

https://dspace.org/

Curiosity in Africans and Europeans

Foden, Bernardina

2015-01-21
http://hdl.handle.net/10539/16660

Downloaded from a DSpace repository. https://dspace.org/

(inducing a longer looking time at these items) than on the incongruous items which were placed at the end of the test. In other studies on the effects of incongruity (Berlyne, 1960, 1963, 1966), the incongruous pictures were either interspersed between the congruous pictures in the series, or else presented simultaneously with their congruous counterparts, while fixation of eye movements were recorded. Novelty of the pictures, would therefore, not have played as great a part in these studies.

(ii) Most of the stimulus material was of an extremely simple nature, but closer inspection of stimulus 11 and 12 revealed that they were more complex than the other stimuli in the series. Figure 8 shows that subjects in both cultural samples looked longer at these pictures than at the other pictures in the series. Thus complexity may also have confounded the results.

Item Equivalence: A correlation of 0,62 was obtained between the 2 cultural samples on the scores on the 18 items in the test. Correcting for attenuation, (cf. Lord & Novick, 1968 p.69) this value becomes 0,88 thus providing some evidence for the idea that the 2 samples responded to the test in a similar manner. Some support for the item equivalence can also be found through inspection of Figure 8.

4. The Novelty Test

The two measures of novelty on each of the 4 series of slides were obtained for each subject, on the two testing

days.

(a) The length of time the subject looked at the novel stimulus multiplied by 100 and divided by the total length of time that the subject looked at the novel stimulus plus the 2 adjacent stimuli.

i.e. $\frac{\text{L.T. (Novel Stimulus)} \times 100}{\text{L.T. (Novel Stimulus + 2 Adjacent Stimuli)}}$

(b) The length of looking time at the first stimulus in a series multiplied by 100 and divided by the total looking time for the first, second and third stimuli combined.

i.e. $\frac{\text{L.T. (1st Stimulus)} \times 100}{\text{L.T. (1st + 2nd + 3rd Stimuli)}}$

Means, S.D.s and reliabilities of these scores are presented in Table 12. t-tests showed that the means of scores attained on (a) and (b) were significantly greater than 33.1/3% for both groups on both testing days ($p < 0,01$), i.e. subjects spent significantly more time looking at the novel stimulus in a series and at the first stimulus of a series than at the immediately adjacent stimuli.

A combined novelty score was obtained by adding the scores obtained on (a) and (b) on each of the two testing days. These scores yielded a reliability of 0,73 for the European sample and 0,51 for the African sample. An analysis of variance, (nested design: cf. Brownlee, 1965,p. 538) (Table 13), carried out the combined novelty score, showed no significant effect due to culture of the sample or day of testing.

Table 12

Novelty Test: Means, S.Ds and Reliabilities (over the two days of testing, corrected for length according to the Spearman-Brown formula) of scores attained on (a) and (b) and on the combined novelty measure.

	(a)			(b)			Combined Novelty Score		
	Mean	S.D.	Rel.	Mean	S.D.	Rel.	Mean	S.D.	Rel.
Africans Day 1 n = 50	43,360	8,405	0,51	38,840	7,181	0,43	82,015	13,096	0,51
Day 2	39,520	7,327		37,360	5,190		76,915	9,404	
Europeans Day 1 n = 47	42,787	8,145	0,64	38,574	4,772	0,29	81,182	10,339	0,73
Day 2	43,234	11,109		37,809	5,261		80,155	12,133	

Table 13

Novelty Test: Analysis of Variance with culture (2x), subject within culture (47x) and Day of Testing as factors.

Source of Variation	Denomination	Mean Square	df num.	df denom.	F
Culture, C	S(C)	51,841	1	92	0,135
Subjects within culture, S(C)		383,557	92		
Day, D	DxS(C)	472,899	1	92	3,37
C x D	DxS(C)	217,665	1		1,55
D x S(C)		140,358	92		

Furthermore, t-tests showed no significant difference in the total length of looking time at all the stimuli between the 2 cultural samples on either day of testing.

For the purpose of graphic representation of the results, the 6 most relevant stimuli in each series (i.e. the 1st, 2nd, 3rd stimulus, the stimulus immediately preceding the novel stimulus, the novel stimulus and the stimulus immediately after the novel stimulus) for each subject were ranked in terms of length of looking time (1 = longest, 6 = shortest looking time). The results are presented in Figure 9, which shows the mean rankings of looking time averaged over subjects and series for the 6 relevant stimuli in both cultural samples. The effect of the novel stimulus and the first stimulus in the series is clearly illustrated; subjects looked longer at both these two stimuli.

If the length of looking time at the first stimulus in a series is regarded as a measure of absolute novelty and the length of looking time at the novel, but similar type of stimulus in the middle of the series, as a measure of relative novelty then this study provides strong confirmatory evidence for Berlyne's (1960, 1963, 1966) idea that both these types of novelty are prepotent in eliciting attention, orientation and exploration, and for other studies carried out in the field which have found similar results.

