5
The Verbal Channel .....	6
The Vocal Channel .....	6
JUDGEMENT OF THE NONVERBAL ASPECTS OF SPEECH BY ADULTS .....	6
Recitation of Nonmeaningful Material .....	7
Recitation of Standard Content .....	8
Filtered Speech .....	8
JUDGEMENT OF THE NONVERBAL ASPECTS OF SPEECH BY CHILDREN .....	10
INCONGRUENT COMMUNICATIONS .....	13
Adult Responses to Incongruent Communications .....	14
Developmental Differences in Response to Incongruent Communications .....	15
Incongruent Communication and Pathology .....	17
STATEMENT OF THE RESEARCH PROBLEM .....	19
Judgement of Standard Nonfiltered Vocal Expressions .....	21
Judgement of Filtered Vocal Expressions .....	22

	<u>Page</u>
METHOD .....	23
SUBJECTS .....	23
APPARATUS .....	23
STICK FIGURE DRAWINGS .....	24
RECORDING OF VOCAL EXPRESSIONS .....	25
RATING OF VOCAL EXPRESSION .....	26
PROCEDURE .....	28
RESULTS .....	30
AGE DIFFERENCES IN ACCURACY OF EMOTIONAL PERCEPTION IN NONFILTERED AND FILTERED TREATMENTS .....	30
RESPONSE BIAS IN NONFILTERED AND FILTERED TREATMENTS .....	41
Activity .....	41
Valence .....	41
DISCUSSION .....	47
GENERAL INFERENCES AND CONCLUSIONS .....	47
FUTURE CONSIDERATIONS .....	51
SUMMARY .....	54
REFERENCES .....	55
APPENDIX A .....	60
APPENDIX B .....	62
APPENDIX C .....	64
APPENDIX D .....	66
APPENDIX E .....	68
APPENDIX F .....	71
APPENDIX G .....	73



	<u>Page</u>
APPENDIX H .....	75
APPENDIX I .....	77
APPENDIX J .....	79
APPENDIX K .....	81
APPENDIX L .....	83
APPENDIX M .....	85
APPENDIX N .....	87
VITA .....	89

## LIST OF TABLES

	<u>Page</u>
1. Analysis of Variance for Mean Number of Correct Identifications of Each Emotional State (Scored Against Adult Ratings) as a Function of Age and Filtering .....	31
2. Significant Differences in Performance between Various Age Groups in the Nonfiltered Condition .....	35
3. Significant Differences in Performance between Various Age Groups in the Filtered Condition .....	36

## LIST OF FIGURES

	<u>Page</u>
1. Stick figures representing the four emotional states .....	25
2. Mean number of correct identifications at successive age levels in nonfiltered and filtered conditions .....	32
3. Mean number of correct identifications of nonfiltered and filtered expressions for each category of emotional meaning at successive age levels .....	37
4. Mean number of correct identifications (across age) for each category of emotional meaning in nonfiltered and filtered conditions .....	40
5. Mean number of negative (versus positive) responses emitted in the nonfiltered condition as a function of age .....	42
6. Mean number of negative (versus positive) responses emitted in the filtered condition as a function of age .....	43
7. Mean number of active (versus passive) responses emitted in the nonfiltered condition as a function of age .....	44
8. Mean number of active (versus passive) responses emitted in the filtered condition as a function of age .....	46

## INTRODUCTION

### OVERVIEW

The present research was directed at examining the ability of children to identify the emotional content of vocal expressions. In general, investigation in this area is based on the assumption that speech is composed of two channels: the verbal channel made up of articulated patterns of sound forming syllables, words, sentences, and other linguistic forms; and the vocal channel containing the nonverbal characteristics of sound such as voice quality and variation of tone. While the verbal channel presumably carries the semantic meaning of a communication, the vocal channel is thought to carry emotional information. In normal speech, of course, both sets of cues are present simultaneously.

A number of investigators have approached the study of communication of emotion in human speech by looking at the vocal channel alone. In their classic investigation, Davitz and Davitz (1959a) tested the hypothesis that emotion can be reliably communicated by the nonverbal aspects of speech. In an effort to remove the semantic meaning from communication, they had speakers recite parts of the alphabet while attempting to simulate various emotions. Adult subjects then made judgements as to what emotion the speaker was intending to communicate.

Electronic filters have also been used to remove semantic meaning before playing the speech samples to judges (Soskin & Kauffman, 1961; Starkweather, 1956). This latter technique may be preferable in that subjects are actually allowed to speak words, rather than trying to express emotion while saying the alphabet. Furthermore, the filtering method can be employed in studies concerned with real-life situations. For example,

taped speech samples obtained from therapy sessions, telephone conversations, radio, television, and so on can be filtered and played to subjects; acting need not be involved.

The literature on vocal expression of emotion, reviewed by Starkweather (1961), Davitz and Davitz (1961), Kramer (1963), Davitz (1964), Mahl and Schulze (1964), and Duncan (1969), indicates that adults in general are able to identify the affective content of both nonfiltered and filtered vocal communications.

While studies in the area involving adult subjects have been fairly well documented, relatively little is known about the development of sensitivity to the emotional meaning of speech. Gates (1927) assessed the ability of children from grades three to eight to judge the emotional content of portions of the alphabet recited to express each of a number of different emotions. She found that there was a progressive increase with age in ability to identify correctly the emotional content of the voice samples.

In another of the studies done with children, Dimitrovsky (1964) attempted to isolate the vocal channel of speech by holding the verbal channel constant. In her speech samples, speakers were asked to recite the same standard paragraph while simulating the emotions of happiness, sadness, love, and anger. When these samples were played to children aged five through twelve, it was likewise found that the ability to judge vocal expression of emotion was positively related to age.

Due to the paucity of developmental research of this kind, it is necessary to gather more information regarding age differences in the perception of nonfiltered, semantic-controlled speech. The initial aim of

this study, then, was to re-examine the problem of development of emotional sensitivity to vocal expression using Dimitrovsky's (1964) standard paragraph technique. In addition, since the filtering method has not been used in research with young subjects, the question arose as to whether children would be able to identify the emotional meaning of speech samples filtered such that there would be no words to "hook onto." Consequently, a further aim of the present research was to explore the relation between a child's age and his ability to identify the affective content of filtered vocal expressions.

In summary, then, the subject for investigation in this study was an assessment and comparison of the ability of children of different ages to judge the emotional content of two kinds of vocal expressions: nonfiltered speech, in which the vocal channel was isolated by speakers reciting a standard paragraph; and filtered speech, in which an electronic device was used to remove semantic meaning. The main thrust of this research was descriptive, in the sense that the chief concern was with the behaviour of children in terms of their responses to the emotional content of communications, rather than with the developmental causes and underlying mechanisms involved.

Preliminary consideration will be given to presenting a theoretical rationale for this type of investigation. Discussion will then be extended to encompass developments in the field using both adults and children as subjects. A detailed presentation of the research proposal itself will follow, along with rationale for its importance to the general area of the development of the communication of emotion.

## EMOTION

As Arnold (1968, 1970) argues, "emotion" is an extremely illusive concept that is impossible to define unequivocally. Throughout history, certain philosophers have assumed that it is emotion that gives rise to physiological change. To illustrate, they felt that if a person is anxious, the anxiety may cause certain bodily changes such as sweating, trembling, and the like (cf., Arnold, 1968). On the other hand, an opposing view holds that emotional experience does not occur until after physiological change; the bodily states precede and cue the emotion which is felt (Lange & James, 1922). Other theories affirm that physiological change and emotional experience occur simultaneously; neither causes the other (Cannon, 1927). Still other viewpoints contend that the emotion we feel is determined by our interpretation of our bodily states (Schachter & Singer, 1962). Further, cognitive theories of emotion assert that the same stimulus does not necessarily always produce the same emotional reaction; emotional experience depends on the organism's interpretation and evaluation of the situation (cf., Arnold, 1970). All of these theories are probably partially valid, and none of them absolutely so. In short, emotion is an intangible which is difficult to come to terms with in an applied sense.

Investigators are well aware of the ambiguity of the concept, and Davitz (1970) has even gone so far as to compile a dictionary and grammar of emotion. Nevertheless, it is generally accepted that emotion is an introspective, idiosyncratic, mentalistic phenomenon (Arnold, 1968). As such, it is difficult for an investigator to establish unequivocally whether or not any given speaker is indeed experiencing, expressing, or conveying a specific emotion in an experimental situation.

In some studies investigating the vocal communication of emotion, affective state is defined by the intent of the speaker (Davitz & Davitz, 1959a, b). Insofar as the communication of emotion is a particularly private and subjective experience, such a definition is a viable one. Other studies, however, have demanded a more stringent operational definition which involves the rating of speech samples by adult judges (Beldoch, 1964; Dimitrovsky, 1964). In these investigations, a sample is not thought to express a particular emotion solely on the basis of speaker's intent, but also because raters agree that the emotion in question is indeed being transmitted. There is a definite external criterion. Despite the ambiguity of the concept, then, emotion contained within speech samples can be defined operationally by speakers' intent, and/or by judges' ratings.

#### VOCAL COMMUNICATION OF EMOTION

A communication between two or more individuals involves much more than the words actually spoken (cf., Starkweather, 1956, 1961). Aside from nonverbal "gestures," a speaker's tone of voice and general vocal qualities affect a listener's perception of what is said. Although we do not commonly attend only to the vocal portion of a communication, it has been suggested that emotion can be reliably communicated by tonal variation and voice quality alone (Davitz, 1964; Davitz & Davitz, 1961; Duncan, 1969; Kramer, 1963; Mahl & Schulze, 1964; Starkweather, 1956, 1961). It appears that in terms of communicating emotion, "how" something is said may be just as (if not more) important than "what" is said.

In normal, everyday conversation, verbal and vocal cues are transmitted and received simultaneously. However, it is possible to separate, at least partially, the vocal and verbal aspects of speech.



### The Verbal Channel

The verbal channel is composed of sound patterns resulting in morphemes, syllables, words, phrases, sentences, and other linguistic forms conveying the semantic meaning of speech. If the words spoken in a particular message are simply written down, the verbal aspects of the communication have, in effect, been isolated. In fact, the verbal information contained in such written messages is what is most commonly examined by traditional content analytic techniques (cf., Holsti, 1969).

### The Vocal Channel

The features of the voice itself, irrespective of the semantic meaning of the words spoken, are thought to come in on what is called the vocal channel. In other words, the term "vocal channel" refers specifically to the tonal qualities of speech. Soskin and Kauffman (1961) were among the first to suggest that this channel carried potential affective information. They felt that the vocal channel was the "carrier" upon which verbal content is superimposed, and that the cues to emotion supposedly reside within this "carrier."

### JUDGEMENT OF THE NONVERBAL ASPECTS OF SPEECH BY ADULTS

Isolating the vocal aspects of speech from the articulated sound patterns is much more difficult than isolating semantic meaning. Simulating human tones through the use of mechanical devices has been attempted, but the necessary equipment is not readily available. Also, some investigators contend that such devices are not yet sophisticated enough to accurately simulate human tones (Soskin & Kauffman, 1961). It is also possible to utilize foreign speech samples in an effort to isolate the vocal channel.

Unfortunately, foreign speech samples are sometimes difficult to obtain, and inflectional variations from language to language may confound results (Soskin & Kauffman, 1961).

#### Recitation of Nonmeaningful Material

Davitz and Davitz (1959a) proposed that emotion could be reliably communicated by nonverbal vocal speech. In order to test this hypothesis, they recorded expressions of emotion by speakers reciting the alphabet. The speakers were instructed to recite portions of the alphabet, first in a "normal" manner, and then simulating ten various emotions (specifically: love, anger, fear, happiness, jealousy, nervousness, pride, sadness, satisfaction, and sympathy). The recordings were played in random order to graduate student judges, who were required to attempt to identify the emotions expressed. The judges made each identification by choosing one of ten adjectives from a check list. It was found, using a binomial estimate of probability, that the possibility was less than one in a hundred that the number of correct identifications made by these judges could have been obtained by chance alone. Hence, the authors concluded that emotions can indeed be transmitted via the vocal channel alone.

A number of similar studies, involving either the recitation of letters or numbers, have also shown that emotion can be communicated accurately by the nonverbal aspects of speech (Davitz & Davitz, 1959b; Dussenberry & Knower, 1939; Knower, 1941, 1945; Pfaff, 1953; Thompson & Bradway, 1950).

There is a subtle problem inherent in this type of research due to the fact that speakers are not allowed to say words. For most humans, words have been inextricably tied to expression of emotion since early childhood. It is likely that it would be difficult and unnatural for speakers to

communicate emotion vocally without using words. Therefore, any attempt to manipulate speakers so that they are not permitted to use words may be somewhat artificial.

#### Recitation of Standard Content

In other studies, the verbal content is held constant by having speakers recite a standard paragraph that is considered to be emotionally neutral. All speakers try to express different emotions by reciting the same neutral verbal content; the emotional feeling communicated depends on vocal rather than verbal cues (Beldoch, 1964; Fairbanks & Pronovost, 1939; Pollack, Rubenstein, & Horowitz, 1960). However, this procedure still depends on speakers to recite specific standard material, thereby precluding the possibility of studying unacted real-life conversations.

#### Filtered Speech

One approach that obviates this problem is to employ an electronic filter to remove the verbal content of speech while leaving most of the emotional cues carried by the vocal characteristics (Starkweather, 1956; Soskin & Kauffman, 1961). When the filtering method is used, it becomes possible to consider samples of speech other than those generated in experimental situations. In fact, investigation can proceed in an unobtrusive or "nonreactive" fashion (cf., Webb, Campbell, Schwartz, & Sechrest, 1966). Voice samples, including those recorded without the speakers' knowledge, can be filtered and played to subjects.

In terms of theoretical rationale for using a filter, the notion is that vowels, which carry most of the tonal qualities in speech, are low frequency sounds. However, consonants, which convey most of the semantic

meaning, are sounds of high frequencies (Soskin & Kauffman, 1961). Essentially, filtering out the high frequency consonant sounds leaves behind the vocal channel -- the tonal characteristics of voice comprised chiefly of vowel sounds -- while removing most of the verbal meaning. The assumption is made that the tonal qualities, left behind when the high frequencies are removed via the filter device, retain the majority of the emotional cues in speech.

