Journal of Criminal Law and Criminology

Volume 6 | Issue 6 Article 4

1916

Capacity to Report upon Moving Pictures as Conditioned by Sex and Age

Edwin G. Boring

CAPACITY TO REPORT UPON MOVING PICTURES AS CONDITIONED BY SEX AND AGE

A CONTRIBUTION TO THE PSYCHOLOGY OF TESTIMONY.[1]

EDWIN G. BORING.

The adequacy of the "picture-test" for the determination of the reliability of report upon a series of events has yet to be demonstrated.[2] It can not be assumed that excellence in the description of a static scene presupposes excellence in the report upon events. On the other hand, the event-test, when it is made to include human action, presents difficulties of accurate control. The disadvantage of both forms of tests can, however, be obviated by the use of the moving picture—the method employed in this experiment.

There is considerable difference of opinion as to the reliability of children as witnesses. Some psychologists regard them as very unreliable, while others maintain that, under usual circumstances at least, they are quite the equals of adults.[4] In the present experiment, children and adults, of both sexes, were employed for the purposes of comparison.

[1]Studies from the Cornell Educational Laboratory No. 21. This work was performed by the author in 1912. The general plan was suggested and supervised by Professor G. M. Whipple now of the University of Illinois, then in charge of educational psychology at Cornell University.

[2]H. GROSS (Zur Frage der Zeugenaussage, *H. Gross' Archiv.* 36; 1910, 372ff.), thinks that ability to report accurately upon a picture does not raise the presumption that the individual would be able to report upon a series of events accurately and doubts the value of the picture-test for jurisprudence. H. B. GERLAN, Zur Frage der Zeugenaussage, *ibid.* 39; 1910, 116ff., on the other hand, believes in the significance of the picture-test for legal procedure.

[4]A. BAGINSKY (*Die Kinderaussage vor Gericht*, 1910), and E. DUPREE (Le temoignage: étude psychologique et médico-légale, *Rev. d. deux Mondes*, 55; 1910, 343), believe that children are in general unreliable. J. VARENDENCK (Les témoignages d'enfants dans un procés retentissant, *Arch. d. Psychol.*, 11; 1911, 171), has shown that the replies of children to implicative and suggestive questions are unreliable, while K. MARBE (Kinderaussagen in einem Sittlichkeitsprozess, *Fortschr. d. Psychol.*, 1; 1913, 339ff.), and MEHL, Beitrag zur Psychologie der Kinderaussage, *Arch. f. Krim.-Anthrop. u. Kriminalistik*, 49; 1912, 193), report instances which show the danger of accepting the testimony of girls under certain emotional conditions. On the other hand, GROSS (*op. cit.*), and R. HEINDL, (Die Zuverlässigkeit von Signalamentaussagen, *H. Gross' Archiv*, 33; 1909, 102) and E. v. KARMAN (Kriminalistische Beiträge III, Kinder als Zeugen, *ibid*, 50; 1913, 231), conclude that children constitute perfectly adequate observers—at least under certain circumstances.

For general summaries of this literature, see G. M. Whipple, *Psych. Bull.*, 7: 1910, 365ff.; 8: 1911, 307ff.; 9: 1912, 264ff.; 10: 1913, 264ff.; 11: 1914, 245ff.; 12: 1915, 221ff. (Most of these have also been reprinted in this JOURNAL.)

SUBJECTS.

Forty-four subjects were used. There were thirteen women, of whom ten were undergraduate students in Cornell University, and one a woman of over fifty years of age. Of the eleven men, three were undergraduates, five were graduate students in psychology, two were professors of psychology; one was over sixty years of age (not a college graduate). The children were taken from a single class in the Ithaca public schools and ranged from 9.5 to 12.6 years of age. There were twelve boys with an average age of 12.3 (P.E.=.49) and eight girls, also with an average age of 12.3 (P. E.=.37.)

MATERIAL.

The material selected for presentation was a scene from a photoplay, which begins suddenly with an encounter between a gentleman and a burglar and continues for about one minute until both men leave the scene.[5] The film is colored blue, in order to produce a moonlight effect, a condition which renders the detail slightly less distinct than it would otherwise be.

The action may be described as follows: The scene is laid in a garden. In the background is the wall of a house, to the right, a wall, about five feet high, in which there is a gate. A path runs from the gate straight across in front of the house to the left of the scene.