Item Equivalence: It can be seen from Figure 9, that the shapes of the graphs for the two cultural groups are highly

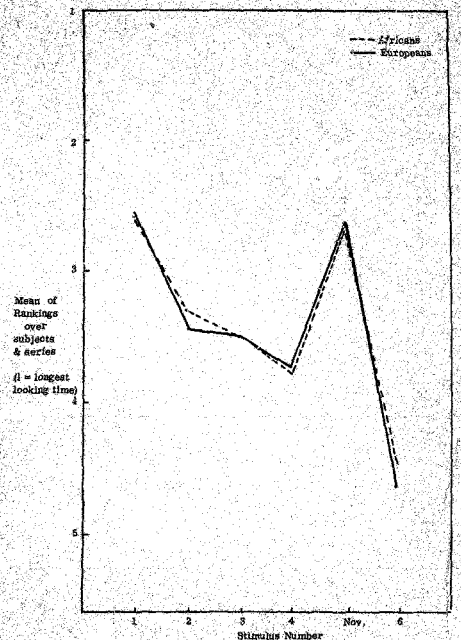


Figure 9. Novelty Test: Mean rankings of looking time averaged over subjects and series.

similar; in fact the correlation of the mean rank order between the two samples was 0,99, thus lending strong support to the notion that the test is item equivalent in the two samples.

5. The Conflict Reaction Time Test

On each of the two testing sessions, the following scores were obtained for each subject:

(i) The Median of his Simple Reaction Time Scores (SRT),

(ii) The Median of all reaction time scores (correct and incorrect for (a) Conflict RT Level 1,
(b) Conflict RT Level 2,
(c) Conflict RT Level 3,

(iii) The Median RT for correct responses only for each of the three levels of conflict,

(iv) The Median of the error responses for each of the three levels of conflict.

(v) The number of error responses made at each of the three levels of conflict.

Means and standard deviations of these medians are shown in Table 14.

The results of t-tests carried out on the SRT data show that RTs for the European sample were significantly faster than for the Africans on both days of testing ($p < 0,01$). No significant difference in median RT between the two days of testing was found for either group.

Table 15 shows the results of the analysis of variance (Nested design, cf. Brownlee, 1965, p.538) carried out on

the medians of all correct responses for the conflict RT data. (It was considered unnecessary to carry out additional analyses on the combined error and correct RTs in view of the close correspondence between "all responses" and "correct responses" (cf. Table 14)). All major factors (Culture, Subjects within culture, level of difficulty and day of testing)were found to be significant. The European sample showed considerably shorter RTs than the African sample at all levels of difficulty.

Table 14

Conflict Reaction Time Test: Means and S.D.s of Median RTs.

			Africans		Europeans	
			Day 1	Day 2	Day 1	Day 2
SRT		Mean	25,660	24,620	18,574	17,309
		S.D	7,879	8,146	3,082	2,902
	All	Mean	62,260	56,830	50,830	51,011
	resp	S.D	12,540	10,827	9,152	8,598
	Corr.	Mean	62,260	56,970	52,277	51,213
CRT	resp.	S.D	12,250	10,709	9,204	6,716
Level	Error	Mean	45,570	36,240	39,351	36,745
1	resp.	S.D	34,782	27,227	19,168	22,256
	All	Mean	63,590	58,750	54,904	53,479
	resp.	S.D	12,562	10,335	9,323	7,704
CRT	Correct	Mean	63,630	58,830	55,277	53,543
Level	resp.	S.D	12,602	10,273	9,204	7,524
2	Error	Mean	47,240	37,860	45,777	29,032
	resp.	S.D	35,980	31,924	31,992	34,481
	All	Mean	66,740	61,910	63,277	59,628
	resp.	S.D	13,182	11,215	13,256	10,396
CRT	Corr.	Mean	66,550	61,990	63,138	59,415
Level	resp.	S.D	12,957	11,105	12,718	9,682
3	Error	Mean	63,940	54,290	65,667	65,457
	resp.	S.D.	38,921	31,974	23,316	26,287

Table 15

Conflict Reaction Time: Analysis of variance with Culture (2x), Subjects within Culture (47x), Treatment (3x) and Day of Testing (2x) as factors.

Source of Variation		Denominator	Mean Square	df.num.	df.denom.	F
Culture C	a	e + b	5614,57	1	30	4,2 *
Subjects within culture S(C)	b	h	1073,69	92	184	25,44**
Treatment,T	c	h	2509,98	2	184	59,48**
Day, D	d	g + n - k	1930,60	1	44	16,05**
C x T	e	h	305,88	2	184	7,284**
C x D	f	j	336,28	1	2	10,14
T x D	g	k	15,59	2	184	-
T x S(C)	h	-	42,24	184		-
D x S(C)	i	k	137,67	92	184	4,18**
C x T x D	j	k	33,18	2	184	1,01
TDS (C)	k		32,97	184		-

Inspection of Figure 10 reveals that there was considerable improvement in RT from Day 1 to Day 2 in the African sample, whereas performance in the European sample remained more or less the same on the two testing sessions. Prolonged training would probably have resulted in further improvements in the African group. Figure 10 also shows a definite order of difficulty, slowest RTs being made to Conflict Level 3 and fastest to Conflict Level 1. This order is more marked in the European than in the African sample as shown by the Culture x Treatment interaction (Table 15). t-Tests carried out on the Data revealed no significant differences in Median RT between the three difficulty levels for the African sample on the first testing session, while only the difference in RT between Level 1 and Level 3 was significant on the second session ($p < 0,05$). For the European sample the difference between Level 1 and Level 3 was significant on both testing sessions ($p < 0,01$).

