Validity and Utility of the Attributional Style Construct at a Moderate Level of Specificity

Craig A. Anderson
University of Missouri—Columbia

Dennis L. Jennings
Bell Communications Research
Livingston, NJ

Lynn H. Arnoult
Rice University

The validity and utility of attributional style has been questioned in recent years. Major criticisms are that attributional style is not cross-situationally consistent, is not measured appropriately, has little construct validity, and contributes little to the prediction of important social behaviors. This article examines these issues with data primarily related to the several different attributional style measures found in Anderson's Attributional Style Assessment Tests. We examined previously published and new data. The results show evidence of convergent and discriminant validity for attributional styles assessed at an intermediate level of specificity. Also, attributional style effects were as large as parallel attribution manipulations in complex social settings. We concluded that both the extreme pessimism of some recent researchers and the broad, sweepingly optimistic claims of some proponents are unwarranted, and that further work on specifying the appropriate level of assessment for attributional style would be useful.

The pervasiveness of depression and related clinical problems, along with recent theoretical advances in attribution theory, have prompted a myriad of researchers to investigate attributional models of these "problems in living" (e.g., Anderson & Arnoult, 1985a, 1985b; Anderson et al., 1983; Coyne & Gotlib, 1983; Feather & Tiggemann, 1984; Russell, 1982; Weiner, 1985; Wimer & Kelley, 1982; Wortman & Dintzer, 1978). We feel that much work is needed on all these questions, but that there are several more pressing issues warranting attention. These issues are (a) is AS a valid construct? (b) is a person's AS general across types of situations, specific to types of situations, or so idiosyncratically specific to particular situations as to render it invalid as an individual difference construct? (c) does AS have important effects in complex social settings, or do situational influences render it impotent?

These issues have been raised most often in the context of studies using Seligman's Attributional Style Questionnaire (ASQ; Seligman et al., 1979). This is due quite simply to the popularity of the ASQ and the reformulated learned helplessness model from which it derives (Abramson, Seligman, & Teasdale, 1978). For example, Cutrona, Russell, and Jones (1984) raised a number of issues and used two data sets based on the ASQ to address the issues. They found only weak evidence of a cross-situationally consistent AS, quite small correlations with depression, and little evidence that ASQ scores predict people's attributions for real negative events in their lives. Their basic conclusions were (a) the concept of AS may not be a valid trait conception, (b) it must be more narrowly defined, and (c) perhaps a more situational orientation to understanding attributional processes for specific events is needed.

Our goal in this article is not to defend the ASQ; its proponents are both responsible for and capable of doing that themselves (e.g., Metalsky, Halberstadt, & Abramson, 1987; Peterson & Seligman, 1984). Indeed, we agree with many of the criticisms and suggestions of Cutrona et al. (1984) and others (e.g.,
Feather & Tiggemann, 1984) concerning development and refinement of the AS concept in general. Rather, we present old and new data relevant to several general issues raised about AS primarily on the basis of different versions of the Attributional Style Assessment Test (ASAT) originally developed by Anderson et al. (1983).

We focus on this limited set of studies for several reasons. First, the ASQ has been subjected to numerous reviews and a recent meta-analysis, so further review of the ASQ is unnecessary (Sweeney, Anderson, & Bailey, 1986). Second, the ASAT scales were specifically designed to distinguish between ASs for different types of situations, and have been administered along with measures of several different types of problems in living (i.e., depression, shyness, and loneliness) that theoretically should show differential relations with the different measures of AS. This allows a better examination of convergent and discriminant validity issues than is usually the case with ASQ studies of depression. Third and perhaps most important, reviews of the ASQ are likely somewhat biased by reporting and publication biases. That is, studies that “work” are more likely to be submitted and accepted for publication than are failures. Thus, it is hard to know whether or to what extent the reviewed findings are accurate reflections of reality (but see Sweeney et al., 1986, for evidence that this bias did not distort their findings). By pooling the results from all of our studies using various versions of the ASAT, we minimize such reporting and publication biases. Finally, we have two studies in which the effects of AS can be compared with the effects of corresponding attribution manipulations in complex social situations. Thus, we can roughly estimate the importance of these two sources of attributions. Before examining the ASAT data we turn to a brief discussion of the issues to be addressed.