There are two actors, who may be designated as the 'man' and the 'burglar.' The man is well-dressed, wearing a silk hat and a light overcoat. He appears to be in evening dress. He is smooth-shaven and has dark hair and a firm, full face. The burglar is very uncouth with ragged clothes, unkempt hair partly covered by a cap, and a face lined and rough. Throughout the scene he maintains a cringing attitude, in striking contrast with the aggressive vigor of the man.

At the beginning of the scene the burglar is crouching before the gate with a sack of plunder beside him. He opens the gate, looks out, and withdraws quickly. Suddenly the man climbs over the wall and springs down upon the burglar, knocking him over and, as he lies prone, kicking the revolver from his hand. The man picks the revolver from the ground and covers the burglar with it, as the latter starts to rise; he leans over and peers into the burglar's face, starting back as if recognizing him. The burglar hangs his head, until the man, who has faced the audience for a moment with an expression of surprise, turns and addresses him. A conversation ensues, in which the burglar makes gestures of despair with his arms and then collapses in the shadow by the wall. The man, after pondering the situation, speaks to the burglar and motions to him to rise. The burglar obeys with great effort, lifts the bag of plunder and, at the direction of the man, carries it off the scene to the left. He returns without the plunder. The man by a gesture calls the burglar's attention to the revolver, and thrusts it into this right overcoat pocket, where he keeps it pointed at the burglar through the coat. He beckons to the burglar to precede him through the gate, and they go out together, the man opening the gate, although the burglar precedes him. Finally their heads can be seen as they pass along outside the wall.[6]

[5]The scene was taken from a film entitled *Van Bibber's Experiment*, produced by Thomas A. Edison, Inc.

[6]This description does not show the number of possible items in the material, but only the main features of the action and the principal details of the setting. The fullest written description obtained was approximately six times as long.

Experimental Procedure.

The Apparatus.—The kinematograph was[7] placed 4.8 meters from a white plaster wall, upon which the film was projected. The picture as shown was approximately 112 by 152 cm., subtending for the observer, who sat about 3.6 meters from the wall and squarely in front of the picture, a vertical angle of 17° and a horizontal angle of 23°. The rate of presentation was controlled by turning the crank of the machine in time with a swinging pendulum bob.

The Instructions.—The observers were not told before the experiment what was to be the exact nature of the proceeding, but were instructed as follows: "Sit in this chair and watch the wall over there. I am going to show you a picture upon the wall. I want you to watch it with your best attention. Be sure to watch it carefully all the time." In this way it was hoped to secure the degree of attention which would ordinarily result from interest. The children perhaps suspected some sort of test, although careful questioning failed to show that they had discovered its real nature. The adults were equally ignorant, excepting those who were borrowed from the psychological laboratory, many of whom suspected from the conditions and the instructions that the test was to be of ability to make a report.

Each subject was cautioned before leaving the laboratory not to tell anyone what had occurred. As the children came from the same class in school, this warning was especially important. It was probably effective, for the teacher of the class was unable to get any information from the children about their experiences, and those of the children who had already been through the test seemed to delight—so the teacher reported—in mystifying the others as to what they were required to do.

The Narrative.—As soon as the presentation was over, the lights were turned on in the room, the subject was seated at a table and asked to "tell just as fully as possible" everything that he had seen in the picture. The experimenter wrote at the subject's dictation. If the children had been obliged to write themselves, they could not have given as full reports. Even adults are less likely to report fully when they are obliged to write. When the narrative was completed, it was read to the subject, who was allowed to make any corrections and was required to say (*a*) to which of his statements he would take oath (oath was explained as "anything you would swear to in court"), (*b*) of which he felt certain, although unwilling to take oath, and (*c*) of which he was more or less doubtful.

[7]The writer is indebted to the Department of Physics of Cornell University for the use of their machine.