Vocal expressions filtered in this manner become garbled; the speech is reduced to a kind of mumble as if heard through a wall (Starkweather, 1956). Soskin and Kauffman (1961) refer to a number of word recognition studies of filtered excerpts which indicate that listeners can only identify two or three percent of the words spoken, and that most of these identified words are merely conjunctions, prepositions, and articles. In addition, these investigators point out that in filtered samples of 50-150 words in length, this level of accuracy would be attained by guessing alone. In short, filtering appears to be a viable technique for removing the semantic meaning from vocal expressions of human speech. Another method, the compression of speech to some multiple of its normal rate, has also been relatively effective in this regard (Goldhaber, 1970; Goldhaber & Weaver, 1968; Orr, 1968; Orr, Friedman, & Williams, 1965).

Utilizing filters that passed frequencies from 100-450 or 150-550 cycles per second, Soskin and Kauffman (1961) found that adult judges agreed upon the emotional meaning of filtered speech samples at a level significantly above chance. Starkweather (1956), also using filtered material, found that adults were able to identify and separate submissive and aggressive speakers when listening to either normal or filtered speech. Subjects judging the

verbal content alone, as recorded on typescripts, were unable to make this differentiation. Interestingly enough, subjects were more efficient in making the submissive-aggressive distinction with filtered than with normal speech. It seems that the presence of verbal content may even constitute some form of distraction for adults making emotional judgements.

Even though speech can be broken down into components, neither the vocal nor verbal aspects of speech should be thought of as being independent. It would be naive to suppose that speech is composed of two mutually exclusive channels. There is an implicit overlap. Therefore, when speech is filtered, a great deal of the tonal quality is probably lost along with the high frequency sounds (Starkweather, 1956). And although tonal variation may account for the majority of emotional cues, the words actually said are no doubt also very important.

Nevertheless, it has been demonstrated clearly that a great deal of emotional information is contained within filtered speech (Soskin & Kauffman, 1961; Starkweather, 1956). This fact, along with the need to examine real-life voice samples, provides ample justification for employing the filtering technique.

#### JUDGEMENT OF THE NONVERBAL ASPECTS OF SPEECH BY CHILDREN

It is evident that adults as a group can recognize the emotional content of speech, even though the extent of this ability seems to be extremely idiosyncratic and variable (cf., Davitz, 1964). However, as Nash (1970) points out, there are relatively few studies assessing the development of emotional sensitivity to vocal expression of emotion.

The earliest, and for a long period, the only experimental study of the development of emotional sensitivity was performed by Gates (1927). Using

recordings of adult speakers reciting the alphabet in an effort to express each of nine emotions, she found that there was a steady increase with age in children's ability to make the correct identifications. Only very tentative conclusions can be drawn from this early study, however, since a short abstract is all that has ever been published and the original data are unavailable.

Almost four decades later, Dimitrovsky (1964) reopened the question of whether the ability to identify emotion in vocal speech improves with age. In her recorded voice samples, chosen from among a larger number compiled by Beldoch (1964), male and female adult speakers recited a standard paragraph while attempting to communicate happiness, sadness, love, and anger. Since all speakers recited the same standard paragraph, considered to be emotionally neutral, expression of emotion was dependent on vocal cues rather than on the words spoken. The different expressions of this standard paragraph were rated by adult judges. For every speech sample accepted for use in the study, the number of judges who identified the item correctly had to exceed that which would have been expected by chance at the .01 level.

Dimitrovsky (1964) then had children from five to twelve years of age attempt to identify the emotion expressed in the speech samples. Children's judgements were scored correct or incorrect solely on the basis of the adult ratings. Dimitrovsky's (1964) results paralleled Gates' (1927) findings, in that there was a progressive and significant increase with chronological age in children's ability to judge emotion communicated by the vocal expressions. That is, with increasing age, children's judgement of emotional state approached more nearly the adult interpretation.

Additionally, Dimitrovsky (1964) demonstrated that girls are more

sensitive to the emotional content of adult voice samples than are boys. This finding was not entirely unexpected, since girls are superior to boys in the acquisition of normal speech (McCarthy, 1953, 1954; Sampson, 1959; Young, 1941). Dimitrovsky (1964) points out that in our culture the female is stereotyped as being more emotionally sensitive than the male. Thus, the superior accuracy of girls in judging the affective content of speech may simply reflect the learning of cultural expectations.

Dimitrovsky (1964) distinguished between emotional states on the basis of valence and activity. With respect to the valence dimension, each emotion was defined as being either positive or negative. Similarly, in terms of the activity dimension, each emotional state was considered to be either active (exuberant) or passive (subdued). Dimitrovsky (1964) found that the negative emotions of sadness and anger are judged correctly by children more often than are the positive emotional states of happiness and love. However, errors tend to be characterized by similarity along the activity dimension, rather than valence. That is, the active emotions of anger and happiness are frequently mistaken for one another, as are the passive emotions of sadness and love.

Furthermore, Dimitrovsky (1964) noted that a response bias seems to exist along both the valence and activity dimensions. In general, children tend to emit more negative than positive responses. Similarly, they tend to respond with passive judgements more often than with active choices.

Davitz and Davitz (1959a) have shown that adults, as well as children, correctly identify expressions of negative emotions more often than positive ones. However, although adults judge the negative emotions more accurately, they do not seem to emit a greater frequency of negative than positive

responses when making emotional judgements (Davitz, 1964). On the other hand, according to Dimitrovsky (1964), children not only make more correct identifications of the negative emotional states, but respond with a greater number of negative judgements as well. Consequently, she concludes that children tend to be characteristically more negative in their judgements of vocal expression than adults.

Fenster (1967), using a similar method, also found that the ability to identify emotional content in neutral vocal expressions increases with age. Using adult and child speakers, he too noted a consistent tendency for adult and child listeners to judge negative vocal expressions more accurately than positive speech samples. However, unlike Dimitrovsky (1964), he found no tendency for the children (or adults) to respond with a preponderance of negative judgements.

Although it has been shown that grade five children can identify the emotional content in speech compressed so as to render the words unintelligible (Boyle, 1970), the development of sensitivity to filtered speech has not been explored. Most investigators to date have had speakers recite neutral verbal content or have utilized the filter device (cf., Starkweather, 1961). A direct comparison of the two methods has not yet been undertaken with either children or adults.

#### INCONGRUENT COMMUNICATIONS

It is possible to evaluate the relative contribution of the verbal and vocal channels of speech when both are present simultaneously. For example, incongruent communications, in which the tone of voice is at variance with the semantic meaning of the words actually spoken, have been employed to investigate whether subjects tend to give more weight to the vocal or to the



verbal channel when judging the emotional content of speech (Bugental, Kaswan, Love, & Fox, 1970; Mehrabian & Wiener, 1967).

#### Adult Responses to Incongruent Communications

In their attempt to discover whether one channel was given more weight than the other by adults making emotional judgements, Mehrabian and Wiener (1967) selected single words to communicate positive, neutral, and negative affective content. Speakers were then required to recite all single words (positive, neutral, and negative) in positive, neutral, and negative tones of voice. Speech samples were recorded and played to undergraduate university students, who rated the emotional content. The results indicated that the vocal channel was dominant. To illustrate, if a positive word was said in a negative tone, the message was generally judged as being negative. This dominance of vocal over verbal cues in incongruent speech has also been noted by other investigators (Fujimoto, 1972).

Nonverbal signals are not only delivered vocally. If the speaker is visible, other nonverbal cues can be transmitted on what is defined broadly as the visual channel. Indeed, certain studies have shown that emotional state can be reliably transmitted via facial expression (cf., Ekman, Friesen, & Tomkins, 1971). Although the present research is confined specifically to the two channels of speech (vocal and verbal), it should be kept in mind that these speech channels often interact with facial expression and other "visual" cues (e.g., proximity, posture, gestures, bodily contact, and so on) in the communication of emotion.

Another study with university undergraduates used a procedure similar to that employed by Mehrabian and Wiener (1967), except that three tones of voice were combined with degrees of attitude communicated by facial expression

(Mehrabian & Ferris, 1967). The results of this study revealed that university students place slightly more weight on facial expression than on tone of voice when judging emotion. There are also other indications that visual cues have more impact than vocal ones on adults and university students (Bugental, Kaswan, Love, & Fox, 1970; Levitt, 1962, 1964).

#### Developmental Differences in Response to Incongruent Communications

It would be of interest to study incongruency from a developmental perspective in order to gain some understanding of how childhood history may contribute to learning to perceive emotion in a normative way, and, possibly, to explain causes of "disturbed" perception and behaviour. Research indicates that adults and university students give more weight to vocal cues than to the verbal channel when judging the affective content of speech (Fujimoto, 1972; Mehrabian & Wiener, 1967; Starkweather, 1956). Research with children, however, has yielded some inconsistencies which suggest the need for more basic research on developmental trends in the perception of emotion transmitted via the vocal channel alone. A discussion of the available research is warranted at this point in order to clarify these issues.

Bugental, Kaswan, and Love (1970) prepared videotapes of acted communications containing incongruence between verbal (typescript of semantic content), vocal (tone), and visual (facial expression) channels. The emotion communicated on each channel was either positive (friendliness, approval, or consideration) or negative (unfriendliness, disapproval, or inconsiderateness). All possible positive-negative content combinations of the three channels (i.e., eight possible combinations) were acted out by two male and two female speakers.

The verbal content, as recorded on typescripts, was rated by at least 25 judges. They all had to be in agreement as to whether each verbal message was positive or negative.

A panel of six judges rated both the vocal and visual information to determine if the intended content was communicated on each channel. To rate the vocal channel, the judges heard the message with the intelligible verbal content removed by a filtering device, and did not view the visual part of the communication. However, when rating the visual channel, the judges were shown the speakers' facial expressions on a television monitor, without hearing the sound portion of the communication.

The videotapes were played to children aged five through twelve, as well as to parents. It was found that "joking" messages (negative script, negative tone, and positive picture) were interpreted significantly more negatively by children than by adults. This age difference was most pronounced when the incongruent communications were enacted by females. Adults appeared to conclude that critical statements said with a smile were not meant to be entirely negative. However, children tended to resolve conflicting messages by accepting the negative component while ignoring the positive aspects of the communication, especially if the speaker was female. These results suggest that communications intended to be joking or sarcastic (i.e., not necessarily negative) may be interpreted negatively by children.

The finding that children reacted most negatively to incongruent communication from female speakers has led to additional research. Bugental, Love, and Gianetto (1971) showed that children interpreted statements from an adult male as being friendlier if he was smiling. However, there was no difference in children's judgements of a message acted by a female when she

was smiling from when she was not. And it seems that when a woman's smile is the only positive aspect of a communication, children judge the message to be negative, while adults tend to regard it as neutral (Bugental, Kaswan, & Love, 1970).

In another related study, it was demonstrated that young children placed less importance on the visual channel (facial expression) when making emotional judgements than did adults (Bugental, Kaswan, Love, & Fox, 1970). This age trend was again strongest for the perception of female messages. What a person says and how he or she says it seem to be more important to children than how that person looked while speaking.

Olsen (1967) ignored the visual component, and looked specifically at the roles of the verbal and vocal channels of speech in incongruent messages. He played incongruent communications, in which the emotional meaning communicated through tone of voice was contrary to that conveyed by semantic content, to children 5, 7, 9, 11, and 13 years of age. In general, the children attended to the verbal portion of the message and discounted the tonal information. Although adults seem to put more emphasis on tone than words when judging expressions of emotion (Fujimoto, 1972; Mehrabian & Wiener, 1967; Starkweather, 1956), Olsen's (1967) results suggest that children respond in just the opposite manner. However, since the preference for the verbal components of the message declined somewhat with age, Olsen (1967) speculated that, with increasing age, children reveal a growing awareness of the subtle interaction between tone and semantic content.

#### Incongruent Communication and Pathology

An excess of incongruent communication from parent to child is thought by some to be involved in the etiology of abnormal behaviour. Specifically,

the double-bind hypothesis has been proposed to account for the origin of schizophrenia (Bateson, Jackson, Haley, & Weakland, 1956). In brief, the theory is that a parent, usually the mother, says one thing to the child with words, while communicating something quite different on the nonverbal level. The contradictions in the message are hidden, inhibited, or denied in some manner, and any attempt by the child to point out the conflicting nature of the communication simply leads the mother to deny any incongruity.

Supposedly, a child is placed in an ambivalent and confusing situation, referred to as a double-bind, when he/she is subjected to such incongruent communication. According to Bateson, Jackson, Haley, and Weakland (1956), continued exposure to ambiguous messages of this sort can cause the child to develop schizophrenic behaviour patterns.

The double-bind hypothesis has generated a good deal of unsuccessful research (cf., Schuchman, 1967). Recently, however, fruitful data have been produced, and ingenious attempts have been made to relate psychopathology and parental use of incongruent communication (Beakel, 1971; Beakel & Mehrabian, 1969; Bugental, Love, Kaswan, & April, 1971; Olsen, 1967).

In a study of major importance, Bugental, Love, Kaswan, and April (1971) prepared videotapes of parent-child communication in "normal" families and in families containing a "disturbed" child. For the purposes of their particular investigation, a "disturbed" child was one referred by school authorities as exhibiting emotional or behavioural difficulties. Adult judges rated these videotaped communications with respect to verbal, vocal, and visual content. It was found that a significantly greater proportion of mothers of "disturbed" children emitted conflicting, incongruent communications than did mothers of the "normal" children. Although

this finding is consistent with the double-bind hypothesis, it suggests that incongruent communication on the part of the mother could be central to the causation of pathology in general, and not limited specifically to schizophrenia.

#### STATEMENT OF THE RESEARCH PROBLEM

It would seem that emotion in speech is communicated in large part on the vocal channel. For example, it is well established that adults in general are able to identify accurately the emotional content of speech containing only vocal cues (Davitz, 1964; Davitz & Davitz, 1961; Duncan, 1969; Kramer, 1963; Mahl & Schulze, 1964; Starkweather, 1961). Starkweather (1956) reports that the vocal channel is more effective in communicating emotion to adult subjects than is the verbal channel. Additionally, when verbal and vocal aspects of speech are incongruent, the vocal symbols seem to carry the most power for university students (Mehrabian & Wiener, 1967; Fujimoto, 1972).