VALIDITY OF ATTRIBUTIONAL STYLE

There are many different meanings of validity that may be addressed, the most important including construct, convergent, and discriminant validity. Essentially, construct validity means that a measure actually measures the theoretical construct it is supposed to measure. To demonstrate construct validity one needs to demonstrate that the measure correlates with other measures (such as scales or behavioral indexes) with which it theoretically should correlate. This is known as convergent validity. In addition, one needs to show that the measure does not correlate with other measures with which it theoretically should not correlate. This is known as discriminant validity. Straightforward but somewhat elaborate procedures have been established for assessing these types of validity. The most common is the multitrait-multimethod matrix procedure of Campbell and Fiske (1959). Construct validity may also be tested experimentally. If experimental manipulations produce theoretically expected changes in a measure, those results support the claim that the measure has construct validity (Nunnally, 1978).

The construct validity of the AS concept has not been rigorously tested. Our data were not specifically generated to examine construct validity issues comprehensively. However, we have assessed several conceptually distinct ASs in two quite different ways. We have also looked at the intercorrelations of AS measures and have examined AS correlations with several measures of different problems in living that should relate to AS in slightly different ways. Finally, we have looked at how ASs measured in two different ways relate to complex social behaviors in two different settings. Thus, our data do address the validity issues raised earlier. It is also important to note that our analyses include all the data we have gathered concerning these issues (both published and unpublished), so that the reporting and publication biases usually present in our field are quite reduced here.

In sum, the studies presented provide good tests of convergent validity, some indication (but admittedly insufficient) of discriminant validity, and thus some evidence of the construct validity.

SPECIFICITY VERSUS GENERALITY OF ATTRIBUTIONAL STYLE

An important theoretical and empirical distinction concerns the AS construct’s specificity versus generality. At the most general level, one could hypothesize that a person’s AS is consistent across all possible situations. This view maintains that regardless of the type of situation (e.g., interpersonal vs. noninterpersonal), individuals’ relative standings on some attribution measure will remain constant.

As pointed out by Cutrona et al. (1984) and others, the prevailing view of AS is that it is quite general. That is, a person’s AS is assumed to be totally cross-situationally consistent. Evidence in support of this assumption would consist of high correlations between attributions for very different types of situations and similar validity coefficients for ASs assessed for very different types of situations. This is implied by the standard ASQ procedure of collapsing AS scores across the affiliation and achievement items before correlating with measures of depression (e.g., Seligman et al., 1979). In addition, Seligman has reported (personal communication, 1979) that the AS correlations with depression do not differ for achievement versus affiliation items in either the successful or unsuccessful situations.

At a less general level, one could hypothesize that a person’s AS is cross-situationally consistent only across situations that are similar in psychologically meaningful ways, but not across very diverse types of situations. This view maintains that within a situation type (e.g., interpersonal failures) the relative standing of individuals will remain fairly constant from one situation to another; but that between different situation types (e.g., interpersonal failure vs. noninterpersonal success) there will be little correspondence.

This view of AS as an individual difference construct at a moderate level of specificity demands that close attention be paid to a taxonomy of situation types. There are, of course, numerous ways of classifying situations. Therefore, finding a situation taxonomy that captures important AS differences may seem to be a formidable task. That task may be guided by theoretical and empirical considerations arising from analyses of the important problems to be linked to AS, such as depression, loneliness, and shyness.

Finally, at the most extreme level of specificity, one could hypothesize that people’s relative standing on some attribution measure would show no consistency even within types of situa-
If this is true, then AS falls apart as an individual difference construct, and one is left with the position that attributions are totally idiosyncratically and situationally determined.

