DOCUMENT RESUME

ED 237 573 TM 830 838

AUTHOR Hample, Dale; Dallinger, Judith M.
TITLE Trust, Solidarity and Target Valence: Some Problems with Our Validity Paradigm.
PUB DATE May 83
NOTE 25p.; Paper presented at the Annual Meeting of the International Communication Association. (Dallas, TX, May 26-30, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Evaluation Methods; Higher Education; *Interpersonal Communication; *Measurement Techniques; *Rating Scales; Research Utilization; Self Disclosure (Individuals); *Test Reliability; *Test Validity
IDENTIFIERS Individualized Trust Scale; *Target Valence; Trust; *Wheeless (L R)

ABSTRACT

Wheeless' trust and solidarity scales are examined under conditions of positive and negative target person. Valence of stimulus is shown to have a strong influence on the performance of the scales. The problems on Wheeless' scales were not detected in the original validation studies because of the normal validation paradigm. That paradigm emphasizes generalizability and scale purity, and therefore tends to risk validity by pooling data across a wide variety of stimulus conditions. The weaknesses of these standard procedures are explained, and suggestions for reform are made. (Author/PN)

TRUST, SOLIDARITY AND TARGET VALENCE: SOME PROBLEMS WITH OUR VALIDITY PARADIGM

By

Dale Hample and Judith M. Dallinger

Western Illinois University

Abstract

Wheeless' trust and solidarity scales are examined under conditions of positive and negative target person. Valence of stimulus is shown to have strong influence on the performance of the scales. The problems in Wheeless' scales were not detected in the original validation studies because of our normal validation paradigm. The weaknesses of these standard procedures are explained, and modest suggestions for reform are made.

Paper presented to the International Communication Association, Dallas, Texas, May, 26-30 1983.

TRUST, SOLIDARITY AND TARGET VALENCE:

SOME PROBLEMS WITH OUR VALIDITY PARADIGM

Wheeless has recently published a series of studies which present evidence for the reliability and validity of scales designed to measure trust and solidarity (1976; 1978; Wheeless & Grotz, 1977). The trust measure, the Individualized Trust Scale, consists of fifteen semantic differential items which measure the trustworthiness of a specific target person.[1] The solidarity instrument uses Likert items--originally nine of them, but twenty in the later version of the scale (Wheeless, 1978). Solidarity concerns the closeness, the identification, and the affiliation one feels for another person. Solidarity and trust are associated with one another empirically, and also correlate significantly with self-disclosure (Wheeless, 1976; 1978; Wheeless & Grotz, 1977). That trust, solidarity and self-disclosure belong to a key cluster of interpersonal variables hardly needs documentation here. The purpose of the present investigation is to re-examine the relationship between trust and solidarity, and to reconsider the validity of the two measures.

The impetus for this study comes from two recent papers which use these instruments to study lying (Hample, 1980; 1982). In both experiments, trust and solidarity are used as dependent measures which are expected to discriminate either between liars and their victims or between different types of liars. That the scales do so is evidence for their construct validity. Some other results in Hample's papers, however, are disquieting.

First, in Hample's factor analyses, solidarity loadings are markedly lower than in the Wheeless reports.[2] Table 1 shows the seriousness of the problem.

INSERT TABLE 1 HERE

In Wheeless' paper, the scale performs admirably, but Hample's results indicate

a marked lack of purity in the scale. Quite a few of the items would not normally be included in the scale, if only Hample's results were used as a criterion. The relatively small sample used in Hample (1980) might allow us to neglect those loadings on grounds of possible matrix instability. But the larger sample in Hample (1982) meets reasonable standards (e.g., McCroskey & Young, 1979), and so must be taken more seriously.

Though the problems with the trust scale are not quite so dramatic, Table 2 shows a similar degradation of loadings. The Wheeless and Grotz (1977)

INSERT TABLE 2 HERE

results are more modest to begin with, and hold up a bit better than those for the solidarity measure. However, most items have lower loadings in Hample's larger study, and three are below .60.