Numerous experiments (Pew,1969; Rabbit, 1967; Rabbit & Vyas, 1970; Schouten & Bekker, 1967) have found that in Choice Reaction Time experiments there is a "trade-off" between speed and accuracy and that speed of responding tends to be inversely related to the number of mistakes the subject makes. IN view of the small number of errors made by each subject (cf. Table 16) and the close relationship between error RT and correct RT for each subject at the three levels of difficulty, the variable "correct-incorrect could not be included in the above analysis of variance, neither could t-tests be applied. Consequently a number

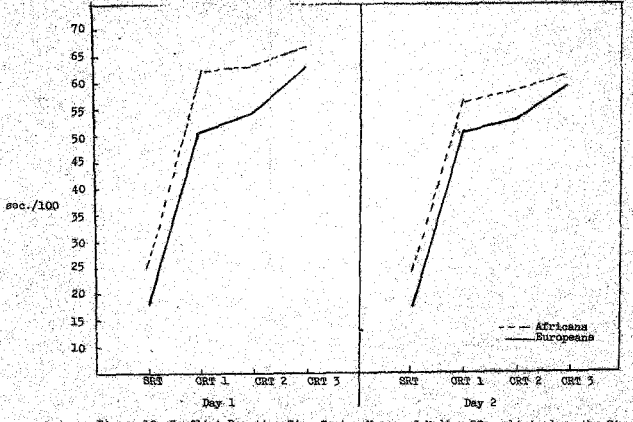


Figure 10. Conflict Reaction Time Test: Means of Median RTs obtained on the Simple Reaction Time Task and Conflict Reaction Time Tasks for the 3 levels of intensity differences between stimuli (all responses).

Table 16

Conflict Reaction Time : Means and S.D.s of the Number of Errors.

		Africans		Europeans	
		Day 1	Day 2	Day 1	Day 2
CRT Level 1	Mean	2,280	1,880	4,660	2,021
	S.D	2,441	2,210	3,377	1,595
CRT Level 2	Mean	1,780	1,460	2,638	1,149
	S.D	1,810	2,261	2,298	1,474
CRT Level 3	Mean	3,340	2,640	5,128	4,213
	S.D	2,861	2,655	2,787	2,331

Table 16(a)

Conflict Reaction Time : t-values for differences between Mean no. of Errors in the two cultural groups. F-values for $p \leq 0,05$ are also given.

		Africans vs. Europeans.		
		df	t	F
CRT level 1	Day 1	84,96	3,96 **	1,92 **
	Day 2	90,88	0,36	1,92 **
CRT level 2	Day 1	95	2,05 *	1,61 **
	Day 2	86,35	0,81	2,35
CRT level 3	Day 1	95	3,11 **	1,05
	Day 2	95	03,09 **	1,30

of Wilcoxon Matched Pairs tests (Table 17) were carried out to test whether Median Correct RT differed significantly from Median Error RT. Median Error RTs were found to be significantly faster than Median Correct RTs in only two cases - Conflict RT Level 1 and Level 2 for the European sample on the second testing session.

Statistics on the distribution of errors are given in Tables 16 and 16 (a). The mean number of errors, as determined by t-tests was significantly greater in the European sample for all levels of difficulty on the first testing session, and for the most difficult level during the second session. A correction for inaccuracy could not be applied, however, as the small number of errors made reliable estimates of error probabilities virtually impossible. Thus although the European sample responded considerably faster than the Africans, the evidance is contaminated by the "accuracy versus speed trade-off".

Studies on differences in reaction time between the two cultural groups abound (cf.Bligh, 1967; Poortinga, 1971). However, since the purpose of this test was to find a measure of conflict in both groups, and absolute differences in RT do not reflect differences in conflict, these will not be discussed further.

Conflict presumably increases when stimuli becomes more similar as generalized response tendencies evoked by the stimuli show more overlap (Berlyne, 1957a, 1963). Both the increase in RT and the increased number of errors and the relationship between speed and accuracy with

increased difficulty in discrimination can be considered as indices of conflict, assuming that perceptual discrimination is the same for subjects in both groups.

Table 17

Conflict Reaction Time Test: Results of Wilcoxon Matched Pairs Tests for differences between Median correct and Median Error RTs.

	Africans (N=50)		Europeans (n=47)	
	Day 1	Day 2	Day 1	Day 2
T	317	234	447	222
CRT Level 1 No of Pairs	36	35	43	38
z	-2,590	-1,335	-0,314	-2,154
Significance	79,55%	18,19%	75,36%	3,13%
T	330	186	381	60
CRT Level 2 No. of Pairs	38	30	40	24
z	-5,87	-0,969	-0,390	-2,571
Significance	55,70%	33,37%	69,67%	1,01%
T	414	354	384	453
CRT Level 3 No. of Pairs	42	40	44	43
z	-0,469	-0,759	-1,301	-0,241
Significance	63,91%	44,76%	19,32%	80,92%

In view of the small number of errors, and the controversy surrounding the rel[illegible] between speed and accuracy (cf. Schouten & Bekk[illegible]97) it was decided to exclude error RTs and to use only "Median correct RTs" in the

determination of the measure of conflict. A regressi line was fitted to the median correct RTs for Conflict Level 1, Level 2 and Level 3 for each subject by the method of least squares and the slope of this line was taken as a measure of the subject's conflict, on the assumption that the steeper the slope of the line, the greater level of conflict for that subject. Means, standard deviations and reliabilities (based on the correlation between the first and second testing session and corrected for length according to the Spearman-Brown formula) of scores are given in Table 18. It can be seen that the scores for the European sample are fairly reliable, while those for the African sample are highly unreliable, so that the test cannot be considered item equivalent for the two groups since the measure is so unreliable for the African sample. In addition there are indications that perceptual discrimination may not be the same for the African and European samples, and that the conflict score in the European sample reflects ability to discriminate as well as conflict per se. Inspection of Table 14 reveals that mean error RTs were shorter than correct RTs for the African group at all levels of conflict and for the European group on conflict RT level 1 and 2. On the third level of conflict, however, error RTs were longer than correct RTs for the European group on both days of testing. This points to the fact that the Europeans had greater difficulty in discriminating between stimuli than the African group. It is therefore regarded as likely that the conflict score of the European sample

Table 18

Conflict Reaction Time Test: Means, S.D.s and Reliabilities (corrected for length of test according to the Spearman-Brown formula) of Conflict scores.