As pointed out in the preceding paragraph, the prevailing view of AS is that it is totally (or nearly totally) cross-situationally consistent. This view would seem to fly in the face of the voluminous literature showing that very few measured constructs are consistent across situations (e.g., Mischel, 1968). Moreover, research from several attribution paradigms has yielded situational differences. The first author has investigated the causal structure of different types of situations, that is, the extent to which the types of causes people consider as possibilities for a situation depends on the type of situation being considered. It has been shown that different types of situations do have different causal structures, and that the causal structure of situations is closely related to attributions made for the situations (Anderson, 1983a, 1985).

Research on the concepts of depression and loneliness also demonstrates some situational specificity suggestive of the importance of distinguishing among different types of situations. For example, the loneliness prototype consists of exclusively interpersonal problems, whereas the depression prototype includes both interpersonal and noninterpersonal problems (Horowitz, French, & Anderson, 1982). Furthermore, research from several laboratories has provided some evidence that AS is most appropriately assessed at an intermediate level of specificity (Anderson & Arnoult, 1985a, 1985b; Anderson et al., 1983; Metalsky et al., 1987).

Our current view is that fairly consistent ASs do exist at a moderate level of specificity (e.g., interpersonal failure situations). We also believe that the generality of ASs is most apparent (or most exaggerated) when assessed for hypothetical (but familiar) situations. In some sense, attributions in these cases are likely to be based on schematiclike judgments.

Consequently, even if AS is perfectly assessed (i.e., without error) we do not expect the intercorrelations of ASs for different types of situations to approach 1. In contrast, viewing AS as completely general across situations does lead to the prediction of perfect correlations, when corrections for the unreliability of the scales are applied. Similarly, we do not expect the cross-situation-type intercorrelations or the within-situation-type reliabilities to be close to 0, as would be the case if there were no cross-situational consistencies in ASs. The studies presented include data allowing examination of all of these predictions.

If attributions influence complex social behaviors only when assessed at a very specific level, then such effects should be quite weak (or nonexistent) when assessed at an intermediate level. That is, if AS is an invalid construct at an intermediate level, then using that intermediate-level construct should result in poor prediction of behavior; conversely, using a very specific attribution should result in a much better prediction of behavior. Our data also allow us to examine this version of the specificity issue.

ATTRIBUTIONAL STYLE AND SITUATIONAL EFFECTS ON SOCIAL BEHAVIOR

The final set of issues to be addressed concerns the effects of AS on consequential, complex social behaviors. Intimately tied to this question is the issue of AS effect size relative to situational effects, and Person × Situation interactions. For some time it has been quite popular for social and personality researchers to bemoan the low percentage of variance explained by individual difference measures. The implicit assumption has been that situational variables do much better. More recently, the interactionist view has also become popular, with the explicit assumption that it is the interaction of situational and personality factors that accounts for the greatest portion of the variability in behavior (e.g., Bowers, 1973; Magnusson & Ender, 1977). However, the empirical basis for the interactionist position is weak at best (see Funder & Ozer, 1983; Sarason, Smith, & Diener, 1975). The percentage of variance accounted for by the situation and by the interaction is not any more impressive in general than the percentage accounted for by the person. (However, see Golding, 1977, for a brief discussion of problems in assessing interactionist effects.)

In the attribution domain, there has been almost no research on the effects of AS on consequential social behaviors. Similarly, practically no attempt has been made to compare the effect size of AS and parallel situational attribution manipulations. Our position is that AS is a valid and influential construct. Further, we believe that when the situation under examination is prototypical (Schatte, Kenrick, & Sadalla, 1985) of the AS under investigation, both AS and parallel attribution manipulations will affect complex social behaviors. We expect that when the attribution manipulation is pitted against AS in an unfamiliar (even though prototypic) setting, the manipulation will have relatively more impact. In general, however, we expect both to have important consequences for social behavior.