It has been found that ability to judge the emotional content of voice samples improves with age (Dimitrovsky, 1964; Fenster, 1967; Gates, 1927). It is not clear, however, whether children are more negative in their judgement of vocal expressions than adults. Dimitrovsky (1964) reports that children emit a greater frequency of negative responses when making emotional judgements than do adults, while Fenster (1967) contends that there is no tendency on the part of children or adults to respond with a preponderance of either positive or negative judgements.

The research on children's response to incongruity is also somewhat contradictory. Olsen (1967) asserts that children faced with incongruent communication tend to accept the verbal message and ignore the tone of voice.

On the other hand, Bugental, Kaswan, and Love (1970) report that children resolve incongruent messages by accepting the negative elements of the communication while ignoring those that are positive.

To iterate, only a few investigations to date have dealt directly with the question of age differences in sensitivity to vocal expression of emotion (Dimitrovsky, 1964; Fenster, 1967; Gates, 1927). As implied previously, although more recent studies have examined the effects of incongruent communications on children of various ages (Bugental, Kaswan, & Love, 1970; Bugental, Love, Kaswan, & April, 1971; Olsen, 1967), such research may be premature in that developmental trends have not yet been explored adequately.

The present study was an attempt to confirm that judgement of the emotional content of standard paragraphs becomes more "adult-like" with increasing age. Another major aim was to gather more information regarding the controversy over whether young children are more negative in their emotional judgements than are older subjects. An effort was also made to examine differences in sensitivity to individual emotional states, and to specify the type and frequency of judgemental errors.

Another goal of the current study was to assess children's ability to judge filtered speech samples in which there were no words to "hook onto." A comparison of children's judgements of nonfiltered and filtered voice samples was designed to determine whether children at various age levels are aided or distracted by words when they attempt to identify emotional cues in speech.

Rather than varying both the sex of the communicator and the sex of the receiver, the research in question was arbitrarily restricted to one communication pattern in which voice samples from female speakers were presented to

male children at different age levels. A great deal of attention has been devoted to the role of the mother in the etiology of psychopathology (cf., Clausen & Kohn, 1960). As discussed previously, while mothers of "disturbed" children produce a significantly greater number of conflicting messages than do mothers of "normal" children, no such difference is evident among the fathers (Bugental, Love, Kaswan, & April, 1971). It has also been demonstrated that incongruent messages are judged more negatively by children if the speaker is female rather than male (Bugental, Kaswan, & Love, 1970). Consequently, it appeared as if the role of the female communicator was of considerable importance and worthy of direct examination.

#### Judgement of Standard Nonfiltered Vocal Expressions

In large part, the present study was patterned after Dimitrovsky's (1964) investigation. In light of her findings, the following major hypothesis was put forward for the nonfiltered condition:

(1) there would be a progressive increase with age in children's ability to identify the emotional content of a nonfiltered standard paragraph.

More specifically, the following secondary hypotheses were also put forward for the nonfiltered treatment:

(a) children would correctly identify the negative emotions (sadness and anger) more often than the positive emotional states (happiness and love).

(b) errors would occur most frequently along the activity dimension such that children would tend to confuse the active emotions (anger and happiness), and, similarly, would also confuse the passive emotional states (sadness and love).



(c) regardless of accuracy, children would respond with negative judgments more frequently than with positive ones.

(d) regardless of accuracy, children would respond with passive judgments more frequently than with active ones.

#### Judgement of Filtered Vocal Expressions

Since filtered speech is only part of a whole, the most general parsimonious assumption was that, although ability to identify filtered vocal expressions might improve as a function of age, children at all age levels would identify filtered speech samples less accurately than nonfiltered ones.

This being the case, it was still reasonable to suppose that the secondary hypotheses (a) through (d), entertained for the nonfiltered condition, would hold for the filtered treatment as well.

However, no information is available on which to base predictions regarding children's sensitivity to filtered speech. It may be, as Starkweather (1956) has found with adults, that the presence of verbal content constitutes some form of distraction for persons making judgements of emotional state. Taking this perspective, it was possible that children judging filtered vocal expressions would identify emotion more accurately than those listening to nonfiltered samples. Further speculation regarding potential outcomes is unwarranted here, however, without reference to specific data.

## METHOD

### SUBJECTS

One hundred and twenty male children from two schools in the Seven Oaks School Division in Winnipeg, Manitoba, Canada, served as subjects. Twenty boys at each age from 6 through 11 took part. No subject at any age level was within two months of a birthday at the time the experiment was run. Also, no child considered to be "slow" or "emotionally disturbed" by the school principal, or whose IQ was known to be above or below the normal range, was included in the sample.

Half of the boys at each age level (i.e., ten subjects) were selected from Seven Oaks School; the other half from Governor Semple School. Both schools are located in West Kildonan, Winnipeg, in a middle class, residential, suburban area. Half the subjects from each age level in each school were assigned at random to either the nonfiltered or filtered condition (i.e., five boys from Seven Oaks School and five from Governor Semple School made up each group).

Twenty "adult," undergraduate male students at the University of Manitoba, ranging in age from 20 to 22, served as judges. They participated in the study in partial fulfillment of the requirement for the introductory psychology course. These judges were assigned at random to either the nonfiltered or filtered treatment groups (i.e., ten per experimental condition).

### APPARATUS

The original speech samples were recorded on a Sony Stereo Tape recorder, model TC-230W. Another set of speech samples was prepared and rendered

unintelligible through the use of an electronic filter. The filter passed frequencies from 100-450 cycles per second with a 60-decibel per octave attenuation at the upper level. A second Sony Stereo Tape recorder was used during the filtering procedure.

The filtering device accepted shielded cables from the tape recorder playing the original material, and accepted another pair of shielded cables that connected the unit to the second tape deck. Once the filter was switched on, the original speech samples were played on the first tape recorder, and the filtered material was recorded on the second tape deck.

Both nonfiltered and filtered materials were presented to subjects via external speakers (Sony Speaker System model SS-23).

#### STICK FIGURE DRAWINGS

In many studies of emotional communication, adult subjects are asked to choose the one particular word from a check list of adjectives that they feel best describes the emotional content of each vocal expression (Davitz & Davitz, 1959a, 1959b; Soskin & Kauffman, 1961). However, when dealing with children of different ages, it is not feasible to employ an experimental task involving either reading or writing. Age differences in these abilities may confound results. Consequently, in an attempt to control for this problem, all subjects in this investigation were asked to make their identification of each speech sample by pointing to one of the four stick figure drawings that represented the four emotional states under consideration. Each picture was drawn on a separate card.

The stick figure drawings, which were identical to those used by Dimitrovsky (1964), are pictured below.

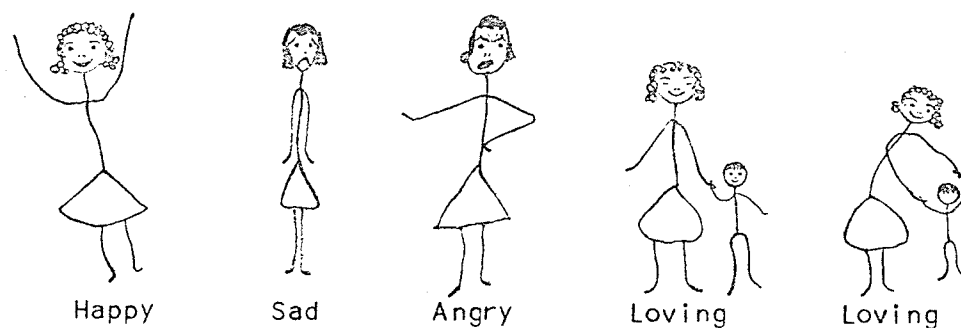


Figure 1. Stick figures representing the four emotional states.

Dimitrovsky (1964) utilized two drawings representing love because she felt that this emotion was the most difficult to portray pictorially. The actual sizes of the figures presented in the current study are illustrated in Appendix A.

#### RECORDING OF VOCAL EXPRESSIONS

Six amateur actresses between 22 and 26 years of age recited a standard three sentence paragraph with the intention of vocally expressing happiness, sadness, love, and anger. Since each of the six speakers expressed each of the four emotional meanings, there were a total of 24 speech samples recorded. The order in which the four emotions were acted was randomized for each speaker. However, it was stipulated that no more than two actresses could be asked to communicate the same emotion first. In addition, the order in which the emotions were expressed had to be different for all speakers. The order in which each of the six actresses was instructed to express the four emotions is detailed in Appendix B.

Each actress was given a typescript of the standard paragraph originated by Beldoch (1964) and later used by Dimitrovsky (1964). The paragraph read:

I'm going out now. I won't be back all afternoon.

If anyone calls, just tell them I'm not here.

Every speaker was asked to read the paragraph aloud twice in order to familiarize herself with the content. Also, each actress was asked to imagine a situation in which she felt extremely happy, sad, loving, and angry prior to attempting to simulate the emotion in question.

After the 24 speech samples of standard verbal content had been recorded, they were arranged in random order. There were again some restrictions to the randomization procedure. No emotion could appear more than twice in a row, and no more than two speech samples from any one speaker could appear in succession. The random arrangement of the 24 nonfiltered voice samples is outlined in Appendix C.

A comparable filtered version was then prepared by filtering this ordering of the original vocal expressions.

#### RATING OF VOCAL EXPRESSIONS

The nonfiltered vocal expressions were played to half of the adult subjects, while the filtered speech samples were played to the remaining ten. A description of the instructions (designed primarily for use with younger subjects), given to both children and adults, follows in the Procedure. The subjects in both adult groups attempted to identify the emotion they thought the speaker was experiencing. They recorded each judgement by pointing to one of the stick figures.

Each of the 24 nonfiltered speech samples was rated as transmitting the emotion intended by the speaker more often than it was judged to communicate any different emotional state. That is, for each speech sample, more adult judges thought that the speaker was communicating the emotion

she was trying to than thought she was communicating some other emotional state. Thus, the adult ratings indicated that the actresses were able to simulate the four emotions quite accurately.

Also, for all nonfiltered speech samples of a particular emotional category, the ratings were in general agreement with the actresses' intent. For example, since each speaker communicated happiness, there were six "happy" speech samples. All ten adults judged each sample, for a total of 60 judgements. These judgements totalled well over a majority (i.e., more than 31/60) in the direction of speakers' intent for each of the four emotional states (Appendix D). A binomial estimate of probability revealed that the number of correct identifications of each emotional state was significantly better than that expected by chance ( $p < .01$ ).

Therefore, on the basis of these ratings, it was concluded that in this particular instance the actresses could indeed simulate the emotion they intended with reasonable accuracy. In this specific situation, speakers' intent seemed to constitute a satisfactory definition of the emotional content of each speech sample.

When scored against speakers' intent, the filtered speech samples were not judged as accurately as the nonfiltered. That is, the adults judging the filtered samples did significantly poorer than those judging the nonfiltered samples ( $F = 39.36$ ,  $p < .01$ ). Still, their judgements of the filtered expressions of each separate affective state totalled more than a majority for every emotion except happiness (Appendix D). In this case, 29 of 60 "happy" samples were judged in accordance with speakers' intent, a level better than chance at the .01 level.

## PROCEDURE

The methodology of the present study paralleled Dimitrovsky's (1964) procedure as closely as possible. All subjects in every age group were seen individually. In order to accustom them to hearing taped speech samples, subjects in the nonfiltered condition listened to recordings of an adult male, an adult female, a male child, and a female child saying: "I'm going out now." Subjects were simply asked to identify whether the speaker was a man, woman, boy, or girl. Children being familiarized with filtered speech listened to the same four speech samples after filtering, and were required to make the same identifications. These subjects in the filtered treatment were told that they were going to hear the voices of people speaking from "far away," and that they would not be able to make out the words clearly. The voice samples were very obvious, and no subject at any age level had any difficulty making the identifications with either nonfiltered or filtered speech.

Next, subjects of all ages in both conditions were shown two cards depicting love, and asked to choose the drawing in which the woman looked as if she loved the child more. By far the greatest number of subjects (i.e., 109/120 children and 19/20 adults) chose the drawing in which the stick figure woman had her arms around the boy, rather than the one in which the woman and child were holding hands. The drawing of love chosen by each subject was then placed in front of him, along with the representations of happiness, sadness, and anger. The drawings were arranged in random order for each subject.

All subjects in each group were then trained to associate each of the four pictorial representations with the appropriate corresponding emotional

state. In other words, they were required to point to the happy woman, the sad woman, the loving woman, and the angry woman. This procedure was repeated until a criterion of two correct identifications of all four drawings was reached. No subject had any difficulty with this task.

Children in the nonfiltered condition were then instructed:

You are going to hear different ladies say the same thing in different ways. Some will sound happy, some sad, some loving, and some angry. Listen to each one and point to the lady it sounds like. Does it sound like the happy lady, the sad lady, the loving lady, or the angry lady?

The nonfiltered recordings were then played to the appropriate subjects, who made their 24 choices. After hearing each speech sample, the child had a ten second interval to point to the stick figure drawing that he thought best represented that particular vocal expression. In the rare instances when a child took longer than ten seconds to respond, the recorder was stopped until he made his judgement.

Children in the filtered group were reminded that the voices they were going to listen to would be coming from "far away," and that they would not be able to hear the words clearly. Aside from this difference, the instructions and procedure in the two conditions were identical.

A more complete description of the instructions is presented in Appendix E.



## RESULTS

## AGE DIFFERENCES IN ACCURACY OF EMOTIONAL PERCEPTION IN NONFILTERED AND FILTERED TREATMENTS

For each subject, the number of correct identifications of each of the four emotions was computed. An analysis of variance was conducted on these data, using a design containing six levels of age, two levels of filtering, and four levels of emotion. The analysis, summarized in Table 1, revealed an age difference in ability to identify the emotional content of the speech samples. As well as the significant age effect ( $F = 18.33, p < .01$ ), the analysis indicated a significant difference between the nonfiltered and filtered conditions ( $F = 153.24, p < .01$ ). Further, the analysis revealed a significant difference in level of recognition for the different emotions ( $F = 57.64, p < .01$ ). There was also evidence of interaction effects between treatment (i.e., nonfiltered or filtered) and emotion ( $F = 9.49, p < .01$ ).