	Conflict Sc.			Conflict Sc./SRT			Conflict Sc./CRT L.l.		
	Mean	S.D.	Rel.	Mean	S.D	Rel.	Mean	S.D.	Rel
Africans (n=50) Day 1	0,95	1,76	-0,35	4,94	9,01	-0,30	2,04	1,70	-0,37
Day 2	1,10	1,01							
Europeans (n=50) Day 1	2,42	2,13	0,70	1,08	7,89	0,70	3,56	2,43	0,66
Day 2	1,82	1,29							

reflected not only conflict but also the ability to discriminate between stimuli.

In a further analysis, an attempt was made to correct for the Law of Initial Values. Differences in initial RT might affect the slope of the regression line. Each subject's score was therefore divided by: (a) his simple RT and (b) his conflict RT Level 1 score on that testing session. The results are presented in Table 18. Reliabilities for the African sample are again negative, indicating that these scores cannot be regarded as a measure of conflict in this group.

6. The Curiosity Questionnaire

Subjects were given a score of 1,2,3 or 4 on each of the 48 items in the questionnaire (4 indicating strong agreement with the alternative pertaining to high curiosity, 1 strong agreement with the alternative pertaining to low curiosity). Means and standard deviations of scores are given to Table 19. An item analysis carried out on the data yielded high values of KR 20 in both groups. t-Tests showed no significant difference in curiosity score between the African and European samples.

Item Equivalence : As a check on the item equivalence of the test, the average score obtained on each of the 48 items was correlated in the two samples. A correlation of 0,45 was obtained. Thus while it can be argued that the test probably does measure the same construct in the two samples to some extent, it can still be possible that

perhaps not all the items in the test are perceived in the same way by the two samples.

In order to discover whether "culture" did indeed have an important effect on the mode of answering the questionnaire, the average scores obtained by the African and European samples on the 48 items in the final questionnaire were correlated with the average scores obtained on these same 48 items by the African and European students who completed the preliminary version of the questionnaire. Inspection of Table 20 shows that intercorrelations were significantly higher within than between the cultural groups.

Table 19.

Curiosity Questionnaire: Means, Standard Deviations and Reliabilities of scores.

	Mean	S.D	Reliability (KR 20)
Africans (n=50)	137,54	16,36	0,91
Europeans (n=47)	141,89	17,45	0,88

Thus, although no significant difference in curiosity scores between the two cultural samples was obtained, Africans and Europeans scored highly on different items in the questionnaire and the questionnaire

Table 20

Curiosity Questionnaire : Correlations of the average scores obtained on the final version of the questionnaire between the African sample (n = 50), the European sample in our present study (n = 47), African male students from the University of the North (n = 155), African female students from the University of the North (n = 72), European male students (n = 33) and European female students (n = 277) from the Johannesburg College of Education.

	African M Fort Hare	African M U of N'th	African F U of N'th	European M Wits.	European M J.C.E	European F J.C.E
African M (Fort Hare)	1,00					
" M (U.of N'th)	0,83	1,00				
" F (U.of N'th)	0,85	0,95	1,00			
Europ. M (Wits.)	0,45	0,51	0,48	1,00		
" M (J.C.E.)	0,51	0,57	0,53	0,73	1,00	
" F (J.C.E.)	0,64	0,67	0,71	0,76	0,74	1,00

can, therefore not be considered item equivalent in respect of the two samples.

7. The Self-Rating Scale

Subjects were given scores ranging from 1 to 7 on each pair of adjectives. (1 indicating highest ranking and 7 lowest ranking on scores pertaining to activity and curiosity). An item analysis was carried out on the data obtained in each group; KR 20 values are given in Table 21 together with the means and standard deviations obtained in each sample. No significant difference in scores between the two groups was found.

Item Equivalence : The average score on each item in the scale was correlated for the two samples. The correlation of 0,81 obtained between the African and European samples can be regarded as strong confirming evidence for the comparability of the scale. Africans and Europeans, then, seem to perceive themselves in the same way as regards curiosity and activity.

8. The Intelligence Tests

Means and standard deviations of the distribution of scores are presented in Table 22, together with the odd-even split-half reliabilities, corrected for length according to the Spearman-Brown formula.

The European sample scored significantly higher on all 4 tests ($p \leq 0,01$). As mentioned earlier, two scores were obtained on the Locations and Letterset tests. Correlations between the first and second scores

Table 21

Self-Rating Scale : Means, S.Ds and Reliabilities of scores.

	Mean	S.D	Reliability
Africans (n=50)	92,38	15,62	0,84
Europeans (n=47)	94,47	20,53	0,88

obtained on these tests were fairly high: 0,91 and 0,63 on the Letterset test and 0,72 and 0,73 on the Locations test for the African and European groups respectively. The lower correlation of 0,63 obtained by the European sample between the two scores obtained on the Letterset test is probably due to the fact that with a 14 minute time limit, the test was far too easy for this group.

The intercorrelations between the various intelligence tests can be found in Table 23. Although both the Letterset and Locations tests supposedly measure inductive reasoning, correlations between these two tests were low, especially for the African sample. In view of the low intercorrelations between the Locations Test and the other intelligence tests, this test was excluded in the determination of the discriminant validity of the curiosity construct and the analysis of the functional equivalence of the curiosity tests (cf. Ch.V).

Table 22

The Intelligence Tests: Means, S.D.s and odd-even split half reliabilities.