OVERVIEW

To examine these issues on the validity, specificity, and behavioral importance of AS we combined and reanalyzed all of our studies on AS. Four had been published previously in psychological journals in some form: two appeared in Anderson et al. (1983), one in Anderson (1983b); and one in Anderson and Arnoult (1985b). We included two additional studies: one a replication of the Anderson and Arnoult (1985b) AS study of depression, loneliness, and shyness; the other Jennings's (1980) dissertation study of AS and attribution manipulation effects on responses to failure at interpersonal persuasion. Of the six studies, four were correlational ones relating ASs to clinical symptoms such as depression, loneliness, and shyness. The other two studies involved preselection of subjects on the basis of their ASs, manipulation of attributions, and observation of performance in interpersonal persuasion contexts. Most of the analyses in this article have not been reported previously and could not be derived from previously published work.

Correlation Studies

Method

Description of Studies

Two of the four correlation studies were originally reported in Anderson et al. (1983). In that article, the development of two versions of the ASAT was reported, along with data examining the relations between
AS and depression and loneliness. For details about scale construction and the basic results, the reader should consult the original article. The main points are as follows. First, a number of situations common to college students (the target population) were described with success and failure outcomes. Pilot subjects wrote open-ended attributions to those situations as if they had been in them. A content analysis of the open-ended responses was conducted to derive a small number of fairly general and inclusive types of attributions. The six resulting categories were strategy, ability, effort, personality trait, mood, and other circumstances.

On the basis of other pilot data, five situations rated as interpersonal and five rated as noninterpersonal were selected for inclusion in the ASAT. Each was presented in both a success and a failure form. Each of these 20 items had brief attributional statements included with it, one for each of the six categories listed in the preceding paragraph. Here is an example of an interpersonal failure item:

You have just attended a party for new students and failed to make any new friends.

a. I used the wrong strategy to meet people.
b. I am not good at meeting people at parties.
c. I did not try hard to meet people.
d. I do not have the personality traits necessary for meeting new people.
e. I was not in the right mood for meeting new people.
f. Other circumstances (people, situations, etc.) produced this outcome.

To complete the questionnaire, a person imagines himself or herself in each situation, then circles the cause that would best explain the stated outcome. This version (ASAT-I) thus contains 20 items allowing assessment of ASs for four types of situations: interpersonal success and failure and noninterpersonal success and failure.

Although numerous attribution measures can be derived from this instrument, we focus on a summary measure labeled apparent changeability by Anderson et al. (1983, p. 133). This measure, now labeled controllability, was calculated for each situation type by combining the number of attributional choices, as follows:

\[
\text{controllability} = (\# \text{strategy} + \# \text{effort}) - (\# \text{ability} + \# \text{trait}).
\]  

In essence, for each of the five items of each situation type, 1 is added to the controllability score if strategy or effort is chosen, 0 is subtracted if ability or trait is chosen, and 0 is added if mood or other is chosen. High scores thus indicate that the subject tends to select as attributions those factors that he or she can change, modify, or control. Low scores indicate a preference for factors that cannot be changed, modified, or controlled. Note that all four of these attributions are generally internal to the person.

The ASAT-I was administered to a large group of undergraduates, along with the Beck Depression Inventory (BDI; Beck & Beck, 1972) and the UCLA Loneliness Scale (LS; Russell, Peplau, & Ferguson, 1978). Approximately 300 subjects completed these forms.

For the second study, a 36-item version of the ASAT was used. In addition to the increase in items, the ASAT-II also differed in the attributional choices offered. Instead of the six described earlier, the ASAT-II included only ability, effort, and strategy choices. The measure of controllability (i.e., changeability) was the negation of the number of ability choices, because the number of ability choices determines the number of strategy + effort choices. (The negation makes the direction comparable to the ASAT-I controllability index.) In this second study, approximately 120 undergraduate subjects completed the ASAT-II, the BDI, and the LS.