Further evidence that the scales were less pure in Hample's studies is also reported in Tables 1 and 2. The percentage of total item variance accounted for by the first (only) factor of the solidarity scale drops in half in Hample's experiments. For trust, the decline is gentler, but still important. This is a critical result because of the normal practice of summing items (rather than using factor scores) to get instrument scores. The original Wheeless results are fairly encouraging about what the item sums measure, but Hample's data challenge the validity of such item totals.

The final difficulty which Hample's papers uncovered has to do with the correlation between trust and solidarity. This coefficient is .65 in Wheeless (1978). In Hample (1980), the figure for liars is .51, and for their victims is .53. In Hample (1982), the correlation is .40. The relationship between trust and solidarity is a theoretically important one. If the two measures do not in fact correlate well, the construct validity of one or both scales is jeopardized. Whether the common variance of these two constructs is closer to 35% or 15% is therefore a significant question.

Several things could account for the differences between Wheeless' and Hample's results. Samples were from different populations, though undergraduates were used in all but Wheeless and Grotz (1977), wherein teachers and their families provided data.[3] Procedures for data collection seem to have been reasonably similar among the studies, except that the Wheeless (1976) and Wheeless and Grotz (1977) subjects filled out scales at home rather than in a classroom setting of some sort. In principle, either the sampling or procedural differences could explain the divergent results, but it seems implausible that the scales should be so sensitive to such minor dissimilarities.

More theoretically interesting departures have to do with the scales' targets. In all of Wheeless' studies, subjects received categories of targets (e.g., classmate/coworker), imagined a specific person who fit the category, and then filled out the scales for that real person. In both of Hample's experiments, subjects' responses referred to fictional targets who had been involved in a conversation which the students had just read. So targets in Wheeless' studies were real and familiar, while those in Hample's projects were not; perhaps the reality of, or experience with, the target influences the amount of error in the data. On this supposition, targets in the present study conform to Wheeless' methodology; if the scales perform poorly here, it will not be because the targets are ephemeral.

The second difference between the types of targets is the one on which this study focuses. All the targets in Hample (1982) and half those in Hample (1980) were liars. Liars are obviously negatively-valenced people, as the means for trust and solidarity in Hample's studies confirm. However, virtually all the targets in Wheeless' studies are positive. In Wheeless (1976), eighteen targets were used; of these, only one ("disliked person") seems to have a negative valence. Wheeless and Grotz (1977) used twenty targets; only disliked male and disliked female are negative. Wheeless (1978) used the Wheeless (1976) list again. Wheeless'

results, of course, are based on combined [illegible] targets. Though his lists are reasonably varied--they are int[illegible] to [illegible]se close and distant relations with males and females--they a[illegible] almost u[illegible]formly positive.

The present study tests the hypo[illegible] that [illegible]nce of target mediates the validity of the trust and solidarit[illegible] ca[illegible] [illegible]articular, negatively valenced targets are expected to stimul[illegible] [illegible] f[illegible]tor loadings and a smaller trust-solidarity correlation than in Whee[illegible] [illegible]iginal reports.

Method

Subjects were 185 undergraduates enrolled in a university-required public speaking class. 56% were male and 44% female. 70% were juniors or seniors. Students received extra credit for participating. Materials were filled out after the regular class meetings, in an ordinary classroom.

Each student was given a booklet and three index cards. The booklet consisted of the solidarity and then the trust scales. The solidarity items were presented in the same order as in Wheeless (1978), and the same instructions were typed at the top of the instrument. The trust items were presented in the same order as in Wheeless and Grotz (1977), with instructions patterned after those for solidarity. Several of the trust items were reversed in polarity to avoid response sets. Following Hample (1980; 1982) the second item was reworded slightly to increase clarity: "I am" was added to trustful/distrustful of this person.