The Interrogatory.—After the narrative the subject was required to answer a questionary. His statement that he was more or less doubtful of his answer, certain of it, or willing to swear to it, was indicated after each question. The form of the questions was varied so as to conform with the terminology and the facts as stated by the subject in his narrative. The two actors, for example, were always called by the names which the subject had given them. In most cases such alteration of questions was slight. The questions, of which Nos. 9, 14, 16, 17, 18 and 24 are implicative questions, follow. Correct answers are given in parentheses:

QUESTIONARY.—*1.* Where did the man come from at the beginning of the scene? (From over the wall.)
2. Where did the burglar come from? (He was already in the yard beside the gate.)
3. Where did the man get the revolver with which he controlled the burglar? (From the ground; indirectly he got it from the burglar.)
4. When the burglar first arose from the ground did he stand or kneel? (Knelt.)
5. Did the man point the revolver at the burglar while he was rising? (Yes.)
6. What did the man do as soon as the burglar had risen? (Peered into the burglar's face and seemed surprised.)
7. Did the man speak to the burglar while he was in this position? (No.)
8. Did the burglar speak to the man? (No.)
9. Did the burglar resist when the man grabbed him by the throat? (The man did not take the burglar by the throat.)
10. What did the man do while the burglar was lying close by the wall? (Turned toward the audience and deliberated.)
11. Did the burglar obey quickly when he was told to get up? (No.)
12. What did the burglar do as soon as he had gotten up? (Went off to the left with the sack.)
13. Did the man point the revolver at him while he was doing this? (No.)
14. Did the burglar pass in front of or behind the rosebush? (There was no rosebush.)
15. When the burglar returned what did the man do? (Showed him the revolver in his hand, and then, placing it in his pocket, covered the burglar with it.)
16. Did the man take the burglar by the right or by the left shoulder? (He did not touch the burglar.)
17. Was the gate open or did the burglar open it? (The man opened it.)
18. At the end of the scene which one remained in the yard? (Both went out.)
19. Was the wall higher or lower than the man's shoulder? (Higher.)
20. What could you see outside over the wall? (Buildings.)
21. Did the gate open toward the men or away from them them? (Toward them.)
22. What sort of a neck-tie did the man wear? (White.)
23. Did the man wear gloves? On both hands? (He wore no gloves.)
24. Did the burglar wear a felt hat or a derby? (A cap.)
25. Did the bag appear heavy? About how heavy? (Yes. About 30 pounds, or as heavy as if filled with dirt.[8])
26. How long did it take to show the picture? (About one minute.[9])

[8]It is obvious from the difficulty with which the burglar lifts the bag that it must have weighed, say, about 30 pounds. When the subjects hesitated in answering, they were asked whether the bag was "as heavy as if filled with feathers—or clothes—or dirt."

[9]The subjects often gave the limits between which they felt sure the correct time must lie. In general, when asked to take oath, they extended the limits considerably.

Treatment of Data

Scoring.—In evaluating the protocols of the subjects it is first necessary to determine the total number of separate items in the material (*P*, in the formulae[10]). It is, however, difficult to decide just what constitutes an item. Is the fact that the man wore evening dress an item? Or are his high hat, his white tie, and his patent leather shoes, all separate items? If so, what about the fact that the shoes were laced, not buttoned, and that the tie was slightly crooked? It is possible to regard as an item any detail about which an independent statement can be made, and it seems to be some such practice that has prevailed. Such a method, however, results in the assignment of equal weight to factors that differ considerably in prominence and importance, since neither the prominence nor the importance of a factor is a function of the number of separate statements that can be made about it. Hence, when a measure of accuracy is desired with regard to the relative importance of detail, it is not fair to score the separate statements equally, to give, for example, equal weight to statements that the man climbed over the wall and that he wore no gloves.[11] Accordingly, we divided the complete description into a few phases of approximately equal importance. Each phase was again divided into such of its minor details as might represent equivalent part-factors of the particular phase. Thus, the recognition of the burglar by the man constitutes a phase of the action which is comparable to the phase in which the man and the burglar converse, and which may be analyzed into five part-phases, itemizing the behavior of the man and the burglar.

The actual separation of the material into phases was empirical.[12] Three graduate students, who had already acted as subjects, were shown the film slowly and with as many repetitions as they desired and were asked to write a complete description, dividing it into sections which corresponded to what they regarded as the main phases. There was great uniformity. One student distinguished sixteen, a second seventeen, and the third eighteen descriptive units.

[10]For the coefficients and their formulae, as well as for general discussion, see Whipple, *Manual of Mental and Physical Tests*, Pt. II, 1915, 20ff.

[11]The issue is this: Shall we consider the total material as adequately represented by a complete, descriptive, verbal mosaic, in which any statement is regarded as of the same value as any other statement; or shall we attempt to weight the separate items in some such way as they are invariably weighted (by primary attention?) in everyday life?

[12]T. G. Hegge (Zur Frage der Bewertung von Aussagen bei Bildversuchen, *Zeitschr. f. angew. Psychol.*, 6: 1912, 51), also has used an empirical method for the determination of the number of items involved in a picture. He employed eight observers for this determination.