	Africans (n = 50)	Europeans (n = 47)
Mental Alertness Test:		
Mean	24,48	49,85
S.D.	6,96	7,01
Rel.	0,87	0,84
Blox:		
Mean	15,90	30,91
S.D.	7,82	7,90
Rel.	0,84	0,[illegible]
Locations (6 mins):		
Mean	2,69	4,80
S.D.	1,83	2,01
Rel.	0,56	0,58
Locations (12 mins):		
Mean	3,99	7,36
S.D.	2,66	2,88
Rel.	0,54	0,67
Letterset (7 mins):		
Mean	6,48	10,82
S.D.	2,63	2,07
Rel.	0,51	0,77
Letterset (14 mins):		
Mean	8,32	13,01
S.D.	3,09	1,69
Rel.	0,58	0,45

Table 23

The Intelligence Test: Intercorrelation of Scores.
African scores are above the
diagonal, European scores below.

	(1)	(2)	(3)	(4)	(5)	(6)
Ment. Alert, (1)	1,00	0,37	0,38	0,37	0,10	0,41
Blox (2)	0,44	1,00	0,46	0,44	0,04	0,01
Lett.set.(7mins)(3)	0,51	0,42	1,00	0,91	-0,07	0,03
Lett.set(14mins)(4)	0,52	0,43	0,63	1,00	-0,07	0,07
Loc. (6 mins) (5)	0,23	0,01	0,27	0,27	1,00	0,72
Loc.(12 mins) (6)	0,44	0,08	0,30	0,36	0,73	1,00

9. The Biographical Questionnaire

Scores on the questionnaire ranged from 1,6 to 5,5 with mean of 3,45 and a standard deviation of 1,03. The reliability of the questionnaire was low, 0,56 as compared with a value of 0,89 obtained by Grant in the original version of the questionnaire. The relationship between scores on the questionnaire and scores on the curiosity tests is discussed in the next chapter, but must be treated with reservation in view of the small standard deviation of scores and the resultant low reliability of the test.

10. Summary and Conclusions

With regard to the cross-cultural comparability of the individual curiosity tests, results are encouraging. Analysis of item equivalence by means of correlations of item scores in each test for the two groups on the Novelty, Incongruity, and Uncertainty tests, the Curiosity Questionnaire and Self-Rating scale and by comparison of the pattern of item preference for the Representational and Non-Representational Complexity tests,reveals that four of the tests (Novelty, Incongruity, Representational Complexity and the Self-Rating Scale) can be considered to be comparable,in that subjects in the two groups responded to the items in a similar way. Although a correlation of 0,80 between the two groups was obtained on scores on items in the Uncertainty test

pointing to item equivalence of the test in the two samples, this result must be treated with some reservations in view of the differences in reliability of the test in the two samples. "Culture of the sample" may well have been responsible for the higher reliability of scores in the African sample.

Of the three tests in the battery which cannot be considered item equivalent for the two groups, the non-representational complexity test was shown to have a cultural bias, probably as a result of the abstract nature of the stimulus material. Item Analysis of the Curiosity Questionnaire revealed that while both groups obtained similar scores on the test, the African and European subjects scored high on different items in the test. The scores on the individual items in the test for both groups were correlated with scores obtained by the African and European samples in the preliminary version of the questionnaire. The intercorrelations obtained were significantly higher within than between cultural groups, showing that "culture of the sample" has an important effect on the mode of answering the questionnaire. The Conflict Score obtained can also not be considered Item Equivalent as it was shown to be highly unreliable in the African group. However, significant differences in reaction time between the two groups on both days of testing, under all conditions, were found.

No interpretation of actual differences in scores on the three tests which cannot be regarded as comparable in the two cultural samples is given. It is interesting that no significant differences emerged between the two groups in the five

tests which can be regarded as more or less item equivalent. i.e. for the stimulus material selected, no differences in preference for representational complexity, novelty, uncertainty or incongruity was found and subjects in the two cultural groups also rated themselves equally on curiosity.

Lack of significant differences between the two cultural groups on the tests which can be regarded as comparable for the samples, can, on first consideration, be regarded as rather a striking finding, since equal results for Africans and Europeans seem to be the exception, rather than the rule, in most cross-cultural studies. On further consideration, however, the results are less surprising. Both groups of subjects were undergraduate psychology students at South African Universities, with similar educational qualifications and very probably similar interests since they were doing the same course at University. In less homogeneous groups, more representative of the general population, greater differences in performance on the tests could be expected and the factors affecting the development of curiosity, discussed in Chapter II could conceivably have a greater effect on the African sample. However, since this study was of an exploratory nature, it was decided to use a technologically sophisticated and educated African sample. If no or only small differences are found between two samples which are highly similar in terms of motivation and education, then it is unlikely that the greater differences which may be foun-

between groups at lower levels of literacy reflects more than readily changeable cultural effects.

As was mentioned earlier, the intelligence tests were included in the battery primarily as a means of testing the discriminant validity and functional equivalence of the curiosity tests in the two cultural samples (discussed in the next chapter). Consequently they were not subjected to as detailed an analysis as the curiosity measures. In contrast to the results on the curiosity tests, large differences in performance between the two groups were found on all the intelligence tests. The Locations test was found to be unsuitable for its purpose in the African sample, even with the extended time limit. In addition, correlations with other intelligence tests were low for both cultural samples. It was consequently excluded from further analysis of the relationship between the intelligence and curiosity measures, discussed in the following chapter.

The study provided confirmatory evidence for several of Berlyne's hypotheses relating to the determinants of curiosity discussed in Chapter I.