Each index card had one of 22 target persons on it. Half these were positive and half negative. The positive targets were (stimuli also used by Wheeless are asterisked): most liked female,* most liked male,* mother,* father,* barber/hairdresser,* minister/priest/rabbi,* favorite high school teacher, roommate,* employer,* coworker/classmate,* and doctor.*[4] The negative targets were: most disliked male,* most disliked female,* conceited male, conceited

female, rude salesperson, male convicted of rape, female prostitute, unfair teacher, dishonest policeman, someone who lies a lot, and someone who doesn't repay debts. These targets were randomly distributed, with the stipulation that each student received at least one of each valence. In line with Wheeless' procedures, students were allowed to trade in targets which were unfamiliar. 14 negative stimuli were turned back (the most frequent were the rapist and the prostitute, at 4 times each), as were 4 positive targets. Since each student provided data three times, the effective sample size is larger than 185. 279 sets of data were obtained for positive targets, and 274 for negatively valenced ones.

Students provided data on computer-readable forms. All scales were seven items, as is also the case in the Wheeless experiments. Students finished in about 25 minutes, and were debriefed.

Results

Induction

The validity of the positive/negative operationalization was tested by comparing the data from each condition on several variables. For solidarity, the groups differed significantly ($\bar{X}_p$=77.1, $\bar{X}_n$=40.3, t=15.62, $p<.001$), as they did for trust ($\bar{X}_p$=72.8, $\bar{X}_n$=32.8, t=25.34, $p<.001$). Subjects also rated each target as being negative person/positive person and close to me/far from me. These variables also produced significant constrasts ($\bar{X}_p$=4.5, $\bar{X}_n$=2.4, t=12.51, $p<.001$, and $\bar{X}_p$=4.2, $\bar{X}_n$=1.9, t=14.03, $p<.001$, respectively). All differences were in the expected directions. The valence induction was clearly successful.

Reliability

In Wheeless and Grotz (1977) the split-half reliability for trust is .92; Wheeless (1978) reports .97 for his 14 item version of the scale. Hample (1980) gives Cronbach's alphas of .90 and .92 for the victim's and liar's data;

Hample (1982) reports an alpha of .92. In the present study, alpha is .97 for all data, .93 for negative targets and .92 for positive ones.

For solidarity, Wheeless (1978) reports a reliability of .96. Hample (1980) gives .87 and .86 as the figures for victims and liars, while Hample (1982) reports a value of .89. The present data yield an alpha of .96 for all data, .93 for negative targets, and .95 for positive targets.

In sum, both scales continue to be extremely reliable, and valence of target is irrelevant.

Trust-Solidarity Correlation

The original report of the trust-solidarity correlation was .65 (Wheeless, 1978). Hample (1980) gives .51 and .53 for liar and victim, while Hample (1982) reports .40. Here, the figure for all data is .67 ($p<.001$), .43 ($p<.001$) for negative targets, and .44 ($p<.001$) for positives.

Valence of target obviously makes a considerable difference in these correlations. Figures 1, 2 and 3 dramatically show what the true relationship

INSERT FIGURES 1, 2 AND 3 HERE

is. Figure 1 is the scatterplot for all data, and displays the expectable pattern which ought to appear for a fairly large postive correlation. Figure 2 shows the same plot for negative targets, however, and Figure 3 gives the positive results. As is plain, the negative and positive patterns are very similar, but are in different portions of the variable space. Both graphs show moderate relationships. Only when Figures 2 and 3 are superimposed does the "clear" relationship of Figure 1 appear. This means that Figure 1 represents a spuriously large correlation coefficient, and that the only valid statistics for this study are .43 and .44. The three plots together provide unequivocal evidence that valence of target mediates the relationship between trust and

solidarity. Since Wheeless (1978) used only one negative target, his result is not as thoroughly confounded as the all data condition here. Even so, the present correlations ought to be preferred to Wheeless' as a more accurate indication of the trust-solidarity relationship, especially if one wishes to generalize to only positive (or negative) targets.