We decided to consider the material as naturally composed of eighteen phases—those phases which were identical in at least two of the three reports. An additional point, covering the duration of the action, was added to the others, making nineteen in all.

The eighteen phases were subdivided by the experimenter into as many approximately equivalent items as were neccessary to cover the particular phase completely. The number of items varied from three to seven for the different phases. The total number for all phases was eighty-five.

EMPIRICAL ANALYSIS INTO PHASES OF THE PRESENTED MATERIAL, WITH THE NUMBER OF SPECIFIC ITEMS FOR EACH PHASE.[13]

1. Man on wall; burglar crouching at gate. (3)
2. Man attacks burglar and gets revolver. (5)
3. Burglar rises to his knees. (5)
4. Man recognizes burglar. (5)
5. Man reflects. (3)
6. They talk to each other. (6)
7. Man looks at revolver, turns, and reflects. (4)
8. Man makes burglar get up. (5)
9. Burglar takes bag away and returns. (5)
10. Man talks to burglar about the revolver. (5)
11. Man directs the burglar to leave. (4)
12. Both go out together. (4)
13. Physical features of man. (4)
14. Clothing of man. (6)
15. Physical features of burglar. (4)
16. Clothing of burglar. (5)
17. Foreground and background. (7)
18. Scene at the right. (4)
19. Duration of the action.

The phases were scored equally. Since the number of items within each phase varied, the single items (part-phases) were not scored equally. The scores were determined upon a percentile basis. The total amount of the material, P, in the formula was taken as 100. Thus each of the nineteen equal phases received the value 5.25. Since the number of items in a single phase varied from three to seven, the values of the separate items varied accordingly from 1.75 to 0.75.

In scoring the answers to the question: "How long did it take to show the picture?", it was necessary to adopt an arbitrary scale. For times less than the minute required for the action, deductions directly proportional to the error were made, so that no time at all would have been scored zero. For estimates greater than a minute

[13] Numbers in parentheses are the numbers of items within each phase.
Twelve phases deal with the action and six with features that remain unchanged. This division into two classes of description was made by almost all observers.
The phases as given differ little from the three empirical lists. One of these lists united Phase 1 with 2 and 11 with 12; another united 8 with 9; the third combined 17 and 18, and separated 6 into two.

the deductions were decreased geometrically. The rule (perfect score, 5.25) was as follows: Score 5.25 for 54—67 secs.; score 4.20 for 42,—54 and 67—85 secs.; score 3.15 for 30—42 and 85—120 secs.; score 2.10 for 18—30 and 120—220 secs.; score 1.05 for 6—18 and 220—400 secs.; score 0 for 0—6 and 400 secs. and over.

Calculation of Coefficients of Report.—The percentages of items reported [n], of items that are rightly reported [n(r)], of items reported with certainty [c], of items that are certain and right [c(r)], of items whose correctness is attested under oath [a], and of items that are attested and right [a(r)] were first determined for the narrative and the deposition taken together—that is to say, an item reported in the one or in the other or in both was scored as if reported once. The averages for the groups are shown in Table I. The number of items reported in the narrative and in the deposition were also determined separately in order to provide a basis for the computation of the spontaneity of report, which is the ratio of these two values.[14]

TABLE I.

Average percentages of total material reported in narrative or in deposition (or in both) by four groups of subjects, 44 in all. Total number of items = P = 100%.

Group		Reported n	Reported and right n (r)	Certain c	Certain and right c (r)	Attested a	Attested and right a (r)
13 Women	Av.	38.2	25.1	35.3	23.7	25.8	18.8
	P. E.	3.1	4.8	3.5	3.1	3.0	3.5
11 Men	Av.	42.1	32.6	40.6	31.2	26.0	22.6
	P. E.	5.7	5.6	5.6	5.9	7.2	6.8
8 Girls	Av.	32.5	20.1	30.0	19.2	17.5	13.4
	P. E.	2.8	2.2	4.2	3.0	6.3	4.2
12 Boys	Av.	35.2	20.5	32.8	19.7	23.5	15.7
	P. E.	3.1	3.6	4.0	3.5	4.0	3.7