The third correlational study was originally reported in Anderson and Arnoult (1985b). In that study, the rating scale format used in the ASQ was adapted to the 20 situations from the ASAT. That is, subjects generated an open-ended attribution for each situation and rated their own attributions on each of several causal dimensions. The rated dimensions included locus (internality), stability, and globality, as in the ASQ. In addition, intentionality and controllability were rated to provide tests of several competing attributional models. This version (ASAT-III) and modified measures of depression (MDI), loneliness (MLS), and shyness (MSS) (see Anderson & Harvey, 1988) were completed by 207 undergraduates. The major finding was that only the controllability and locus dimensions contributed significant unique increments to the prediction of depression, loneliness, and shyness. Therefore, we include here the results of the unit-weight two-dimension attributional styles (UWAS), computed as follows:

For failure situations, \[
\text{UWAS} = \text{control score} - \text{locus score}. \tag{2a}
\]

For success situations, \[
\text{UWAS} = \text{control score} + \text{locus score}. \tag{2b}
\]

For failure, high UWAS scores indicate controllable and external attributions. For success, high UWAS scores indicate controllable and internal attributions. Note that the conclusions to be drawn later in this article do not change if the AS measure in the third and fourth studies included only the controllability dimension.

The fourth correlational study was a replication of the third, with a fourth version of the ASAT (ASAT-IV). This version was the same as the ASAT-III, except that 36 items (as in the ASAT-II) were used instead of 20. In this study, 206 undergraduates completed the ASAT-IV, the MDI, the MLS, and the MSS. Again, we include here only the composite AS results (i.e., the UWAS). This study has not been previously reported.

Analyses

For all analyses, the results of the forced-choice studies (ASAT-I and -II) and the rating scale studies (ASAT-III and -IV) were averaged separately, using the \( \tau \) to \( \tau \) weighted averaging procedure. We conducted four types of analyses. First, we computed reliabilities of the AS measures using Cronbach’s alpha. If attributions are completely specific to each situation, then the reliabilities should be about 0. Second, we computed the intercorrelations of the AS measures for different types of situations. If AS is extremely idiosyncratic and not at all cross-situationally consistent, these correlations should essentially be 0. Third, we computed the intercorrelations of AS measures corrected for unreliability (Pedhazur, 1982). If AS is very general across situations, these corrected intercorrelations should be perfect (1). Corrected intercorrelations of moderate size would suggest moderate specificity in AS, whereas corrected intercorrelations near 0 would indicate that AS is specific to each of the four types of situations used. Fourth, we computed (and averaged across studies) the pattern of AS correlations with the clinical symptoms. These correlations provide further evidence regarding the convergent and discriminant validity of the AS construct.

Results and Discussion

Attributional Style Intercorrelations

The reliabilities, raw intercorrelations, and corrected intercorrelations of the AS measures for the four situation types are presented in Tables 1 and 2. The diagonals of Tables 1 and 2 contain the reliability estimates (alpha) of the AS measures. For both the forced-choice versions and the rating scale versions the reliabilities are modest, ranging from .51 to .60. These are about the same as for other AS measures (e.g., Cutrona et al., 1984), and are sufficient for use in large sample studies. More crucial to our discussion is the fact that these moderate levels of internal consistency pro-
vide good evidence that within situation types there is a modest level of cross-situational consistency.

The raw intercorrelations between AS assessed for different types of situations are presented above the diagonals of Tables 1 and 2. These correlations are uniformly positive and of small to moderate magnitude. All but one are significantly different from 0, demonstrating some cross-situational consistency.¹ That is, these results also suggest that AS is a valid construct at either a moderate or a general level; otherwise, the intercorrelations would have to be essentially 0.

The raw intercorrelations discussed in the preceding paragraph give one a picture of the lower limit of the true relations among AS measures. To get a picture of the upper limit, we corrected these estimates using the reliabilities shown on the diagonals of Tables 1 and 2. If AS is quite cross-situationally consistent, as has been assumed in the past, the corrected intercorrelations should approach unity. The corrected intercorrelations, presented below the diagonals, did increase in size but did not generally approach 1. This is in line with our moderate level of specificity hypothesis.

Additional comparisons of interest concern the relative magnitude of cross-situational consistency as a function of the similarity of the items in the different AS measures. The items differed on two dimensions, outcome (success or failure) and interpersonalness (interpersonal or noninterpersonal). Recall that each situation was presented twice, once as a success and once as a failure. Thus, to the extent that subjects perceived any demand to respond consistently, one might expect AS correlations that are based on the same situations (e.g., the success and failure versions of “attempting to cheer up your roommate”) to be artificially inflated and relatively larger than other intercorrelations.