Factor Analyses

Table 3 presents the trust loadings for the present study. Those from Wheeless and Grotz (1977) are also reproduced for convenience. For the all data condition, loadings are actually a little better than in the original study.

INSERT TABLE 3 HERE

For either negative or positive targets, however, loadings worsen. The percentage of variance explained is correspondingly high for all data, and much lower for each valence condition. The latter figures are, however, quite close to those reported in Hample's studies (see Table 2). Sample sizes, even in the valence conditions, are large enough to have generated stable correlation matrices, and should therefore have had little effect on these results.

The parallel solidarity loadings are in Table 4. The all data condition does not do quite as well as the original report, but loadings are still comparable.

INSERT TABLE 4 HERE

Once again, both valence conditions produce worse loadings. Variance accounted for is lowest for these latter columns as well, although neither result is as poor as Hample reports (see Table 1). As before, sample sizes in the present study are sufficient.

Table 5 presents the results of statistical comparisons among the studies reported in Tables 3 and 4. The r_c coefficients reported in the body of the

table are measures of association between the loadings resulting from the various data sets. This congruence coefficient is similar (but not identical) to

INSERT TABLE 5 HERE

Pearson's r. It is a convenient index of factor similarity, and significance levels for it have been published (see Cattell, 1978, pp. 251-255 for the calculation formula and discussion, and pp. 568-569 for the significance table). As the results clearly indicate, the trust and solidarity scales have the same essential patterns in each data set.

Discussion

The present results confirm the reliability and basic structure for both the trust and the solidarity instruments. Target valence, however, certainly mediates the relationship between these two measures, and may well be mediating inter-item relationships within each scale as well.

The mechanism of valence mediation is this. Negative and positive targets stimulate radically different ratings on the items, as the induction results show. The effect of this is that negative targets produce low ratings on the items and positive targets receive higher scores. Negative and positive stimuli will therefore each have clusters of points (for any two given scale items) in different corners of the variable space. When all the data are included in a correlation, the two clusters together produce a pattern corresponding to a strong correlation, even though the relationship may be mediocre or completely absent in one or both clusters. The high correlations and concomitantly high loadings which appear when data for positive and negative targets are jumbled together are therefore spurious and should be discarded. The comforting results for the all data condition in the present study are hopelessly confounded, and so are meaningless for ordinary purposes. Wheeless' use of one or two negative

targets in each of his original studies contaminates them as well. Whether those few targets are sufficient to explain the differences between his results and those for the present positive valence condition is an empirical question which can be answered conclusively only by reanalysis of Wheeless' data. In any case, the best currently available information on the characteristics of the trust and solidarity scales is that which is reported here for the two valence conditions.

This outcome suggests a more general and serious problem with the use of standardized scales. Any variable whose conditions are radically different on two instruments may well mediate those scales' association. Of course, one might be fortunate enough that the two variables' association could be the same for each level of the potential mediator. The present study shows, however, that such mediation can be quite important. And the present results are not an especially drastic example. If the two clouds of points in Figures 2 and 3 were a little rounder, the overall correlation between trust and solidarity would barely be affected. The correlation within each valence condition could be zero, though. The two component clusters could even have slightly negative correlations and still produce a fairly high positive overall correlation. This can happen with any scale which in used in conjunction with any variable whose conditions differ radically on the scale measure. This, of course, is a common design.

We are therefore confronting a fundamental problem for our validation paradigm: generalizability can be bought at the price of validity. By including positive and negative, male and female, close and distant, expert and inexpert targets in our validating studies, we seem to be assuring ourselves that the results will not be specialized for one kind of target. But the current investigation shows that if the menagerie of stimuli really does contain clearly distinct conditions, the reality is harsher. The effort to promote generalizability has sacrificed the quality of what is being generalized. A related motive--to

maximize the scale variance--leads to exactly the same design (a large variety of stimuli) and carries the same risk as well. Maximizing variance makes a multi-item scale come out more cleanly, and this report may well explain why.