[14]Spontaneity of report—number of items reported in narrative ÷ number of items reported in deposition. In figuring the percentile values of items reported in narrative and in deposition the scores for each point had to be altered somewhat from their amounts when the narrative and deposition were taken together, for the narrative cannot be expected to include such additional points as are introduced by the implicative questions in the deposition, and the deposition does not, of course, cover many of the details which may appear in the narrative. When the deposition was treated separately, all the questions were considered as of equal value. It should be remembered that the spontaneity here is the ratio of the *percentages* of the items reported in the narrative and of those reported in the deposition, as based upon the total number of items in the narrative and in the deposition *respectively*. Since a complete answer to the interrogatory does not involve as much material as does a complete narrative, the probability of the spontaneity being grater than unity is less than would have been the case if absolute numbers of items had been used instead of percentages. The comparative status of different groups of subjects with respect to spontaneity of report remains, of course, unaffected.

The various coefficients were computed for each of the subjects separately. The averages and the probable errors [15] were then found for the various combinations of subjects according to age or sex, *i. e.*, for women, men, girls, boys, adults, children, females, and males.

TABLE II.

Average coefficients of report based upon narrative and deposition and computed from Table I.

Coefficient	Formula		Women	Men	Girls	Boys	Females	Males	Adults	Children
Range of report	n / P	Av.	.382	.421	.325	.352	.360	.384	.400	.341
		P. E.	.031	.057	.028	.031	.034	.049	.043	.059
Spontaneity of report	Note / p	Av.	.257	.291	.152	.145	.217	.215	.273	.148
		P. E.	.057	.065	.041	.044	.068	.073	.062	.042
Range of knowledge	n (r) / n	Av.	.669	.753	.622	.578	.651	.662	.708	.596
		P. E.	.059	.071	.056	.074	.060	.098	.076	.064
Accuracy of report	n (r) / c	Av.	.722	.761	.668	.658	.702	.703	.741	.662
		P. E.	.070	.071	.055	.058	.068	.079	.069	.059
Assurance	c / n	Av.	.916	.969	.918	.929	.917	.945	.938	.924
		P. E.	.038	.019	.073	.062	.052	.046	.041	.067
Reliability of assurance	c (r) / c	Av.	.688	.745	.642	.601	.671	.670	.714	.618
		P. E.	.057	.067	.062	.074	.062	.089	.068	.071
Warranted assurance	c (r) / n	Av.	.634	.742	.591	.556	.618	.645	.683	.570
		P. E.	.054	.069	.061	.069	.057	.105	.075	.067
Assured accuracy	c (r) / n (r)	Av.	.946	.980	.947	.963	.947	.971	.962	.957
		P. E.	.030	.013	.056	.034	.039	.024	.026	.042
Tendency to to oath	a / n	Av.	.665	.620	.495	.612	.600	.642	.643	.595
		P. E.	.065	.083	.128	.137	.106	.077	.075	.117
Warranted tendency to oath	a (r) / n	Av.	.478	.536	.392	.414	.445	.472	.504	.405
		P. E.	.062	.084	.109	.066	.073	.089	.076	.082
Unwarranted tendency to oath	a(w) / n	Av.	.177	.084	.103	.198	.155	.170	.139	.190
		P. E.	.079	.032	.080	.096	.068	.064	.038	.078
Reliability of oath	a (r) / a	Av.	.719	.869	.799	*.667	.748	*.764	.788	.716*
		P. E.	.092	.066	.084	.090	.089	.103	.087	.098

*One indeterminate case omitted.

[15] A Wales adding machine was used for figuring the averages. The probable errors were obtained by use of a Comptometer after K. Dunlap's method for the determination of an M. V. (*Psych. Rev.*, 20: 1913, 154ff.). The addition and subtraction were done directly on the machine, the division on a slide-rule, and the final conversion of the M. V. into the P. E. by a muliplication table for 845 (since P. E.=.845 M. V.).

Since the probable errors are large and the significance of the relationships between the various groups is not immediately apparent, the differences for six pairs of groups (women-men, girls-boys, females-males, women-girls, men-boys, and adults-children) and the probable errors of these differences [16] were computed and are shown in Table III. This table gives in heavy-faced type all those differences in which the probability that the difference is in the direction indicated is greater than 100 to 1. Those differences in which the probability is greater than 10 to 1, but less than 100 to 1, are in italics.[17]

TABLE III.