The mean scores for all age levels in the nonfiltered and filtered conditions are presented in Figure 2 (and in tabular form, along with the standard deviations, in Appendix F). An inspection of the means shows that the subjects judging the nonfiltered vocal expressions scored consistently higher at all ages than did those listening to filtered samples. In both conditions, however, there was a gradual and progressive increase in number of correct identifications made from every younger to every older age group. As Figure 2 indicates, the curve for the nonfiltered condition rises gradually from a low of 10.4 for the six-year-olds to a high of 15.0 for the eleven-year-olds. A test for trend for the nonfiltered treatment revealed that the linear trend between age and mean number of correct identifications was significant ( $F = 46.10, p < .01$ ). The trend for residual

TABLE 1

Analysis of Variance for Mean Number of Correct Identifications  
of Each Emotional State (Scored against Adult Ratings)  
as a Function of Age and Filtering

Source	SS	df	MS	F
Between subjects	237.82	119		
Age	61.39	5	12.28	18.33*
Filter	102.67	1	102.67	153.24*
Age x filter	1.64	5	.33	.49
Subjects within groups	72.12	108	.67	
Within subjects	724.50	360		
Emotion	221.34	3	73.78	57.64*
Age x emotion	31.19	15	2.08	1.63
Filter x emotion	36.46	3	12.15	9.49*
Age x filter x emotion	19.23	15	1.28	1.00
Emotion x subjects within groups	416.28	324	1.28	
Total	962.32			

\* $p < .01$

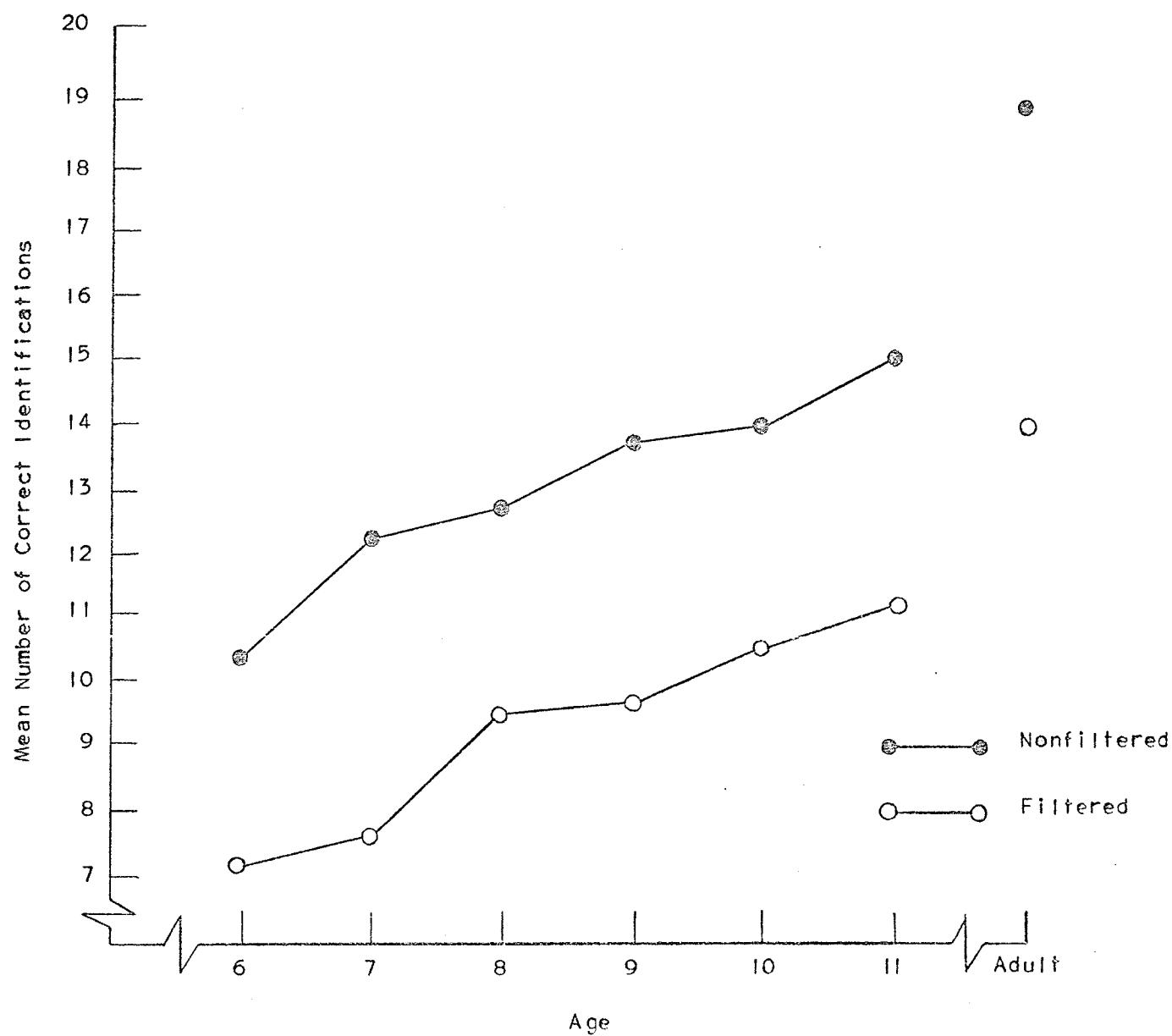


Figure 2. Mean number of correct identifications at successive age levels in nonfiltered and filtered conditions.

components (i.e., quadratic, cubic, and quartic) was not significant ( $F < 1$ ), indicating no significant deviations from linearity. Similarly, the curve for the filtered condition rises slowly but steadily from 7.2 for the six-year-olds to 11.2 for the eleven-year-olds. A test for trend on the filtered data also yielded a significant linear trend ( $F = 43.60$ ,  $p < .01$ ) and no significant deviations from linearity for the residual components ( $F < 1$ ).

The six-year-olds judging the filtered speech scored lower on the average than any other group, with a mean score of only 7.2 correct out of a possible 24. However, a binomial estimate of probability indicated that this performance level was significantly greater than that expected by chance alone ( $p < .05$ ). The performance of every other group in the experiment was therefore also better than chance ( $p < .01$ ).

Since the overall  $F$  was significant for the age, filter, and emotion factors, it was possible to perform post-hoc comparisons. All groups tested were of equal size, and the chief concern was with pairwise differences between means. For these reasons, the Tukey ratio was chosen for the purpose of making certain multiple comparisons. This method has been outlined in detail elsewhere (Kirk, 1969; Winer, 1962).

The Tukey ratio indicated a significant difference at each age level between the mean scores obtained in the nonfiltered and filtered conditions. At every age, from the six through to the eleven-year-olds, subjects listening to nonfiltered speech scored consistently and significantly higher than those hearing filtered samples (i.e., for six-year-olds  $q = 6.18$ ,  $p < .01$ ; for seven-year-olds  $q = 8.66$ ,  $p < .01$ ; for eight-year-olds  $q = 6.18$ ,  $p < .01$ ; for nine-year-olds  $q = 7.73$ ,  $p < .01$ ; for ten-year-olds  $q = 6.72$ ,  $p < .01$ ; and for eleven-year-olds  $q = 7.34$ ,  $p < .01$ ).

In terms of age effects, there was a significant difference between the performance of the youngest (six-year-old) and the oldest (eleven-year-old) children in both the nonfiltered ( $q = 8.89, p < .01$ ) and filtered ( $q = 7.73, p < .01$ ) conditions. Some other comparisons between means within each experimental condition also yielded significance. Within the nonfiltered treatment, there were significant differences between: the six and seven-year-olds ( $q = 3.48, p < .05$ ); the six and eight-year-olds ( $q = 4.44, p < .01$ ); the six and nine-year-olds ( $q = 6.37, p < .01$ ); the six and ten-year-olds ( $q = 6.96, p < .01$ ); the seven and nine-year-olds ( $q = 2.90, p < .05$ ); the seven and ten-year-olds ( $q = 3.48, p < .05$ ); the seven and eleven-year-olds ( $q = 5.41, p < .01$ ); and the eight and eleven-year-olds ( $q = 4.44, p < .01$ ). As an aid to the reader, a graphic display indicating the significant differences in performance between the various age groups in the nonfiltered condition is presented in Table 2.

Within the filtered condition, additional significant differences were found between: the six and eight-year-olds ( $q = 4.44, p < .01$ ); the six and nine-year-olds ( $q = 4.83, p < .01$ ); the six and ten-year-olds ( $q = 6.38, p < .01$ ); the seven and eight-year-olds ( $q = 3.48, p < .05$ ); the seven and nine-year-olds ( $q = 3.86, p < .01$ ); the seven and ten-year-olds ( $q = 5.41, p < .01$ ); the seven and eleven-year-olds ( $q = 6.76, p < .01$ ); the eight and eleven-year-olds ( $q = 3.28, p < .05$ ); and the nine and eleven-year-olds ( $q = 2.90, p < .05$ ). Table 3 shows these significant differences in performance level between the various age groups in the filtered treatment.

The mean scores attained by each age group in each condition for each of the four emotional states are presented in Figure 3 (and in tabular form, along with the standard deviations, in Appendix G). The figure shows that, in the nonfiltered condition, the level of recognition was highest for the

TABLE 2

Significant Differences in Performance between Various Age  
Groups in the Nonfiltered Condition

Age	6	7	8	9	10	11
6		*	**	**	**	**
7				*	*	**
8						**
9						
10						
11						

\*\*  $p < .01$

\*  $p < .05$

TABLE 3  
Significant Differences in Performance between Various Age  
Groups in the Filtered Condition

Age	6	7	8	9	10	11
6			**	**	**	**
7			*	**	**	**
8						*
9						*
10						
11						

\*\*  $p < .01$

\*  $p < .05$

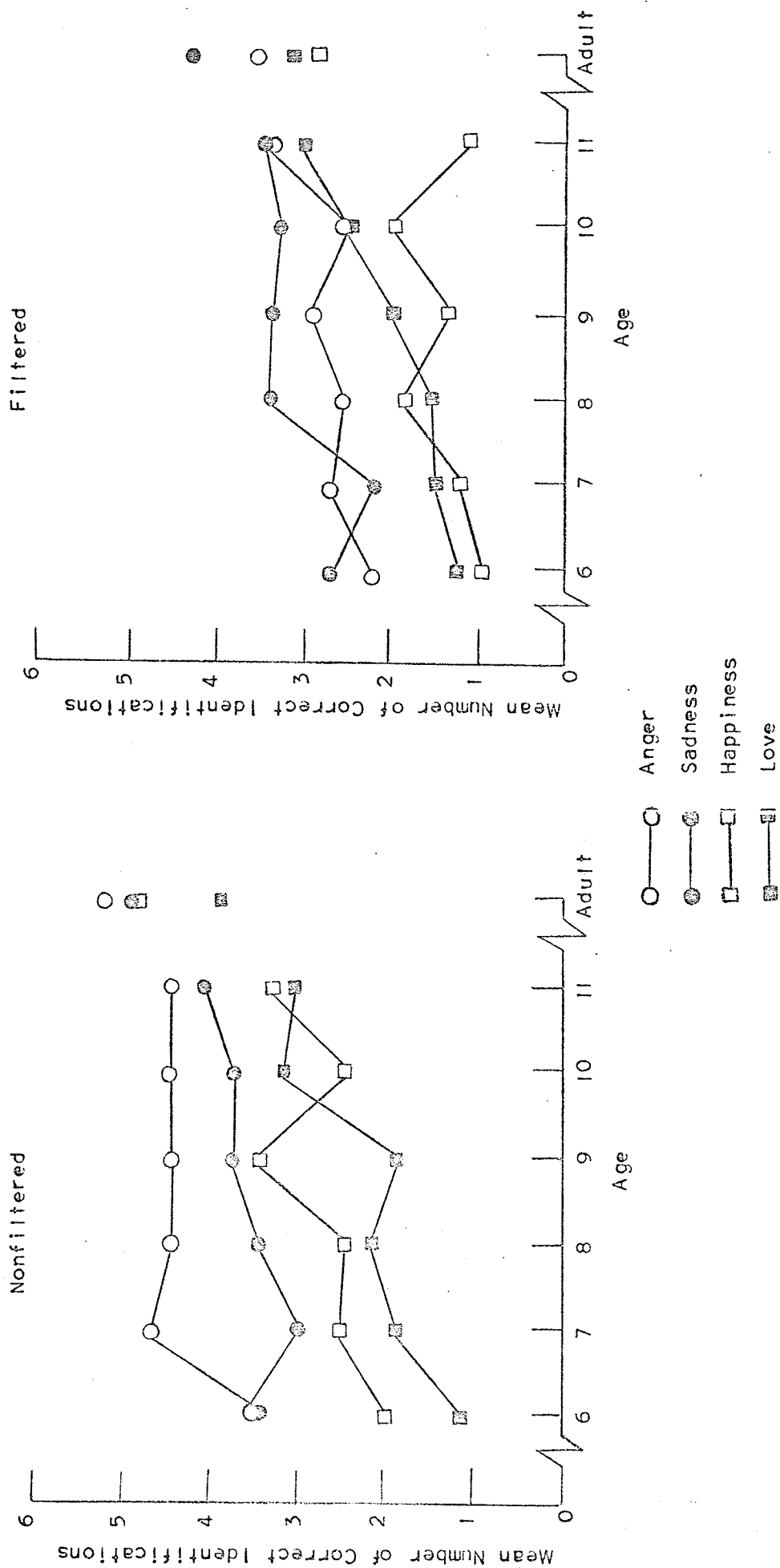


Figure 3. Mean number of correct identifications of nonfiltered and filtered expressions for each category of emotional meaning at successive age levels.



"angry" vocal expressions. "Sad," "happy," and "loving" items followed respectively. Only once, when the ten-year-old group correctly identified "loving" items more often than "happy" ones, was this order of relative difficulty violated. The Tukey ratio revealed that these differences in recognition level were statistically significant: anger was judged more accurately than sadness ( $q = 5.14, p < .01$ ); sadness more accurately than happiness ( $q = 6.16, p < .01$ ); and happiness more accurately than love ( $q = 3.31, p < .05$ ).