All this is not an argument against using a variety of targets to validate measures. But it is a strong argument against combining the data from those conditions. We need to have separate results for any stimulus conditions which make a difference--any which do a particularly good job of increasing scale variance. We can only agglomerate data when we have established that the conditions are unimportant. Once having shown that a particular contrast (e.g., male/female) makes no difference, data can be pooled, and the results may be confidently generalized over that contrast without risking invalidity or spuriousness. This ought to be our standard procedure, and ought to replace our present usual practice, which is represented here by Wheeless' work, but which is certainly not unique to him.

The present study can obviously serve as an example of how to test whether a third variable mediates two others. Perhaps not so evident is the application of these techniques to a single scale validation study. If a scale has several items, their intercorrelations should be compared between various conditions of the stimulus. If two items correlate well in the whole data set, but poorly (or even quite differently) between stimulus conditions, the researcher should be alerted to the possibility of mediation.

This procedure is not especially difficult to implement (provided it is planned for), and will produce better, more valid scales. Later researchers need to know, for example, that the correlation between trust and solidarity for a single target is really .44, not .67, and that trust and solidarity account for about 50% of the scales' sums' variance, not 60%. Of course, later workers ought to do their own re-validation of any standardized scales, precisely to locate problems similar to the one studied here.

Conclusions

Wheeless' scales are reliable and have the same structure, regardless of target valence. However, target valences does mediate the trust-solidarity association, as well as the scales' inter-item associations. These results are entirely expectable, given our usual scale validation procedures. That paradigm emphasizes generalizability and scale purity, and therefore tends to risk validity by pooling data across a wide variety of stimulus conditions. Accordingly, these practices ought to be reformed.

Notes

[1]Wheeless (1978) actually dropped one item because it failed to load.

[2]This is the first report of any of Hample's factor loadings, though they were discussed in passing in both papers. Neither of those investigations was centrally concerned with validating the scales, and the factor analyses were only conducted as a matter of procedure.

[3]Wheeless (1976) used both regular undergraduates and adults enrolled in extension graduate courses in communication in the classroom. The loadings on the (nine item) solidarity scale were quite similar for the two samples. This suggests that the different sample in Wheeless and Grotz (1977) probably did not generate non-comparable data.

[4]A few of the asterisked targets only correspond approximately to Wheeless' targets--for example, most liked female to best female friend.

References

CATTELL, R.B. The scientific use of factor analysis in behavioral and life sciences. New York: Plenum Press, 1978.

HAMPLE, D. Empirical evidence for a typology of lies. Paper presented to the Speech Communication Association, Louisville, 1982.

HAMPLE, D. Some effects of an intentional communication rule violation: The case of superior-to-subordinate lies. Paper presented to the Speech Communication Association, New York City, 1980.

McCROSKEY, J.C. & YOUNG, T.J. The use and abuse of factor analysis in communication research. Human Communication Research, 1979, 5, 375-382.

WHEELESS, L.R. A follow-up study of the relationships among trust, disclosure, and interpersonal solidarity. Human Communication Research, 1978, 4, 143-157.

WHEELESS, L.R. Self-disclosure and interpersonal solidarity: Measurement, validation, and relationships. Human Communication Research, 1976, 3, 47-61.

WHEELESS, L.R. & GROTZ, J. The measurement of trust and its relationship to self-disclosure. Human Communication Research, 1977, 3, 250-257.