COMPARISON OF GROUPS. *Differences* between the various groups of subjects with respect to the different coefficients and based on narrative and deposition. "+" indicates that the difference is in the direction indicated at the top of the column; "–" that the difference is in the opposite direction. (*E. g.*, "Women minus men = .150" means that the men exceed the women by .150 with respect to that coefficient.) *Italics* indicate that the probability that the difference is in the direction indicated is greater than 10 to 1; HEAVY-FACED TYPE, that the probability is greater than 100 to 1.

[16] Whipple, *op. cit.*, Pt. I, 1914, p. 27.

[17] By the use of E. L. Thorndike's tables (*Theory of Mental and Social Measurements*, 1913, 200), it can be shown that, when a mean is 3.45 times as great as its P. E., there is only one chance in 50 that a single deviation will exceed the mean. The probability that a difference will be in the direction indicated is twice the probability that it will not vary by an amount greater than itself, since the variation is just as likely to be an increase as a decrease. This probability is 10,000 when the ratio of the difference to its P. E. is 5.50; 1000, when the ratio is 4.60; 100, when the ratio is 3.45; and 10, when the ratio is 1.90.

Coefficient	Formula		Women minus Men	Girls minus Boys	Females minus Males	Women minus Girls	Men minus Boys	Adults minus Children
Range of report	n/P	Av.	*−.039*	*−.027*	*−.024*	**+.057**	**+.069**	**+.059**
		P. E.	.019	.013	.008	.013	.019	.017
Spontaneity of report	Note/p	Av.	−.034	+.007	+.002	**+.105**	**+.146**	**+.125**
		P. E.	.023	.019	.021	.021	.021	.016
Range of knowledge	n (r)/n	Av.	**−.084**	+.044	−.011	+.047	**+.175**	**+.112**
		P. E.	.023	.027	.024	.024	.026	.021
Accuracy of report	n (r)/c	Av.	−.039	+.010	−.001	+.054	**+.103**	**+.079**
		P. E.	.029	.026	.022	.027	.028	.019
Assurance	c/n	Av.	**−.053**	−.011	−.028	+.002	+.040	+.014
		P. E.	.012	.031	.015	.028	.019	.017
Reliability of assurance	c (r)/c	Av.	−.057	+.041	+.001	−.046	**+.144**	**+.096**
		P. E.	.026	.031	.023	.027	.029	.021
Warranted assurance	c (r)/n	Av.	**−.108**	+.035	−.027	−.043	**+.186**	**+.113**
		P. E.	.026	.029	.025	.026	.029	.021
Assured accuracy	c (r)/n (r)	Av.	**−.034**	−.016	*−.024*	+.001	+.017	+.005
		P. E.	.009	.002	.010	.021	.010	.011
Tendency to oath	a/n	Av.	+.045	*−.117*	−.042	**+.170**	+.008	+.048
		P. E.	.031	.060	.028	.045	.047	.030
Warranted tendency to oath	a (r)/n	Av.	−.058	−.022	−.027	*+.086*	**+.122**	**+.099**
		P. E.	.031	.043	.024	.042	.032	.024
Unwarranted tendency to oath	a(w)/n	Av.	**+.093**	*−.095*	−.015	*+.074*	**−.114**	*−.051*
		P. E.	.024	.040	.020	.036	.029	.019
Reliability of oath	a (r)/a	Av.	**−.150**	**+.132**	−.016	*−.080*	**+.202**	*+.072*
		P. E.	.032	.039	.029	.039	.033	.030

In order to study suggestibility, the six implicative questions of the interrogatory [18] were isolated and the entire computation outlined above was carried through for them separately. These questions were considered as equal in value. The final results, in the form of differences between groups, are shown in Table IV.

[18] Nos. 9, 14, 16, 16, 18 and 24.

TABLE IV.

SUGGESTIBILITY. *Differences*, based on the implicative questions taken alone, between the various groups of subjects with respect to the different coefficients. Symbolism the same as in Table III.