As Figure 3 also shows, the level of recognition seemed to shift in the filtered condition. When speech was filtered, "sad" items were judged correct most often, with "angry," "loving," and "happy" items following in that order. The order of relative difficulty was consistently maintained in the filtering condition as well, with the only exceptions being at the seven-year-old level (where "angry" items were judged correctly more often than "sad" ones) and at the eight-year-old level (where "happy" items were judged with more accuracy than "loving" ones). Although the difference in recognition level between sadness and anger did not reach acceptable levels of statistical significance ( $q = 2.28, p > .05$ ), the Tukey ratio indicated that anger was judged significantly more accurately than love ( $q = 5.14, p < .01$ ), and that love, in turn, was judged more accurately than happiness ( $q = 3.65, p < .01$ ).

In terms of valence, the Tukey ratio revealed that the negative emotions (anger and sadness) were judged more accurately than the positive emotional states (happiness and love) in both the nonfiltered ( $q = 10.39, p < .01$ ) and filtered ( $q = 8.11, p < .01$ ) condition. With respect to the activity dimension, the active emotions (anger and happiness) were identified correctly more often than the passive emotions (sadness and love) in the nonfiltered

treatment ( $q = 4.22, p < .01$ ). However, in the filtered condition judgement of the passive emotions improved such that they were identified more accurately than the active emotional states ( $q = 2.97, p < .05$ ).

Aside from the significant age, filter, and emotion effects, the analysis summarized in Table I also revealed a significant filter by emotion interaction. In an effort to explore this interaction further, the mean scores for all subjects (irrespective of age) for each emotional category in both the nonfiltered and filtered conditions are presented in Figure 4 (and in tabular form, along with the standard deviations, in Appendix G). The figure shows that, for each emotional state, subjects in the nonfiltered condition attained higher mean scores than those in the filtered treatment. Further, the differences between the mean scores achieved in the nonfiltered versus the filtered condition were greater for the active emotions of anger ( $4.38 - 2.75 = 1.63$ ) and happiness ( $2.73 - 1.47 = 1.26$ ) than for the passive emotions of sadness ( $3.63 - 3.08 = .55$ ) and love ( $2.25 - 2.00 = .25$ ). The Tukey ratio revealed that these differences were significant for the emotional states of anger ( $q = 11.91, p < .01$ ), happiness ( $q = 9.24, p < .01$ ), and sadness ( $q = 4.01, p < .01$ ), but not for love ( $q = 1.82, p > .05$ ).

The table outlined in Appendix I gives the distribution of responses to each stimulus for the total group of subjects in both the nonfiltered and filtered conditions. The table shows that errors are characterized by similarity along the activity dimension. For example, for both treatment groups, "happy" vocal expressions were mistaken for "angry" more often than for "sad" or "loving." Similarly, "sad" items tended to be mistaken for "loving," "loving" for "sad," and "angry" for "happy." Binomial estimates of probability revealed that errors in both the nonfiltered and filtered conditions occurred along the activity dimension significantly more often than would have been expected by chance alone ( $p < .01$ ).

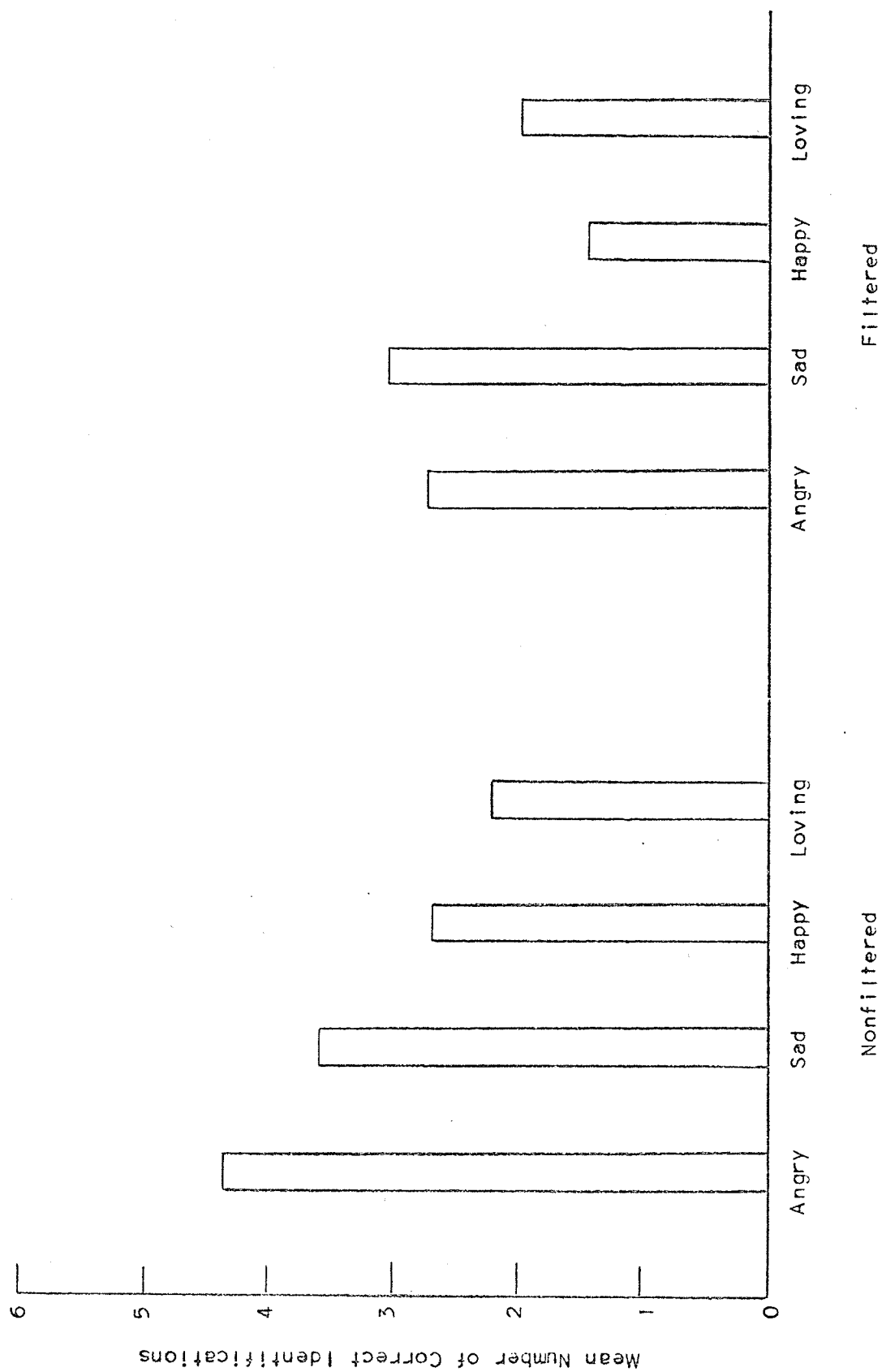


Figure 4. Mean number of correct identifications (across age) for each category of emotional meaning in nonfiltered and filtered conditions.

## RESPONSE BIAS IN NONFILTERED AND FILTERED TREATMENTS

The number of responses emitted by each age group for each category of emotional meaning in the nonfiltered and filtered conditions are presented in tabular form in Appendix J.

### Valence

Considering the valence dimension alone, Figure 5 shows that, for the nonfiltered condition, more negative (angry and sad) than positive (happy and loving) responses were emitted by subjects of all ages. Since an increase in negative responses was reflected by a concomitant decrease in number of positive judgements, and vice versa, the one curve in Figure 5 is the mirror image of the other. A binomial estimate of probability indicated that, overall, the bias in the direction of negativity was significant ( $p < .01$ ). However, although Figure 5 would seem to suggest that the tendency to emit negative responses decreased with age, a one-way analysis of variance (Appendix K) performed on the negative response data revealed that older subjects did not respond significantly less negatively than younger ones ( $F < 1$ ).

Similarly, Figure 6 shows that children of all ages responded more negatively in the filtered condition as well. As in the nonfiltered treatment, a binomial estimate of probability revealed that significantly more negative than positive responses were emitted by subjects overall ( $p < .01$ ). A one-way analysis of variance (Appendix L) also indicated that there were no significant differences in negativity among the various age groups listening to filtered speech ( $F < 1$ ).

### Activity

With respect to the activity dimension, Figure 7 shows that, in the

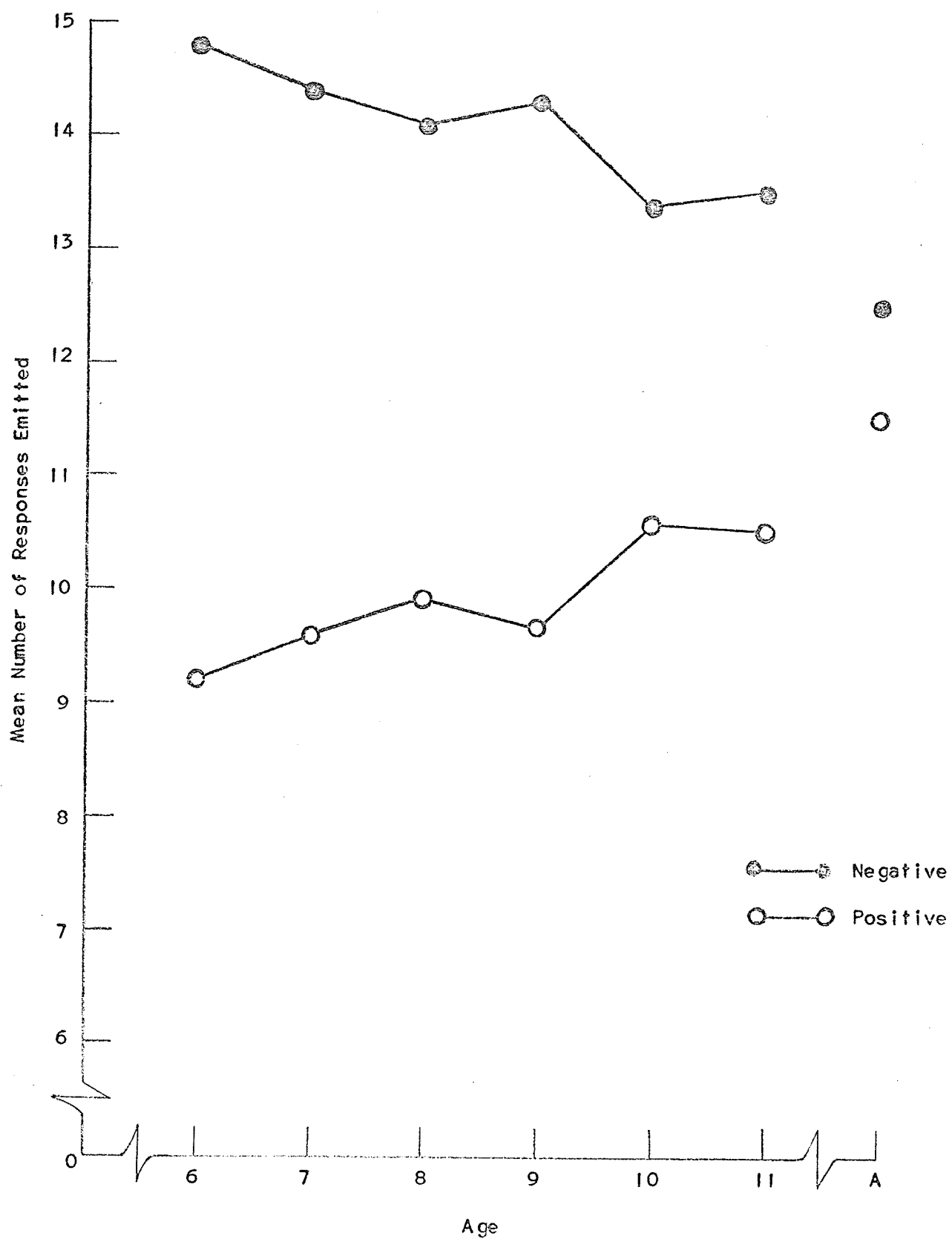


Figure 5. Mean number of negative (versus positive) responses emitted in the nonfiltered condition as a function of age.

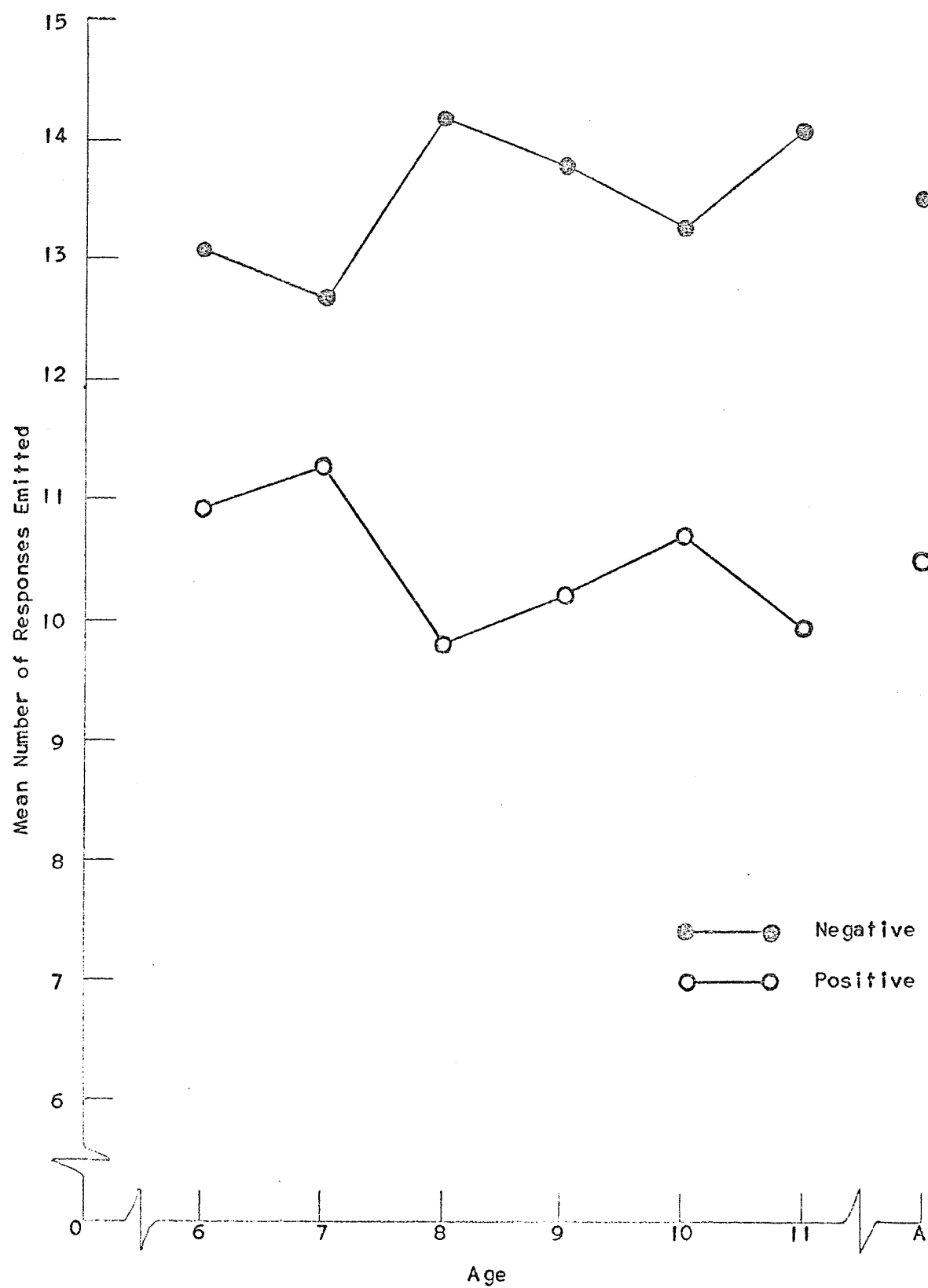


Figure 6. Mean number of negative (versus positive) responses emitted in the filtered condition as a function of age.