Table 1

Scale Loadings Frequencies on the Solidarity Measure

Loading	Wheeless (1978)	Hample (1982)	Hample (1980) liars	Hample (1980) non-liars
.90 or more	3			
.80 - .89	9			1
.70 - .79	5	2	2	1
.60 - .69	3	7	5	3
.50 - .59		4	2	4
.40 - .49		3	3	7
.30 - .39		4	3	1
.20 - .29			5	2
.10 - .19				1
% of Variance	63%	34%	30%	31%
N	385	202	82	82

Table 2

Scale Loadings Frequencies on the Trust Measure

Loading	Wheeless (1978)	Hample 1982)	Hample (1980) liars	Hample (1980) non-liars
.90 or more				
.80 - .89	8	1	1	
.70 - .79	3	5	6	4
.60 - .69	3	6	3	7
.50 - .59	1	2	4	2
.40 - .49		1	1	2
% of Variance	59%	47%	49%	44%
N	261	202	82	82

Table 3

Factor Loadings for Trust

Item	Wheeless and Grotz (1977)	All Data	Negative Target	Positive Target
1	.86	.90	.83	.81
2	.78	.90	.81	.82
3	.56	.74	.54	.67
4	.60	.81	.69	.71
5	.81	.84	.72	.72
6	.67	.81	.70	.69
7	.83	.86	.74	.76
8	.75	.70	.56	.51
9	.83	.74	.63	.52
10	.82	.83	.76	.57
11	.79	.88	.83	.71
12	.84	.76	.59	.54
13	.84	.75	.54	.58
14	.84	.85	.74	.66
15	.66	.72	.47	.61
% of Variance	59%	68%	50%	48%
N	261	518	261	257

Note: Item numbering follows Wheeless and Grotz (1977, p. 254). Item 3 is the one dropped by Wheeless (1978). Item 2 is the one slightly reworded here. The sample sizes reported are the smallest number of subjects providing data for any item; pairwise deletion of missing values was in force.

Table 4

Factor Loadings for Solidarity

Item	Wheeless (1978)	All Data	Negative Target	Positive Target
1	.91	.85	.77	.84
2	.74	.70	.51	.68
3	.74	.79	.67	.70
4	-.62	.52	.37	.49
5	.79	.77	.69	.71
6	-.72	.60	.41	.63
7	.75	.73	.68	.70
8	-.63	.70	.56	.56
9	.83	.72	.71	.56
10	-.81	.63	.57	.61
11	.81	.75	.64	.72
12	.80	.67	.52	.75
13	-.68	.67	.53	.54
14	.83	.76	.80	.66
15	-.91	.83	.76	.82
16	.84	.78	.74	.69
17	.84	.81	.74	.79
18	-.84	.68	.64	.57
19	.94	.84	.76	.83
20	.80	.68	.61	.69
% of Variance	63%	55%	44%	49%
N	385	545	271	269

Note: Item numbering follows Wheeless (1978, p. 147). The scores for all negative items were reflected prior to factoring in the present study, so the differences in signs are meaningless. The sample sizes reported are the smallest number of subjects providing data for any item; pairwise deletion of missing values was in force.