Both the African and European subjects preferred more complex patterns to the simpler ones up to a certain level of complexity, which can be called the subject's optimal or preferred level of complexity, and furthermore, when given the choice, subjects preferred this optimal level of complexity to stimuli which belonged to a higher or lower complexity level.

While the uncertainty data did not provide clear-cut

evidence that uncertainty induces perceptual curiosity in that subjects in both groups did not look significantly longer at blurred than at clear pictures, this was probably due to unsuitability of the stimulus material. A supplementary study relating length of looking time at the pictures in the main study to correct identification of the pictures by a number of independent judges showed a significant negative relationship between the number of judges correctly identifying the pictures in the supplementary study and the length of looking time at the pictures by the subjects in the main experiment.

The novelty test, too provided confirmation for the hypothesis that subjects will look longer at novel than at familiar stimuli; this fact applying to both relative and absolute novelty.

Confirmation for the hypothesis that subjects will look longer at incongruous stimulus material was found only in the African group, however, order of presentation of the stimuli or the simplicity of the stimulus material itself may have been at fault here.

Although evidence was found that conflict (as measured by a lengthening of reaction time) does increase with the nearness in equality of the competing response tendencies in both samples, the effect was more marked in the European sample. However, the measure of conflict chosen can not be considered satisfactory for the testing of Berlyne's hypothesis that conflict is the variable underlying the motivational effects of the collative variables

(cf. Chapter I, P. 13) since a negative reliability for the conflict score was obtained in the African sample, and there were indications in the European sample, that ability to discriminate between stimuli on the third level of conflict in the experiment, rather than conflict per se was being measured.

CHAPTER V

ANALYSIS OF THE NATURE AND COMPARABILITY OF THE CURIOSITY CONSTRUCT: DISCUSSION AND CONCLUSIONS

1. Interrelationships between the Curiosity Tests

In Chapter I, it was pointed out that although curiosity is usually considered a unitary construct, there is no direct evidance that this is actually the case. Furthermore tests measuring novelty, complexity, incongruity, uncertainty and conflict, regarded as the prime determinants of curiosity by Berlyne (1960) have been used individually as measures of curiosity in a wide variety of studies without any attempt, beyond comparison with teachers' or self ratings' of curiosity to verify that curiosity is actually being measured in these tests.

In this study, the principal aim was to investigate the interrelationships between tests tapping the most important collative variables, complexity, novelty, incongruity and uncertainty. If curiosity, as defined by Berlyne (1960, 1963,1966) is in fact a unitary construct, then one would expect that intercorrelations between scores on tests of the various collative properties would be high. It would also be expected in terms of Campbell & Fiske's (1959) requirements for discriminant validity that intercorrelations between the curiosity measures and other tests such as the intelligence

tests would be low: i.e. considerably lower than between the curiosity tests themselves.

The intercorrelation matrices for the curiosity and intelligences tests in the African and European samples are shown in Tables 24 and 25. Conflict scores have been delekted from the matrix in view of their lack of reliability and validity in the two samples. In the matrix of the African sample, an additional variable, scores on the biographical questionnaire is included. Correlations of all curiosity measures, with the exception of the curiosity questionnaire and the self-rating scale which were administered on one occasion only, are based on combined scores on the two testing sessions. In the original data, low scores on the self-rating scale were regarded as indicative of high curiosity; in order to avoid confusion, as negative correlations with other curiosity tests would indicate positive relationships, the signs of the correlations between the self-rating scale and other tests were reversed.

Inspection of the correlation matrices for the two groups reveals that although correlations between the intelligence measures are low (mean r for the African sample = 0,13, mean r for the European sample = 0,12) they are no lower than between the curiosity measures themselves (mean r for Africans = 0,17, mean r for the Europeans = 0,15) only a few correlations reaching significance. The correlation between the two complexity tests was significant in both samples, indicating that these two tests do to some extant tap the same construct in the two groups. A significant correlation, in both samples,

Table 24

Intercorrelations of scores in the African Sample.

		1	2	3	4	5	6	7	8	9	10	11
Non-Rep. Complexity	1	1,00										
Rep. Complexity	2	0,44**	1,00									
Uncertainty	3	0,04	0,40**	1,00								
Incongruity	4	-0,21	0,00	0,42*	1,00							
Novelty	5	0,18	-0,04	0,06	0,05	1,00						
Curiosity scale	6	0,04	0,23	-0,02	-0,14	-0,17	1,00					
Self-rating scale	7	0,15	0,14	-0,05	-0,18	0,11	0,41**	1,00				
Mental Alertness	8	0,26	0,04	0,32*	0,06	0,14	0,31*	0,12	1,00			
Blox	9	0,00	0,08	0,05	0,22	-0,11	0,01	0,12	0,37**	1,00		
Letterset	10	0,19	0,18	0,08	0,12	0,12	0,19	0,01	0,38*	0,46**	1,00	
Biographical Scale	11	0,17	0,03	0,06	-0,02	0,27	-0,09	-0,19	0,06	0,19	0,16	1,00

Table 25

Intercorrelation of scores in the European Sample.