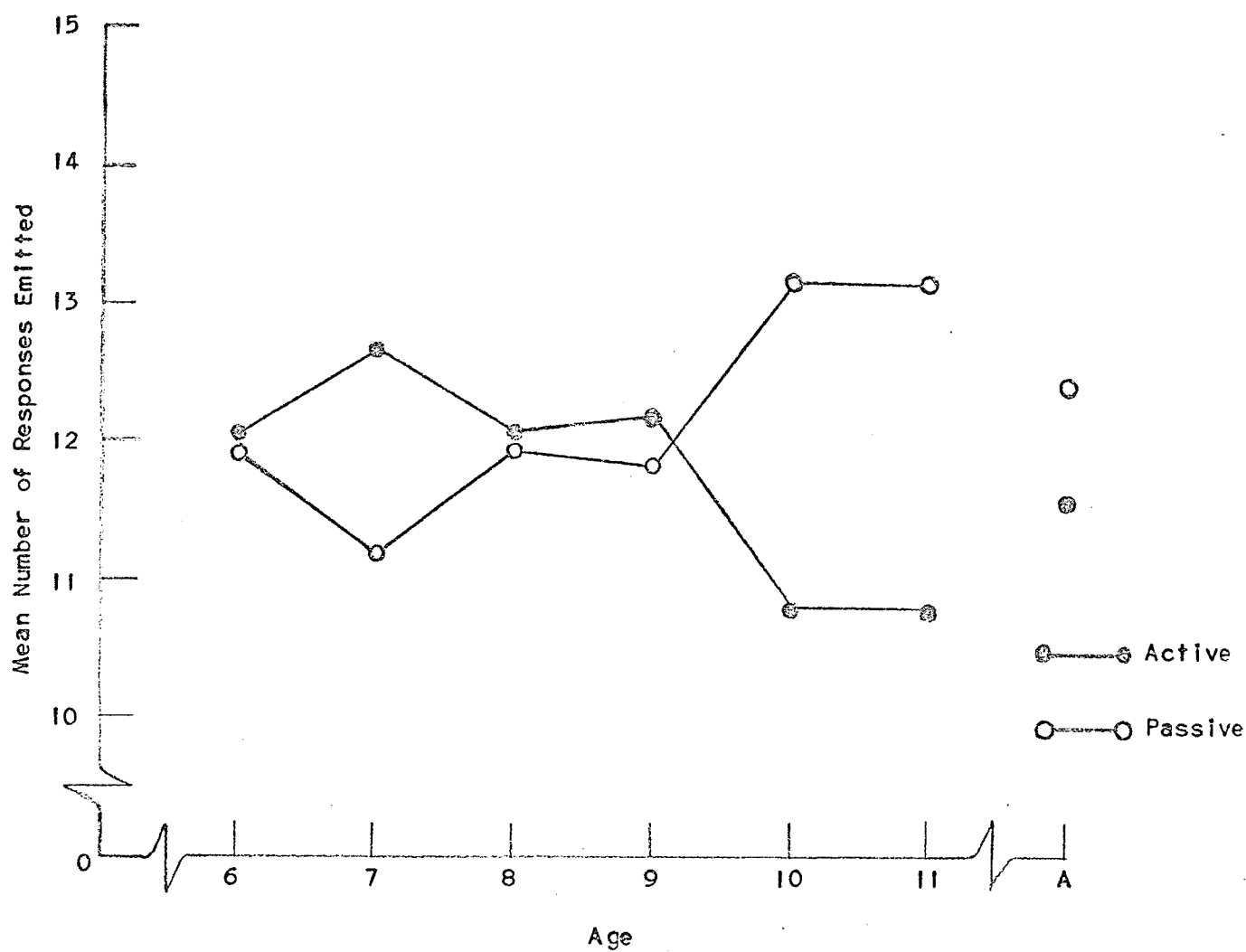


Figure 7. Mean number of active (versus passive) responses emitted in the nonfiltered condition as a function of age.

nonfiltered condition, young children (i.e., six through nine years of age) emitted more active (angry and happy) than passive (sad and loving) responses. Older children (i.e., ten and eleven years of age), on the other hand, responded more frequently with passive judgements than with active ones. The "passive" curve in Figure 7 is the mirror image of the "active" curve, since an increase in number of active judgements was reflected by a corresponding decrease in passive responses, and vice versa. Although a binomial estimate of probability failed to reveal any significant overall difference between number of active and number of passive judgements emitted ( $p > .05$ ), a one-way analysis of variance (Appendix M) performed on the active response data indicated a significant age difference in the tendency to respond with active judgements ( $F = 2.89, p < .025$ ). The Tukey ratio revealed that this result was due solely to the fact that the eleven-year-old children made significantly fewer active (i.e., more passive) responses than the seven-year-olds ( $q = 1.9, p < .05$ ).

Figure 8 shows that children of all ages in the filtered condition emitted more passive than active responses. A binomial estimate of probability indicated that this bias in the direction of the passive emotional states was significant ( $p < .05$ ). However, a one-way analysis of variance (Appendix N) revealed that there was no significant difference with age in the number of active (or passive) responses emitted in the filtered treatment ( $F < 1$ ).



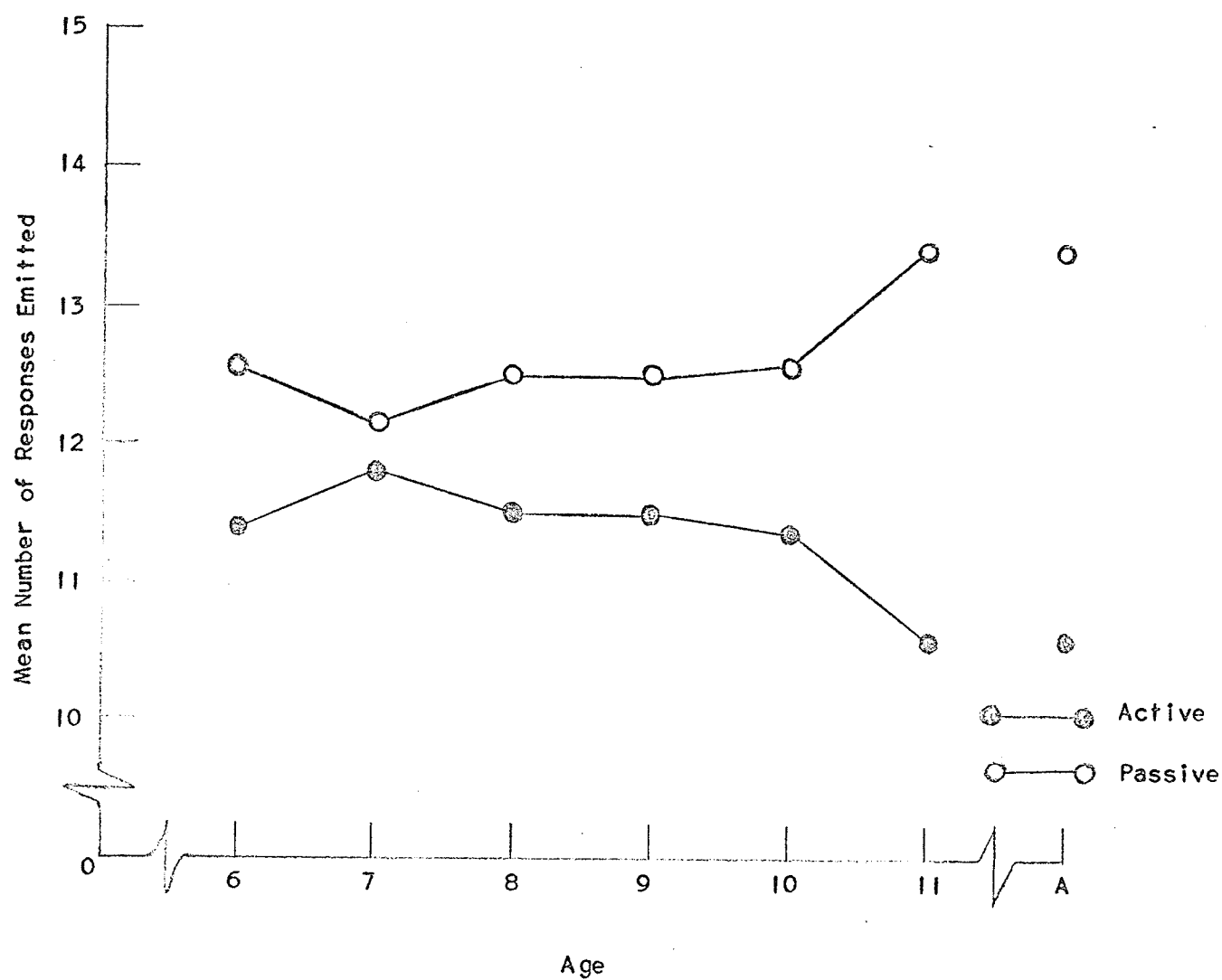


Figure 8. Mean number of active (versus passive) responses emitted in the filtered condition as a function of age.

## DISCUSSION

### GENERAL INFERENCES AND CONCLUSIONS

The results of the present study support the major hypothesis I for both the nonfiltered and filtered conditions, in that there was a steady and progressive increase with age in children's ability to identify the emotional content of speech. This finding, that children become more "adult-like" in their judgement of emotion in speech as they get older, is consistent with earlier results reported by Gates (1927), Dimitrovsky (1964), and Fenster (1967).

It has been suggested that infants six months of age can respond to tonal variation (cf., Dimitrovsky, 1964). While this research did not explore sensitivity prior to age six, there was a linear relationship between age and mean number of correct identifications. It seems reasonable to speculate, then, that development of emotional perception may begin in infancy and gradually improve throughout childhood. Certainly, the present investigation provides further evidence that considerable growth in ability to identify the emotional content of speech occurs in childhood between the ages of six and eleven.

As expected, subjects at all age levels in the filtered condition performed less accurately than those hearing nonfiltered speech. The fact that subjects of all ages do more poorly when listening to filtered material does not support Starkweather's (1956) contention that words may be distracting for persons making emotional judgements. It seems rather, that something of the voice quality is lost when speech is filtered, and that certain cues to emotional state are removed. Indeed, this research provides empirical evidence that investigators such as Mehrabian and Wiener (1967)

are justified in not using the filtering technique on the grounds that it somehow modifies the vocal aspects of speech while removing the semantic content. Therefore, even though the filter device makes possible the examination of unacted, real-life voice samples, using it as an unobtrusive tool may well result in the reduction of tonal information. It is possible, however, that extremely sophisticated filters being developed at present may be more efficient in terms of leaving intact the tonal qualities of speech (cf., Starkweather, 1967, 1969).

Exploration of the filter by emotion interaction provides some more precise information as to what cues are lost during filtering. The interaction appears to be due to the fact that subjects in the nonfiltered condition performed considerably better than those in the filtered treatment when judging the active emotional states, but only slightly better when hearing the passive emotions. That is, the differences between the mean scores attained in the nonfiltered versus the filtered condition were greater for the active emotions of anger and happiness than for the passive emotions of sadness and love. In the case of love, as opposed to the other three emotional states, there was no significant difference between the mean scores in the nonfiltered and filtered treatments. It would appear, then, that cues essential to the perception of activity may be lost when speech is filtered, while those involved in the communication of passivity are left more intact. In other words, it is likely that the vocal signals characteristic of active vocal expressions (e.g., extreme tonal variation and rises and falls in pitch) are removed by filtering.

The secondary hypotheses concerning the dimension of valence were upheld. Hypothesis (a), that children would judge negative emotions more

accurately than positive ones, was supported for the nonfiltered and filtered treatments. Subjects of all ages in both conditions made significantly more correct identifications of the negative than of the positive emotional states. This finding was consistent with the earlier work by Dimitrovsky (1964) and Fenster (1967), who both reported a similar emotion effect.

Only Fenster (1967) has attempted to account for this sensitivity to the negative emotions. He theorized that children (as well as adults) make more correct judgements of negative than positive stimuli because they have learned through experience that an error in recognizing the negative meaning of a communication can have serious consequences. For example, if a child fails to perceive a negative message from his parents, he is more likely to be punished than if he does not understand a positive communication. In effect, then, Fenster (1967) accounts for the accuracy of negative judgements in terms of what he calls "survival value." Although this explanation is post-hoc and untested, it should serve to provide some direction for future research.

Hypothesis (c) was also concerned with the valence dimension. This hypothesis, that children would emit more negative than positive judgements, was supported in the nonfiltered and filtered conditions. Subjects in both treatment groups exhibited a significant response bias in the direction of negativity. So although the current data are not consistent with Fenster's (1967) assertion that there is no tendency for children to respond with a preponderance of negative judgements, they support Dimitrovsky's (1964) finding that children of all ages do indeed emit more negative than positive responses. Earlier research has suggested that adults, unlike the children

in the present study, do not exhibit a response bias towards the negative emotions (cf., Davitz, 1964; Dimitrovsky, 1964). Thus, as Dimitrovsky (1964) suggests, it may be characteristic of children to interpret vocal expressions in terms of negative emotional meanings. That is, children may be more negative or "pessimistic" than adults. Although analyses of variance did not reveal a significant decrease with age in the tendency to emit negative responses, Figure 5 suggests that, in the nonfiltered condition at least, subjects may in fact become gradually less negative with increasing age. It is possible that an age trend of this nature may be uncovered if the experiment were extended to include groups over eleven years of age.