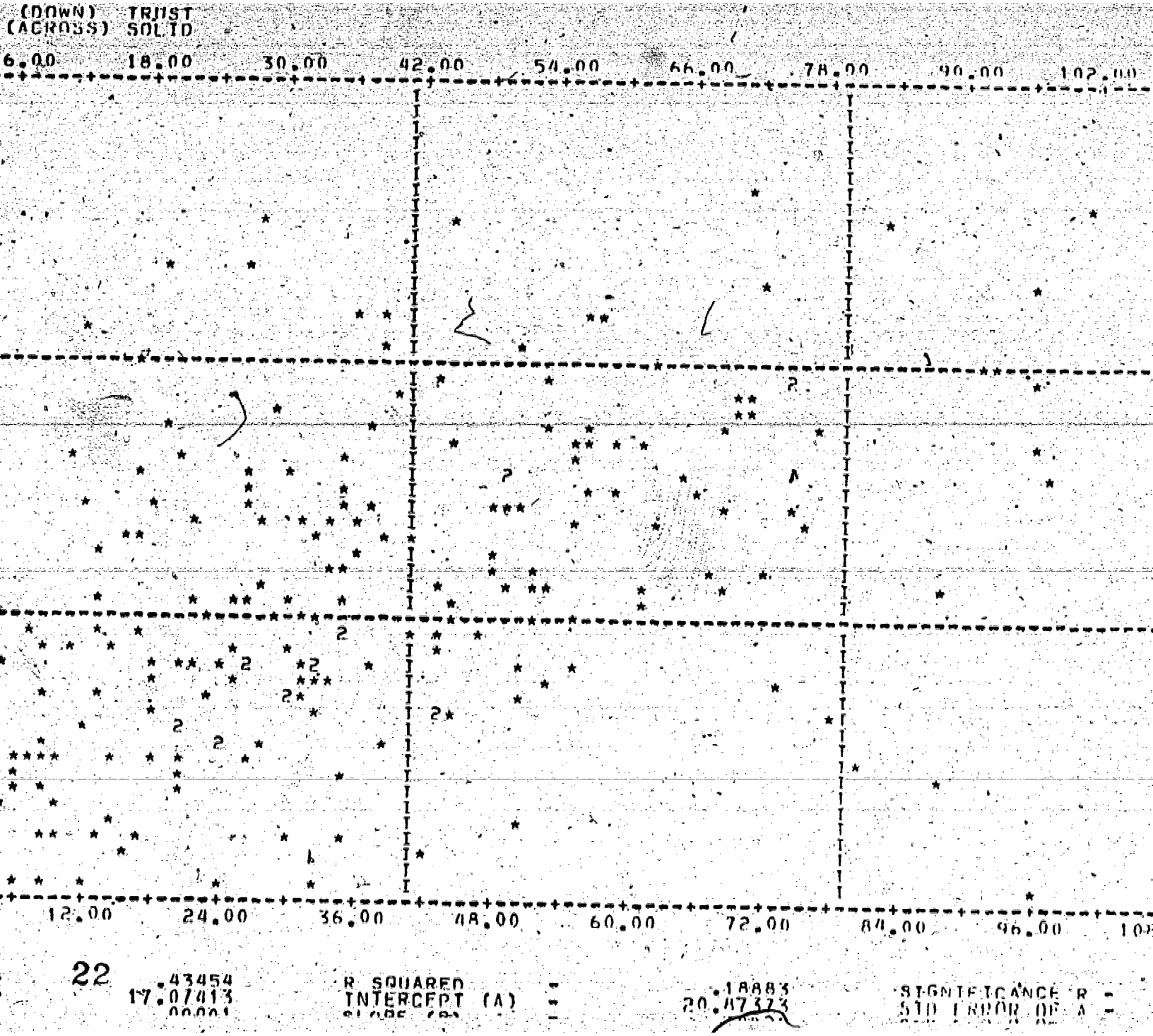
Table 5

Congruence Coefficients (r_c) for Trust and Solidarity

TRUST	Wheeless and Grotz (1977)	All Data	Negative
All Data	.993		
Negative Target	.989	.996	
Positive Target	.961	.996	.992

SOLIDARITY	Wheeless (1978)	All Data	Negative
All Data	.997		
Negative Target	.993	.995	
Positive Target	.995	.996	.989