		1	2	3	4	5	6	7	8	9	10
Non-rep.Complexity	1	1,00									
Rep. Complexity	2	0,45**	1,00								
Uncertainty	3	0,24	0,09	1,00							
Incongruity	4	0,11	0,00	0,17	1,00						
Novelty	5	-0,07	0,03	0,06	0,09	1,00					
Curiosity	6	0,15	0,10	0,01	-0,17	0,03	1,00				
Self Rating Scale	7	0,07	0,20	-0,07	-0,27	-0,22	0,38**	1,00			
Mental Alertness	8	0,15	0,06	0,09	0,21	0,09	-0,17	-0,26	1,00		
Blox	9	0,11	0,22	-0,22	-0,05	0,03	-0,12	-0,19	0,44**	1,00	
Letterset	10	0,08	0,24	-0,24	-0,07	0,19	-0,06	-0,13	0,43**	0,42**	1,00

was also found between scores on the curiosity questionnaire and the subjects' self-ratings on curiosity, i.e. Subjects who saw themselves as curious, interested, active etc., also obtained high scores on the curiosity questionnaire, thus adding to the face validity of the questionnaire. However, scores on the curiosity questionnaire, in the African sample also correlated significantly with scores on the mental alertness test, detracting from the discriminant validity of the questionnaire. It is noteworthy that none of the other curiosity measures in this study correlated significantly with either the curiosity questionnaire or the self-rating scale on curiosity, thus casting serious doubt on the suitability of the collative variables, as used in this study, as measures of curiosity.

While no further significant correlations between tests of the collative variables and other tests were observed in the European group, the correlations between the uncertainty test, and incongruity test and between the uncertainty test and the representational complexity test were significant in the African sample. This would at first lead one to suppose that the relationship between these three collative variables is closer in the African than in the European group; however, a correlation of 0,00 was found between the incongruity and representational complexity test scores, and furthermore, the scores on the uncertainty test also correlated significantly with scores on the mental alertness test in the African sample. It is therefore considered more likely that the uncertainty test lacks

discriminant validity in the African sample. Campbell & Fiske (1959) state that tests can be invalidated by high correlations with other tests which purport to measure different traits.

On the whole the correlations between the collative variables, the curiosity questionnaire and self-rating scale and between the curiosity measures and intelligence tests do not provide support for the hypothesis that curiosity is a unitary construct when seen in terms of Berlyne's (1966) definition of curiosity as: "a state of heightened arousal brought about by incomplete perception of a sector of the stimulus field which leaves the subject in some uncertainty as regards its characteristics." which regard the collative stimulus properties as prime determinants of curiosity.

However, in terms of the definition given in this study, as a motive to acquire or transform incoming stimulation under circumstances that offer no immediate value in terms of classic biogenic needs or acquired drives, an alternative interpretation of the results is possible. In terms of this definition, curiosity can be seen as a state of the organism rather than as being caused by the collative properties of stimuli, these determining not whether curiosity will be aroused but which stimuli will be explored, i.e.the way in which curiosity manifests itself. The low intercorrelations in this study can then be explained by saying that curiosity manifests itself differently in different subjects without any implications for the unitary-or-other nature of

the curiosity construct. The conclusions reached in this study, that curiosity is a multifaceted rather than unitary construct applies only if Berlyne's definition of curiosity is accepted. The results of this study do, however, cast serious doubt on the validity of earlier studies which used novelty, complexity, incongruity and uncertainty as measures of curiosity without verification that curiosity rather than merely preference for novelty, complexity, etc. is being measured.

2. Comparability of the results in the African and European samples

The requirements for comparability of test scores accross cultures were discussed in Chapter II. Not only must the individual tests be item equivalent in the two cultural groups being compared but the tests must be shown to be functionally equivalent and score equivalent as well. However, as discussed in Chapter II, a number of difficulties are encountered in the determination of functional equivalence as advocated by Poortinga (1971); viz. by means of the comparison of correlation matrices. First of all, meaningful comparison of correlation matrices requires larger samples than used in this study, and secondly, meaningful determination of functional equivalence implicitly assumes that the various tests being considered should measure the same complex of factors, a requirement also not fulfilled in this study. Consequently no definite conclusions as to the functional equivalence of the battery as a whole can be drawn.

However, some significant correlations values which did emerge between the curiosity tests and between these tests and the intelligence tests and between the intelligence tests themselves can be discussed briefly, although the results must be treated with reservations in view of the size of the samples.

Correlations between the two complexity tests and between these tests and the intelligence tests were sufficiently similar to warrant the conclusion that the complexity tests are functionally equivalent in the two samples. Nevertheless, the two tests cannot be considered comparable in the two samples since the representational complexity test was not item equivalent (cf.Chapter IV) and furthermore the two tests cannot be considered score equivalent in that the Africans scored considerably lower than the Europeans in the non-representational complexity test, while no significant difference between the two samples was found on the representational complexity test so that the requirement for score equivalence, namely, that the predicted score on one test is the same for all subjects who had an equal score on the other test, irrespective of whether they belong to the same group or not, is not met.

Similar correlations in the two groups were also found between the curiosity questionnaire and the self-rating scale. The curiosity questionnaire cannot however be considered functionally equivalent in the two samples since the correlation with the mental alertness is significantly higher in the African sample ($p < 0,01$, cf. Edwards,1953,

p.131). In view of the higher correlations between the uncertainty test and the non-representational complexity tests and between the uncertainty test and the incongruity test, the uncertainty test can also not be considered functionally equivalent in the two samples.

The intelligence tests can however be considered to be functionally equivalent since a Box's test for equality of covariance matrices (cf.Morrison, 1967, p.151) in the intelligence tests showed the matrices to be essentially similar in the two cultural groups. This finding runs counter to that of Poortinga (1971) who found intercorrelations between the Blox and Mental Alertness tests which were substantially lower in the European group.

It must be emphasized again, however, that conclusive interpretations cannot be made on the basis of these results in view of the lack of homogeneity of the tests and the size of the sample.