COEFFICIENT	Formula		Women minus Men	Girls minus Boys	Females minus Males	Woemn minus Girls	Men minus Boys	Adults minus Children
Range of report	n/P	Av.	**−.232**	−.065	*−.160*	−.087	+.080	−.018
		P. E.	.054	.066	.046	.072	.046	.045
Range of knowledge	n (r)/n	Av.	*−.156*	+.106	−.003	+.114	+.076	**+.247**
		P. E.	.072	.066	.061	.085	.047	.056
Accuracy of report	n (r)/c	Av.	*−.145*	+.082	−.028	+.073	**+.310**	+.095
		P. E.	.076	.102	.069	.105	.072	.054
Assurance	c/n	Av.	*−.125*	+.042	−.046	−.043	**+.133**	+.048
		P. E.	.054	.064	.046	.077	.034	.094
Reliability of assurance	c (r)/c	Av.	*−.171*	+.156	−.002	+.053	**+.380**	*+.226*
		P. E.	.073	.084	.064	.097	.054	.094
Warranted assurance	c (r)/n	Av.	*−.211*	*+.166*	−.011	−.040	**+.417**	**+.235**
		P. E.	.074	.077	.055	.087	.063	.051
Assured accuracy	c (r)/n (r)	Av.	−.080	+.071	−.019	−.005	+.056	−.017
		P. E.	.043	.041	.029	.045	.037	.027
Tendency to oath	a/n	Av.	+.091	−.190	−.031	+.182	−.099	+.027
		P. E.	.097	.118	.074	.112	.103	.077
Warranted tendency to oath	a (r)/n	Av.	+.012	+.067	+.058	+.125	**+.180**	*+.159*
		P. E.	.067	.068	.057	.081	.051	.052
Unwarranted tendency to oath	a(w)/n	Av.	*+.097*	**−.258**	−.090	+.065	**−.299**	*−.133*
		P. E.	.044	.061	.048	.064	.040	.051
Reliability of oath	a (r)/a	Av.	−.084	**+.382**	*+.149*	−.053	**+.413**	+.131
		P. E.	.083	.111	.077	.122	.074	.070

RESULTS.

In the following discussion no account is taken of any differences between groups where the probability that the difference is of the order shown is less than 10 to 1. It is assumed that some significance attaches to a value when its probable correctness [19] is over 10. No especial emphasis, however, is placed upon any figure unless its probable correctness is over 100. The results have been evaluated with those cases in mind also, in which the probable correctness is greater than 1,000, 10,000 and 100,000.

[19] The term, "probable correctness," is used of differences in this paper to denote the probability that a value will not vary in the direction of zero by more than its magnitude, *i. e.*, that its sign will remain unchanged.

Sex-Differences.—Inspection of the third column of Table III fails to reveal any marked general sex-differences. The differences are very small in relation to the size of the coefficients, and their significance is low; only two show a probability of correctness greater than 10. These two are range of report and assured accuracy, in which the males exceed the females.

If we turn to the first two columns, however, we find more positive results. There is evidence that the boys exceed the girls in range of report, in tendency to oath, and also in unwarranted tendency to oath, and that the girls exceed the boys in reliability of oath. The significance of none of these coefficients is very high, although the last is undoubtedly reliable. With men and women the differences are more marked. The greatest difference occurs in the unwarranted tendency to oath, in which the women exceed the men. The men exceed the women in range of report, range of knowledge, assurance, warranted assurance, assured accuracy, and reliability of oath. In all of these factors, except the range of report, the probable correctness is greater than 100.

It appears, then, that there is little difference apparent, with the material used, between boys and girls, whereas there is quite a marked superiority of the men over the women with respect to six (perhaps seven) coefficients. This conclusion accords with the general psychological principle that even those mental sex-differences, which are large in adults, are relatively slight in childhood.

Age—Differences.—With respect to age (see last three columns of Table III) we find more differences, greater differences, and differences of greater significance. The most positive differences occur between the men and the boys. Here we find the men superior to the boys in all cases but three—assurance, assured accuracy, and tendency to oath—which show no significant difference. The probable correctness is large for all of the coefficients (over 10,000 for four). The differences between the women and the girls is less consistent. The women exceed the girls with respect to both range and spontaneity of report. Examination of the coefficients for attestation reveals that the women exhibit a greater tendency to oath, but that, although they show a greater warranted tendency to oath, they also show a greater unwarranted tendency and a less reliability. The women are not so cautious as the girls, and for that reason are less accurate in reporting under oath.

Since sex-differences are apparent, a general comparison of adults with children loses much of its significance. There is no doubt, however, that the adults must be accredited with superiority over the children in range and in spontaneity of report.

Suggestibility.—For the results based upon the implicative questions the reader should consult Table IV.

A sex-difference in suggestibility appears among the adults. The men exceed the women in range of report, range of knowledge, accuracy of report, assurance, reliability of assurance, warranted assurance, and unwarranted tendency to oath. With one exception (range of report) the probable correctnesses are not high, but, since the differences themselves are very large and are all in the same direction, there seems to be ample justification for the conclusion that a sex-difference exists, and that the men were less suggestible than the women.