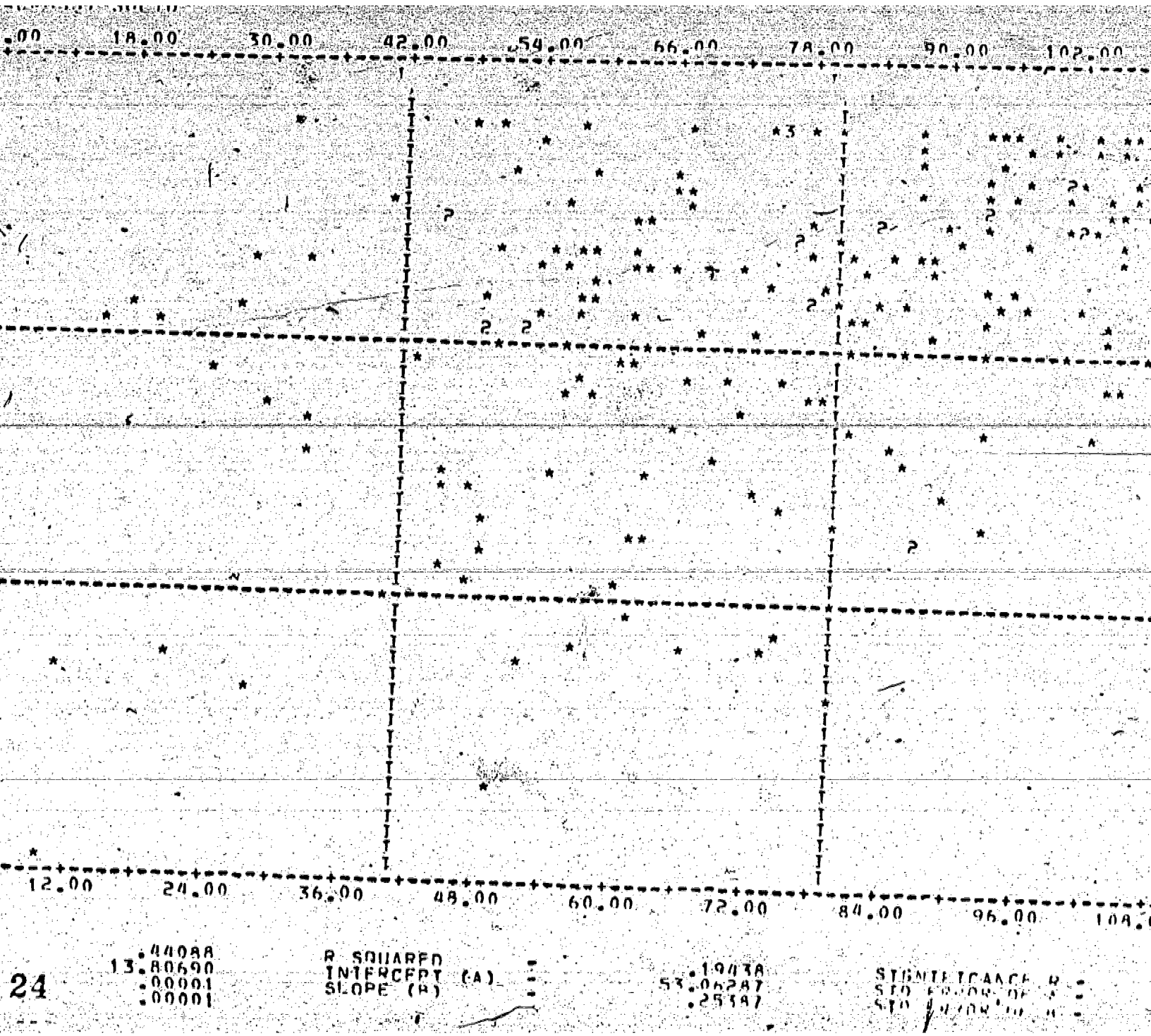
Note: Calculations were performed on loadings of five significant digits for the present study, and two significant places for Wheeless' data. The negative signs in Wheeless' results were ignored in the calculations because the present sutdy reflected negative items' scores prior to factoring, and Wheeless did not. All coefficients reported above are statistically significant at p .001.



22

.43454 R SQUARED = .18883 SIGNIFICANCE R =

17.07413 INTERCEPT (A) = 20.47373 STD ERROR OF A =



.44088 R SQUARED
13.80690 INTERCEPT (A) = 53.06287 STUDENT'S T
.00001 SLOPE (B) = .25197 STD ERROR OF B =
.00001