No significant relationships between the Biographical Questionnaire and the curiosity tests was found in the African sample. This may be the result of the small standard deviation of scores on the questionnaire and the resulting low reliabilities. Nevertheless the questionnaire provides little evidence that the results of the tests are affected by the degree of urbanization of the subjects. The use of Grant's original urban-rural scale, being longer and more comprehensive, might have yielded more significant results. Furthermore, the subjects in the African sample were a highly selected group; in a group more representative of

the general population, the factors affecting curiosity in Africans, discussed in Chapter II, might have resulted in a higher correlation between scores on the tests of the collative properties and the biographical questionnaire being observed.

3. Summary and Conclusions

Two main aims in carrying out this study were specified in Chapter III. No evidence was found for the hypothesis that the collative stimulus properties, regarded as the prime determinants of curiosity by Berlyne (1960, 1963, 1966), do in fact tap the same construct, in that correlations between tests of the collative variables were low in both cultural groups. Furthermore correlations between the tests of the collative stimulus properties and the curiosity questionnaire and self-rating scale were low, further detracting from the face validity of the tests of the collative stimulus properties as tests of curiosity. The results of this study thus cast serious doubts on the validity of earlier studies (cf. Chapter I) which have used tests of one or other of the collative variables individually as measures of curiosity and lead to the conclusion that curiosity, as defined by Berlyne, should be regarded as a multifaceted rather than unitary construct. If, however, curiosity is defined as a motive to acquire or transform incoming stimulation under circumstances that offer no immediate reinforcement in terms of classic biogenic needs or acquired drives, an alternative interpretation of the

results, in which the collative stimulus properties are seen not as determinants of curiosity but as manifestations of curiosity can be given.

However, the possibility cannot be excluded that failure to find confirmatory evidence for Berlyne's implicit hypothesis that curiosity is a unitary construct, is partly due to the artificiality of the experimental situation and stimulus material. According to Berlyne (1960, cf.Chapter I, p.3) research on curiosity should concentrate on the problem of "stimulus selection" and try to provide an answer to the question: "To what stimulus will this animal or person respond?" rather than on the traditional behaviouristic problem of "response selection" which asks the question: "What response will this animal/person make to this stimulus?". It can be argued that in many studies on curiosity including some by Berlyne himself (Berlyne, 1962, 1963a, 1966, 1970; Berlyne et al., 1963; Berlyne & McDonnel, 1965) and in this study as well, in all tests of the collative variables, with the exception of the complexity tests, we are still attempting to find the answer to the second rather than the first question. A less structured situation, characterized by a greater freedom of response to a number of different stimuli might yield more fruitful and valid results.

Furthermore, although it seems likely that in the two complexity tests, the response measure, stated preference, is an indication that a greater interest was evoked by the preferred as opposed to the non-preferred item, there is

no evidence that the measure of "curiosity" used in the other tests, viz. length of looking time, gives a true indication of the intensity of curiosity aroused by the stimuli.

A further criticism which can be made of the findings in this study is that the subjects used were a highly selected group of students who can perhaps be considered more "curious" than the general population. It can be argued that if the experiments were repeated with subjects more representative of the general population, the larger individual differences which would appear might result in stronger relationships between the various collative stimulus properties being observed. However, for reasons discussed later in this section it was decided to use a homogenious group of subjects.

Although the study failed to provide evidence for Berlyne's implicit hypothesis that curiosity is a unitary construct, confirmatory evidence for several other hypotheses about the collative variables were found. These were discussed in some detail in Chapter IV, but for the sake of completness of the discussion, they will be repeated briefly here. Subjects in both cultural groups were found to have an optimal or preferred complexity level and when given the choice, preferred stimuli at this level of complexity to stimuli which belonged to a higher or lower complexity level. The novelty test provided evidence that subjects in both cultural groups would look longer at novel than at familiar stimuli, this fact applying to both absolute and relative novelty. Con-

firmation for the hypothesis that subjects will look longer at incongruous than at congruous stimulus material was found only in the African group, while the hypothesis that uncertainty induces perceptual curiosity and a consequently longer looking time was confirmed in the European group only. It is considered, however, that excessive simplicity and faulty presentation of the stimulus material may have been responsible for the lack of significant findings in these two tests. The conflict measure derived in the conflict reaction time experiment was considered unsatisfactory for testing the hypothesis that conflict is responsible for the motivational effects of the collative variables, due to the low reliability and validity of the measure in the two cultural groups. However, the failure to find any significant correlation between scores on tests of the various collative stimulus properties casts doubt on the existence of any underlying factor, be it curiosity or conflict.

The second main aim of this study, to investigate curiosity in a group of African and European students and to discover whether the determinants of curiosity operate in a similar manner in an African and European group, was discussed in relation to Poortinga's (1971) criteria for cross-cultural comparability. There were indications that 5 of the individual tests and questionnaires (representational complexity, uncertainty, novelty, incongruity, and the self rating scale) can be considered item equivalent in that the order of item difficulty in the two cultural groups being

Author Foden Bernardina
Name of thesis Curiosity in Africans and Europeans. 1973

PUBLISHER:
University of the Witwatersrand, Johannesburg
©2013

LEGAL NOTICES:

Copyright Notice: All materials on the University of the Witwatersrand, Johannesburg Library website are protected by South African copyright law and may not be distributed, transmitted, displayed, or otherwise published in any format, without the prior written permission of the copyright owner.

Disclaimer and Terms of Use: Provided that you maintain all copyright and other notices contained therein, you may download material (one machine readable copy and one print copy per page) for your personal and/or educational non-commercial use only.

The University of the Witwatersrand, Johannesburg, is not responsible for any errors or omissions and excludes any and all liability for any errors in or omissions from the information on the Library website.