There is little discrepancy between the replies of the girls and those of the boys to the implicative questions, except in regard to attestation. The girls show a much less unwarranted tendency to oath and a much greater reliability of oath than the boys. This difference reflects the fact that the girls were more cautious in taking oath than the other groups. One girl even refused to take oath to any statement at all.

The principal difference with respect to age for the implicative questions occurs between men and boys. There is no indication of any difference at all between women and girls; there is not even a single significant coefficient. The men, however, exceed the boys in accuracy of report, assurance, reliability of assurance, warranted assurance, warranted tendency to oath, unwarranted tendency to oath, and reliability of oath. In no case do the boys exceed the men. The values, however, although somewhat greater, show a great similarity to those obtained for the general material, so that it is probable that the superiority of the men is due principally to their general superiority and not to any particularly high degree of suggestibility on the part of the boys.

Conclusions.

Before drawing conclusions as to the relative reliability of the various groups, it is necessary to recall that the constitution of the group of men and that of the group of women were not exactly the same. The men included more graduate and fewer under-graduate students than did the women, they were slightly older on the average, and several of them were familiar with the general form of the experiment. This disparity in favor of the men, which was not thought of at the time as prejudicial, now coincides—unfortunately for the formulation of positive conclusions—with the fact that in general the men appear to be superior as witnesses to both women and boys, whereas between women and girls and between girls and boys there

is a much less striking difference. That a knowledge of the type of experiment on the part of the greater number of men was not likely to produce a decided general difference may perhaps be maintained on the basis of the reports of Baade [20] and of Lipmann.[21] Baade found no effect from a repetition of the experiment, and Lipmann found with repetition no improvement other than a tendency to caution. As the men in this study were more cautious than the women there may be reason to look to their training for the explanation of this trait,[22] but it is doubtful if their knowledge of conditions could have had any other effect. It is further improbable that the general superiority of the men over the women can be merely a function of age, for the women do not show as great a difference when compared with the girls. Thus we are constrained to accept the differences as significant.

We may state, then, that there is considerable presumption of a difference in excellence of report between men and women in favor of the former; that no such sex-difference is apparent in childhood; and that the reports of adults are more adequate and accurate than those of children. The sex-difference which appears with advancing age reflects a greater change in the males than in the females; the age-difference is more marked in the former than in the latter.

Striking specific differences occur in range of report and in spontaneity of report, especially in the latter. In these respects the adults are markedly superior to the children. Furthermore, the range of report for males is somewhat greater than that for females, a fact which again reflects the greater age-difference for males.

With regard to attestation we may note that the men were the most careful and accurate witnesses. The girls came next, being especially cautious and much more careful than the boys.[23] The women made the worst showing, since they combined a high degree of willingness to take oath with great unreliability of oath.

SUMMARY.

1. The moving picture presents a satisfactory and an easily and accurately controlled form of event-test.

[20] W. BAADE (Aussage uber die physikalische Demonstrationen, 1 Abhl., *Zeitschr. f. angew. Psychol.*, 4: 1911, 189ff.

[21] O. LIPMANN, Aussage uber die physikalische Demonstrationen, 2 Abhl., *ibid.*, 4: 1911, 312ff.

[22] A. FRANKEN (Aussageversuche nach der Methode der Entscheidungsund Bestimmungsfrage bei Erwachsenen und Kindern, *ibid.*, 6: 1912, 174ff.), finds that his more gifted subjects excel through a greater tendency to caution.

[23] H. B. L. Vos (Beiträge zur Psychologie der Aussage bei Schulkindern: Analyse d. Aussage über eine gehörte Erzählung, Eigenbericht, *ibid.*, 4: 1911, 37 5ff.), alsofinds girls more cautious than boys.

2. In the material of such a test the total number of items is best determined by an empirical division into phases made by several observers.

3. In the present experiments a sex-difference, which did not exist in childhood, was found in adulthood. The men were more accurate and careful reporters than the women.

4. A general age-difference, in favor of the adults, was especially evident in comparison of men with boys. The difference between the women and the girls was much less marked—a relationship which is consistent with the sex-difference in adulthood.

5. The men were less suggestible than were the women or the children of either sex.

6. The children showed less spontaneity of report and less accuracy in general. They showed, however, considerable caution under oath, much more than was exhibited by the women.