Psychologically, outcomes are powerful organizing themes. Thus, one might expect intercorrelations between same outcome AS measures (e.g., interpersonal failure and noninterpersonal failure) to be the highest. If subjects responded to each item accurately and without perceived consistency demands, then the same outcome correlations should be the highest; the same interpersonalness correlations and the correlations that were not the same on either should be about equal and somewhat lower. The results followed this latter predicted pattern exactly. The average intercorrelations for same outcome, same interpersonalness, and same on neither AS measure were .40, .17, and .14, respectively.

In sum, these data provide strong evidence that AS is cross-situationally consistent when assessed at a moderate level of specificity. In addition, they provide some evidence of both convergent and divergent validity. The AS measures that should theoretically correlate the highest with each other did so; those that should correlate only weakly also did so as well.

### Attributional Style Correlations With Problems in Living

We conducted further examinations of the construct validity of AS by comparing correlations among AS measures and measures of depression, loneliness, and shyness. These correlations are presented in Tables 3 and 4.

Theoretically, depression is less specifically tied to interpersonal situations than is loneliness or shyness. This is because depression contains both interpersonal and noninterpersonal features, whereas loneliness contains only interpersonal features (e.g., Horowitz et al., 1982). We also assume shyness to be primarily interpersonal, but the relevant prototype analysis (comparable to Horowitz et al., 1982) has not been conducted.

Furthermore, because success is seldom disruptive regardless of its perceived causes, we expected the interpersonal–noninterpersonal specificity of correlations with various problems in living to appear only (or more strongly) in failure settings. In short, we expected an interactionlike pattern of correlations among the problems in living and AS for failures. Specifically, within the failure situations the depression–AS correlations should be more equal across the interpersonal–noninterpersonal distinction than the loneliness and the shyness–AS correlations. Stated differently, the average of the loneliness and shyness correlations with interpersonal failure AS and the depression correlations with noninterpersonal failure AS should be stronger (i.e., more negative) than the average of the remaining failure AS–problems-in-living correlations. The results, as shown in Tables 3 and 4, support this interaction prediction. Unfortunately, there are no established procedures for testing this prediction.

We used a conservative procedure to get some idea of the reli-

---

¹ All p levels are based on two-tailed tests.
Table 3
**Weighted Average Correlations From Forced Choice Studies: Controllability Attributional Styles as Predictors of Depression and Loneliness**

<table>
<thead>
<tr>
<th>Situation type</th>
<th>Interpersonal failure</th>
<th>Noninterpersonal failure</th>
<th>Interpersonal success</th>
<th>Noninterpersonal success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>-.36*</td>
<td>-.22*</td>
<td>-.02</td>
<td>-.04</td>
</tr>
<tr>
<td>Loneliness</td>
<td>-.44*</td>
<td>-.18*</td>
<td>.08</td>
<td>-.10</td>
</tr>
</tbody>
</table>

*Note. Ns are approximately 420. *p < .001.

ability of the obtained interaction pattern. First, we transformed the correlations to z scores using the standard r-to-z transformation. The variance of such z scores is, of course, 1 divided by N. Now if each z were independent (i.e., based on separate subjects), then a linear combination of the zs such as would be used to test the interaction pattern would have as a variance the sum of the variances of the zs. Thus, to test the interaction pattern one would simply create a contrast reflecting the predicted pattern of correlations (zs) and divide by the square root of the sum of the variances. This is a Z test. For the results presented in Table 3, this test would consist of summing the loneliness-interpersonal failure and the depression-noninterpersonal failure zs, and subtracting from this total the sum of the depression-interpersonal and loneliness-noninterpersonal failure zs. This contrast score would then be divided by the square root of 4 divided by 420. This same procedure could also be applied to the data in Table 4, with appropriate contrasts applied to the shyness results included.