Hypotheses (b) and (d) were concerned specifically with the activity dimension. As predicted in hypothesis (b), errors occurred most frequently along this dimension, such that anger and happiness were mistaken for one another, as were sadness and love. This finding is not surprising if one considers the speech samples themselves. The emotions of anger and happiness tended to both be expressed in loud and forceful tones. Due to the similarity between them (in terms of gross vocal cues), it follows that these two active emotions would be mistaken for one another. Likewise, it is possible that the passive emotional states of sadness and love were confused because the vocal expressions of each were quiet and subdued.

Hypothesis (d), that children would respond with more passive than active judgements, was only partially supported in the nonfiltered condition. Although Dimitrovsky (1964) had found a bias in the direction of the passive emotional states for subjects of all ages, only the older children (i.e., ten and eleven-year-olds) in the present investigation emitted more passive than active responses. This inconsistency is difficult to explain

without reference to additional data. The results herein did indicate, however, that there was a significant age effect in the nonfiltered condition, with older subjects emitting fewer active (more passive) judgements than younger children.

In the filtered treatment, however, hypothesis (d) was upheld, as subjects of all ages emitted more passive than active responses. As implied previously, this result was likely due, in large part, to the fact that the filter seemed to eradicate vocal cues essential to the judgement of activity.

#### FUTURE CONSIDERATIONS

The present study was, in many respects, exploratory in nature. It was established that children could identify the emotional meaning of filtered vocal expressions at levels significantly above chance (even though they performed less accurately than those judging nonfiltered material). Further, an improvement in sensitivity with age and a response bias in the direction of negativity were noted in both the nonfiltered and filtered conditions. Now that a foundation has been laid, there are a number of directions future programmes of research in the area may take.

In retrospect, it can be seen that a few modifications of the methodology would be in order before attacking new theoretical issues. In the first place, it may well be that Dimitrovsky's (1964) standard paragraph: "I am going out now . . .," is not neutral. Indeed, it might be negative, in that it could cause children to feel they are being left alone. Since children hearing filtered (as well as nonfiltered) speech emitted more negative than positive judgements, it seems unlikely that the response bias was due to the semantic content of the standard paragraph. Nonetheless, it would be a wise precaution to have children rate future paragraphs, rather

than assuming neutrality.

In terms of the economics of research, it might also be wise to compare the performance of subjects seen individually with those run in groups. If there were no significant differences, a number of subjects could be seen simultaneously, instead of one at a time. In the group situation, it might be possible to modify the procedure such that the experimenter would not have to record each subject's response himself. To illustrate, each child could be provided with a series of drawings of the four stick figures. For each speech sample, the subject would make his judgement by placing a check mark under the figure of his choice, rather than pointing. Also, since by far the greatest number of subjects thought that the drawing of the woman with her arms around the child was the most loving, it seems unnecessary to use two drawings representing this emotional state.

The sensitivity of preschool children to the emotional content of speech remains essentially unexplored (Dimitrovsky, 1964). In addition, it has not been established unequivocally whether or not the tendency to emit more negative than positive responses decreases with age. Consequently, there is little doubt that this research should be extended to include an assessment of the performance of both younger and older subjects.

One of the more obvious questions that has also been neglected deals with "who communicates most effectively to whom" (cf., Davitz, 1964). Aside from varying the sex of the adult communicators and the child receivers, it would also be worthwhile to consider the responses of children and adults to vocal expressions from child speakers. As well, research could be undertaken to explore the potential role the vocal channel may play in pathology. For example, an effort could be made to determine whether "disturbed"

children respond differently to the emotional content of messages from "normal" and "disturbed" communicators. Due to the fact that it may be desirable to use natural, unacted voice samples, the filter device could be especially valuable for this type of work.

Future research could also involve a consideration of social class variables. Schmidt and Hore (1970) have demonstrated that mothers considered to be from high socioeconomic status groups looked at their preschool children, and reciprocated glances from them, more often than mothers from lower social classes. Further, in the same investigation, it was shown that women from lower class groups used more physical contact when communicating with their children. If such class differences exist in this type of nonverbal communication, it is possible that nonverbal vocal signals may also vary as a function of socioeconomic status.

At an even broader level, it has been suggested that there are cross-cultural differences in nonverbal behaviour in terms of gestures; body orientation, personal space, and so on (Argyle, 1972). It is worth asking whether children from other cultures are more (or less) sensitive to emotion in speech than are Canadian children. Again, the filter would likely be involved in any empirical test of this notion, since it could be used to control for language differences among subjects.

It would be remiss not to re-emphasize a most important point. Throughout this study, the vocal channel alone has been examined. However, in real life, communication channels often interact with one another, and combinations of emotional cues from the various channels occur simultaneously. In short, "what" a person says, "how" he says it (in terms of tone, facial expression, and the like), and "where" he says it are often simultaneously



involved in communicating emotion. Verbal, vocal, and visual cues all play a part in most interpersonal relationships (cf., Davitz, 1964; Reusch, 1957). It must always be kept in mind that the channels of communication are not mutually exclusive; the vocal channel does not exist independent from verbal and visual information.

Thus, the ultimate aim of much of the research in the area must be to consider all channels in combination. The work dealing with incongruency represents a noteworthy first step in this direction. For example, it has already been shown that it is important for parents to be consistent on the nonverbal and verbal levels when communicating to their children (Bugental, Love, Kaswan, and April, 1971). At this time, however, it is not clear whether children place differential weightings on the cues coming in on different channels than do adults. The fact that parents may inadvertently confuse and harm their children by communicating in an incongruent fashion is of practical value in regard to child rearing. Perhaps the most pressing need at the present time is to consider further this possibility.

#### SUMMARY

Semantic-controlled (nonfiltered) and semantic-free (filtered) vocal expressions were played to separate groups of boys six through eleven years of age. In both the nonfiltered and filtered treatments, there was a progressive increase with age in ability to identify the emotional content of the speech samples. However, children listening to the nonfiltered speech performed more accurately than those hearing the filtered samples. Negative emotional states (anger and sadness) were judged more accurately than positive emotions (happiness and love) by children in both conditions.

## REFERENCES

- Argyle, M. Nonverbal communication in human social interaction. In R.A. Hinde (Ed.), Nonverbal communication. Cambridge: University Press, 1972.
- Arnold, M.B. (Ed.), The nature of emotion. London: Penguin Books Ltd., 1968.
- Arnold, M.B. (Ed.), Feelings and emotions. New York: Academic Press, 1970.
- Bateson, G., Jackson, D.D., Haley, J., and Weakland, J.H. Toward a theory of schizophrenia. Behavioral Science, 1956, 1, 251-264.
- Beakel, N.G. Parental verbal and nonverbal communication and psychopathology. Dissertation Abstracts International, 1971, 31(7-B), 4325.
- Beakel, N.G., and Mehrabian, A. Inconsistent communications and psychopathology. Journal of Abnormal Psychology, 1969, 74, 126-130.
- Beldoch, M. Sensitivity to expression of emotional meaning in three modes of communication. In J.R. Davitz (Ed.), The communication of emotional meaning. New York: McGraw-Hill Book Co., 1964.
- Boyle, V.A. Visual stimulation and comprehension of compressed speech. Dissertation Abstracts International, 1970, 30(11-B), 5221.
- Bugental, D.E., Kaswan, J.W., and Love, L.R. Perception of contradictory meanings conveyed by verbal and nonverbal channels. Journal of Personality and Social Psychology, 1970, 16, 647-655.
- Bugental, D.E., Kaswan, J.W., Love, L.R., and Fox, M.N. Child versus adult perception of evaluative messages in verbal, vocal, and visual channels. Developmental Psychology, 1970, 2, 367-375.
- Bugental, D.E., Love, L.R., and Gianetto, R.M. Perfidious feminine faces. Journal of Personality and Social Psychology, 1971, 17, 314-318.
- Bugental, D.E., Love, L.R., Kaswan, J.W., and April, C. Verbal-nonverbal conflict in parental messages to normal and disturbed children. Journal of Abnormal Psychology, 1971, 77, 6-10.
- Cannon, W.B. The James-Lange theory of emotions: A critical examination and an alternative theory. American Journal of Psychology, 1927, 39, 106-124.

- Clausen, J.A., and Kohn, M.L. Social relations and schizophrenia: A research report and a prospective. In D.D. Jackson (Ed.), The etiology of schizophrenia. New York: Basic Books, 1960.
- Davitz, J.R. (Ed.), The communication of emotional meaning. New York: McGraw-Hill Book Co., 1964.
- Davitz, J.R. A dictionary and grammar of emotion. In M.B. Arnold (Ed.), Feelings and emotions. New York: Academic Press, 1970.
- Davitz, J.R., and Davitz, L.J. The communication of feelings by content-free speech. Journal of Communication, 1959a, 9, 6-13.
- Davitz, J.R., and Davitz, L.J. Correlates of accuracy in the communication of feelings. Journal of Communication, 1959b, 9, 110-117.
- Davitz, J.R., and Davitz, L.J. Nonverbal vocal communication of feeling. Journal of Communication, 1961, 11, 81-86.
- Dimitrovsky, L. The ability to identify the emotional meaning of vocal expressions at successive age levels. In J.R. Davitz (Ed.), The communication of emotional meaning. New York: McGraw-Hill Book Co., 1964.
- Duncan, S. Nonverbal communication. Psychological Bulletin, 1969, 72, 118-137.
- Dusenberry, D., and Knower, F.H. Experimental studies of the symbolism of action and voice. II. A study of the specificity of meaning in abstract tonal symbols. Quarterly Journal of Speech, 1939, 25, 67-75.
- Ekman, P., Friesen, W.V., and Tomkins, S.S. Facial affect scoring technique: A first validity study. Semiotica, 1971, 3, 37-58.
- Fairbanks, G., and Pronovost, W. An experimental study of pitch characteristics of the voice during the expression of emotion. Speech Monographs, 1939, 6, 87-104.
- Fenster, C.A. Vocal communication of emotional meaning among adults and children. Dissertation Abstracts International, 1967, 28(4-B), 1964-1965.
- Fujimoto, E.K. The comparative communicative power of verbal and nonverbal symbols. Dissertation Abstracts International, 1972, 32(7-A), 4152.
- Gates, G.S. The role of the auditory element in the interpretation of emotions. Psychological Bulletin, 1927, 24, 175. (Abstract)
- Goldhaber, G.M. Listener comprehension of compressed speech as a function of academic grade level of the subjects. Journal of Communication, 1970, 20, 167-173.

- Goldhaber, G.M., and Weaver, C.H. Listener comprehension of compressed speech when the difficulty, rate of presentation, and sex of listener are varied. Speech Monographs, 1968, 35, 20-25.
- Holsti, O.R. Content analysis for the social sciences and humanities. Don Mills: Addison-Wesley Publishing Co., 1969.
- Kirk, R.E. Experimental design: Procedures for the behavioral sciences. Belmont, California: Brooks/Cole Publishing Co., 1969.
- Knower, F.H. Analysis of some experimental variations of simulated vocal expressions of the emotions. Journal of Social Psychology, 1941, 14, 369-372.
- Knower, F.H. Studies in the symbolism of voice and action: V. The use of behavioral and tonal symbols as tests of speaking achievement. Journal of Applied Psychology, 1945, 29, 229-235.
- Kramer, E. Judgement of personal characteristics and emotions from nonverbal properties of speech. Psychological Bulletin, 1963, 60, 408-420.
- Lange, C.G., and James, W. The emotions. Baltimore: Williams and Wilkins Co., 1922.
- Levitt, E.A. The relationship between vocal and facial emotional communicative abilities. Dissertation Abstracts International, 1962, 23(5), 1783.
- Levitt, E.A. The relationship between abilities to express emotional meanings vocally and facially. In J.R. Davitz (Ed.), The communication of emotional meaning. New York: McGraw-Hill Book Co., 1964.
- Mahl, G.F., and Schulze, G. Psychological research in the extralinguistic area. In T.A. Sebeok, A.S. Hayes, and M.C. Bateson (Eds.), Approaches to semiotics. The Hague: Mouton and Co., 1964.
- McCarthy, D. Some possible explanations of sex differences in language development and disorders. Journal of Psychology, 1953, 35, 155-160.
- McCarthy, D. Language development. In L. Carmichael (Ed.), Manual of Child Psychology. New York: Wiley, 1954.
- Mehrabian, A., and Ferris, S.R. Inference of attitudes from nonverbal communication. Journal of Consulting Psychology, 1967, 31, 248-252.
- Mehrabian, A., and Wiener, M. Decoding of inconsistent communications. Journal of Personality and Social Psychology, 1967, 6, 109-114.
- Nash, J. Developmental psychology: A psychobiological approach. New Jersey: Prentice-Hall, Inc., 1970.
- Olsen, P.T. Response to incongruent verbal-vocal communications in children and early adolescents. Dissertation Abstracts International, 1967, 28(5-B), 2147.

- Orr, D.B. Time compressed speech — a perspective. Journal of Communication, 1968, 18, 288-292.
- Orr, D.B., Friedman, H.L., and Williams, J.C.C. Trainability of listening comprehension of speech discourse. Journal of Educational Psychology, 1965, 56, 148-156.
- Pfaff, P.L. An experimental study of the communication of feeling without contextual material. Speech Monographs, 1954, 21, 155-156. (Abstract)
- Pollack, I., Rubenstein, H., and Horowitz, A. Communication of verbal modes of expression. Language and Speech, 1960, 3, 121-130.
- Ruesch, J. Nonverbal language and therapy. Psychiatry, 1957, 18, 323-330.
- Sampson, O.C. The speech and language development of five-year-old children. British Journal of Educational Psychology, 1959, 29, 217-222.
- Schachter, S., and Singer, J. Cognitive, social and physiological determinants of emotional state. Psychological Review, 1962, 69, 379-399.
- Schmidt, W.H., and Hore, T. Some nonverbal aspects of communication between mother and preschool child. Child Development, 1970, 41, 889-896.
- Schuham, A.I. The double-bind hypothesis a decade later. Psychological Bulletin, 1967, 68, 409-416.
- Soskin, W.F., and Kauffman, P.E. Judgement of emotion in word-free voice samples. Journal of Communication, 1961, 11, 73-80.
- Starkweather, J.A. Content-free speech as a source of information about the speaker. Journal of Abnormal and Social Psychology, 1956, 52, 394-402.
- Starkweather, J.A. Vocal communication of personality and human feelings. Journal of Communication, 1961, 11, 63-72.
- Starkweather, J.A. Computer-assisted learning in medical education. Canadian Medical Association Journal, 1967, 97, 733-738.
- Starkweather, J.A. Measurement methods for vocal information. In G. Gerbner, O.R. Holsti, K. Krippendorff, W.J. Paisley, and P.J. Stone (Eds.), The analysis of communication content. New York: John Wiley and Sons, Inc., 1969.
- Thompson, C.W., and Bradway, K. The teaching of psychotherapy through content-free interviews. Journal of Consulting Psychology, 1950, 14, 321-323.
- Webb, E.J., Campbell, D.T., Schwartz, R.D., and Sechrest, L. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand-McNally, 1966.

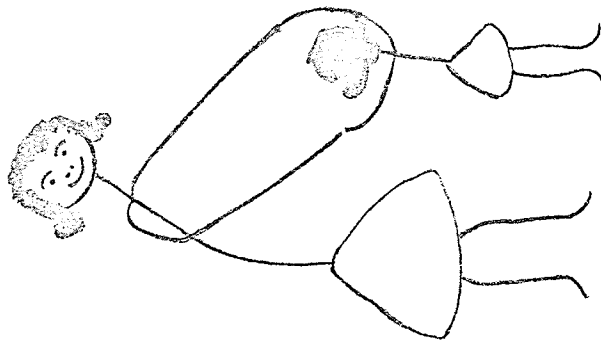
Winer, B.J. Statistical principles in experimental design. New York: McGraw-Hill Book Co., 1962.

Young, F.M. An analysis of certain variables in a developmental study of language. Genetic Psychology Monographs, 1941, 23, 3-141.

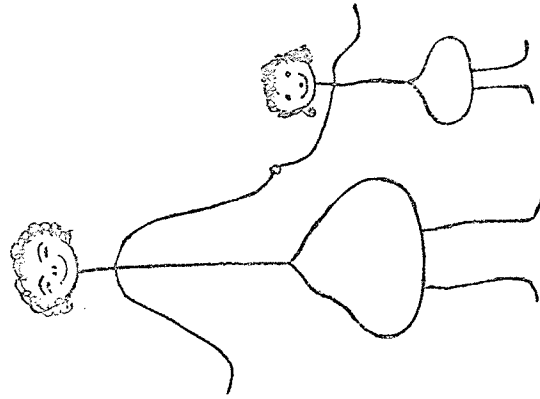
## APPENDIX A

Actual Sizes of Stick Figure

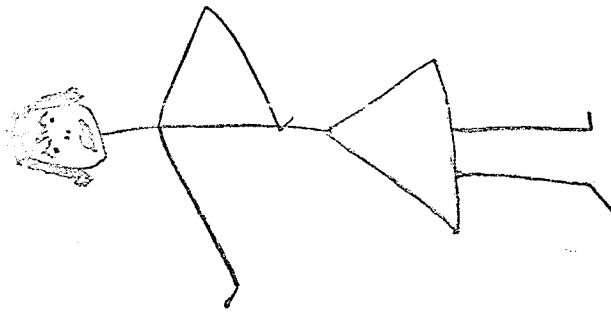
Drawings Presented



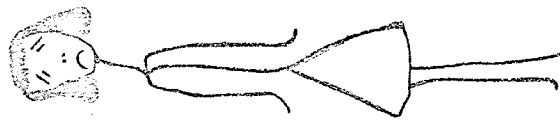
Loving



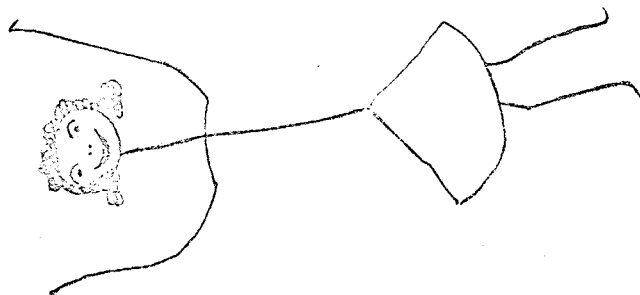
Loving



Angry



Sad



Happy



## APPENDIX B

Orders in which Each Actress Expressed the  
Four Emotional States

C1       S, A, H, L.

C2       H, S, L, A.

C3       L, S, A, H.

C4       H, L, A, S.

C5       A, H, S, L.

C6       S, L, H, A.

C = Communicator

H = Happy

S = Sad

L = Loving

A = Angry

Restrictions:

(a) no more than two actresses  
could be asked to communicate  
the same emotion first

(b) the order in which the four  
emotions were acted was different  
for each speaker

## APPENDIX C

### The Random Arrangement of the 24 Vocal Expressions

(The orders are identical for both the nonfiltered and filtered conditions, since the filtered tape was prepared directly from the randomized nonfiltered material)

C4 C1 C1 C3 C3 C6 C2 C2 C5 C6 C2 C3 C3 C2 C1 C5 C6 C4 C5 C4 C6 C4 C5 C1  
A L S S A L S A H S H H L L H A H S L H A L S A

C = Communicator

H = Happy

S = Sad

L = Love

A = Angry

For example, "C4"  
A represents the speech sample in  
which actress number 4 simulated anger

#### Restrictions:

- (a) no emotional state can appear more than twice in a row
- (b) no more than two speech samples from any one speaker can appear in succession

## APPENDIX D

Total Number of "Correct" Responses by Adult Judges (Scored against  
Speaker's Intent) for Each Emotional State in the  
Nonfiltered and Filtered Conditions

Emotion	Nonfiltered	Filtered
Happy	49	29
Sad	49	43
Loving	39	32
Angry	52	36

Maximum possible score = 60

## APPENDIX E

Instructions Read to Subjects in the Non-  
filtered and Filtered Conditions

The instructions were identical for the children and adult judges. Subjects in the filtered condition were read the instructions presented below. Subjects in the nonfiltered treatment heard the same instructions, with the words in parentheses deleted.

"You are going to hear the voices of different people speaking (from far away. Because they are far away you will not be able to hear clearly the words they are speaking.) I would like you to tell me whether the person speaking is a man, a lady, a boy, or a girl. That is, after you hear the voice of the person, you will tell me if you think it was a man, a lady, a boy, or a girl."

Each of the four tapes was played, and after each the subject was asked to make his judgement. (i.e., "Was that a man, a lady, a boy, or a girl?").

"Now I am going to show you two drawings of a lady and a child."

The two cards with the stick figure representations of love were presented.

"Look at each drawing carefully. Can you point to the lady that looks as if she loves the child more? Which lady do you think loves the child more?"

The drawing of love chosen was placed in front of the subject, along with the three other stick figure diagrams. For each subject the drawings were arranged in random order.

"Here you see drawings of a happy lady, a sad lady, a loving lady, and an angry lady."



The order or the emotions here depended on the random arrangement of the cards.

"Look at each drawing carefully. This lady is happy, this one is sad, this one is loving, and this one is angry."

As this description was given, the experimenter pointed to each appropriate drawing.

"Now, can you point to the happy lady, the sad lady, the loving lady, and the angry lady?"

The order in which each subject had to identify the drawings was completely random. The procedure was repeated until a criterion of two correct identifications of all four drawings was reached.

"You are now going to hear different ladies say the same thing in different ways. (The ladies are speaking from far away, so you will not be able to hear clearly the words they are saying.) Some will sound happy, some sad, some loving, and some angry. Listen to each one and point to the lady it sounds like."

Each speech sample was played.

"Does it sound like the happy lady, the sad lady, the loving lady, or the angry lady?"

The response of the subject to each of the 24 speech samples was recorded.

## APPENDIX F

Mean and Standard Deviations of Number of Correct  
Identifications at Successive Ages in  
Nonfiltered and Filtered Conditions

Age	Nonfiltered		Filtered	
	Mean Score	SD	Mean Score	SD
6	10.4	1.20	7.2	1.81
7	12.2	2.26	7.7	1.10
8	12.7	.90	9.5	1.50
9	13.7	.90	9.7	1.48
10	14.0	1.09	10.5	1.69
11	15.0	1.41	11.2	1.66
A	18.9	1.89	14.0	1.41

Maximum possible score = 24

A = Adults

## APPENDIX G

Means and Standard Deviations of Number of Correct Identifications  
at Successive Age Levels in Nonfiltered and Filtered Conditions  
for Each Category of Emotional Meaning

Age	Nonfiltered										Filtered									
	Happy					Sad					Loving					Angry				
	Mean	SD	Mean	SD	Mean	Mean	SD	Mean	SD	Mean	Mean	SD	Mean	SD	Mean	Mean	SD	Mean	SD	Mean
6	2.0	.77	3.6	.80	1.2	1.2	.87	3.6	.92	1.0	.63	2.7	.78	1.3	1.10	2.2	.98	2.2	.98	2.2
7	2.6	1.28	3.0	.77	1.9	1.9	1.58	4.7	1.00	1.3	.90	2.2	.98	1.5	.81	2.7	.64	2.7	.64	2.7
8	2.5	.81	3.5	.81	2.2	2.2	1.08	4.5	.92	1.9	.70	3.4	1.36	1.6	1.11	2.6	1.28	2.6	1.28	2.6
9	3.5	1.02	3.8	.75	1.9	1.9	1.04	4.5	.67	1.4	1.11	3.4	1.56	2.0	1.18	2.9	.83	2.9	.83	2.9
10	2.5	.67	3.8	1.33	3.2	3.2	1.08	4.5	.67	2.0	1.18	3.3	1.19	2.6	.92	2.6	1.02	2.6	1.02	2.6
11	3.3	1.10	4.1	.70	3.1	3.1	1.22	4.5	.67	1.2	.75	3.5	.92	3.0	1.34	3.5	1.20	3.5	1.20	3.5
A	4.9	.70	4.9	.54	3.9	3.9	1.22	5.2	.75	2.9	.54	4.3	1.10	3.2	1.30	3.6	1.43	3.6	1.43	3.6

Maximum possible score = 6

A = Adults

## APPENDIX H

Means and Standard Deviations of Number of Correct Identifications  
for Each Category of Emotional Meaning in  
Nonfiltered and Filtered Conditions

	Nonfiltered		Filtered	
	Mean	SD	Mean	SD
Angry	4.38	.92	2.75	1.09
Sad	3.63	.96	3.08	1.26
Happy	2.73	1.10	1.47	.97
Loving	2.25	1.36	2.00	1.25

Maximum possible score = 6

## APPENDIX I

Frequency in Each Response Category for Each Stimulus Category  
for the Total Nonfiltered and Filtered Groups



STIMULUS	RESPONSE							
	Nonfiltered				Filtered			
	Happy	Sad	Loving	Angry	Happy	Sad	Loving	Angry
Happy	213	36	71	100	117	53	88	162
Sad	51	267	78	24	56	228	101	35
Loving	50	176	174	20	50	164	152	54
Angry	50	29	26	315	113	50	56	201

Maximum possible score = 420

## APPENDIX J

Number of Emitted Responses in Each Category of Emotional Meaning for Each  
Age Group in the Nonfiltered and Filtered Conditions



## APPENDIX K

Analysis of Variance for Mean Number of Negative Responses Emitted in the  
Nonfiltered Condition as a Function of Age

Source	SS	df	MS	F
Age	14.68	5	2.94	.71
Error	223.90	54	4.15	
Total	238.58	59		

## APPENDIX L

Analysis of Variance for Mean Number of Negative Responses Emitted in the  
Filtered Condition as a Function of Age

Source	SS	df	MS	F
Age	17.73	5	3.55	.92
Error	207.20	54	3.84	
Total	224.93	59		

## APPENDIX M

Analysis of Variance for Mean Number of Active Responses Emitted in the  
Nonfiltered Condition as a Function of Age



Source	SS	df	MS	F
Age	27.88	5	5.58	2.89*
Error	104.10	54	1.93	
Total	131.98	59		

\* $p < .025$

## APPENDIX N

Analysis of Variance for Mean Number of Active Responses Emitted in the  
Filtered Condition as a Function of Age

Source	SS	df	MS	F
Age	8.13	5	1.63	.43
Error	206.80	54	3.83	
Total	214.93	59		

## VITA

- Name: - Kenneth Wilfred McCluskey
- Place and date of birth: - Winnipeg, December 18, 1949
- Education: - West Kildonan Collegiate Institute, 1964-1967  
 - University of Manitoba, 1967-1971  
 B.A. (Honours, Psychology), May, 1971  
 - Faculty of Graduate Studies, University of Manitoba, M.A. Programme, 1971-1972  
 M.A. (to be conferred October, 1974)  
 - Faculty of Graduate Studies, University of Manitoba, Ph.D. Programme, 1972-1974
- Experience: - Teaching Assistant, Introductory Psychology, University of Manitoba, 1970-1971; 1971-1972  
 - Teaching Assistant, Child Development, University of Manitoba, 1972  
 - Teaching Assistant, Adolescent Development, University of Manitoba, 1973  
 - Research Assistant, University of Manitoba, Department of Psychology, 1971-1972; 1972 (summer); 1973  
 - Graduate student head of subject pool for psychological research, University of Manitoba, 1973-1974  
 - Consultant, Seven Oaks School, Winnipeg, Manitoba  
 - Reviewer-consultant, texts in psychology, Choice, American Library Association, Chicago, Illinois, 1973-1974
- Publications: Niemi, R.R., and McCluskey, K.W. Age differences in the perception of the emotional content of speech. Address to the Canadian Psychological Association, June, 1974.
- McCluskey, K.W., Niemi, R.R., and Ferrer, C. Diferencias culturales en la percepcion de las emociones (Cultural differences in the perception of emotion). Informacion Nacional, 1973, 9, 30-31.