The obvious problem with this procedure is that the correlations (and the zs) were not independent. The effect of nonindependence in this situation is to reduce the error variance. That is, the variance of a linear combination of nonindependent parameters is the sum of the variances (as in the preceding paragraph) minus the covariances. We were unable to compute the amount of covariance to be subtracted in this rather unusual circumstance, so nothing was subtracted. Thus, to the extent that the correlations within Table 3 and Table 4 are not independent, the procedure outlined in the preceding paragraph results in too large an estimate of error variance and too small an estimate of the Z test statistic. In fact, the variables involved are all intercorrelated at a fairly high level, guaranteeing that our procedure produces overly conservative results.

Nonetheless, we used this procedure to test our interaction prediction that within failure situations, the depression–AS relations would be more equal across the interpersonal–noninterpersonal distinction than would the loneliness and shyness–AS relations. This was done independently for the data in Tables 3 and 4. Finally, we combined the results of these two separate tests using a chi-square procedure suggested by Winer (1971). Even this conservative procedure resulted in a reliable confirmation of our prediction, \( \chi^2(4) = 12.79, p < .03 \).

In addition to this overall interaction analysis, we conducted a series of t tests for differences between correlated correlations, based on our averaged data in Tables 1 through 4. Using a .01 significance criterion (to protect against excessive Type I errors with the large number of tests involved), we found that the depression correlations with interpersonal and noninterpersonal failure AS were not different in either the forced-choice or the rating scale versions, but that the same comparisons for loneliness and shyness were all significantly different, all ts = 2.70, ps < .008. In essence, tests of the construct validity of our moderately specific AS constructs were supportive.

Two qualifiers to these claims are warranted here. First, note that in Table 3 loneliness did correlate reliably with noninterpersonal failure AS. Second, note that in Table 3 the AS–depression correlation was slightly (although nonsignificantly) smaller for noninterpersonal failure than for interpersonal failure. Thus, although the preceding significance tests of differences between correlations support both the convergent and divergent validity claims for AS, the pattern is not perfect. Although we cannot be sure of the reason for these discrepancies, we have a suspect. Both discrepancies came from studies in which the measures of loneliness and depression had not been modified to eliminate confounding items. Anderson and Harvey (1988) have shown that several items on these scales have higher factor loadings on the other construct (i.e., higher loading on depression for a loneliness item, and higher loading on loneliness for a depression item). The results in Table 4 were based on measures of depression, loneliness, and shyness modified to be as factorially pure as possible; there were no anomalies in those results.

Finally, one very striking difference between the forced-choice and rating scale results emerged. Success AS measures were uniformly predictive of problems in living when the rating scale method was used, but were uniformly unpredictive when the forced-choice method was used. Perhaps the limited choices available in the forced choice ASAs were simply too limited in scope and thus insensitive to success AS differences. This possibility and others warrant further work.

Further refinements in measurement techniques are certainly desirable, as are more sophisticated studies of the validity and specificity of the AS constructs. It is clear nonetheless that AS assessed at moderate levels of specificity is a valid construct (actually, set of constructs) showing both convergent and discriminant validity. We turn now to our final question concerning the effects of AS and attribution manipulations on complex social behaviors.

**Studies of Complex Social Behaviors**

A second approach to the questions concerning the validity and utility of our view of AS is to assess the impact of AS differ-
ATTRIBUTIONAL STYLE VALIDITY AND UTILITY

Table 4
Weighted Average Correlations From Rating Scale Studies: Unit-Weight Attributional Styles as Predictors of Depression, Loneliness, and Shyness

<table>
<thead>
<tr>
<th>Situation type</th>
<th>( \text{Personal problem} )</th>
<th>( \text{Interpersonal failure} )</th>
<th>( \text{Noninterpersonal failure} )</th>
<th>( \text{Interpersonal success} )</th>
<th>( \text{Noninterpersonal success} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression</td>
<td>-21**</td>
<td>-15*</td>
<td>-20**</td>
<td>-23**</td>
<td></td>
</tr>
<tr>
<td>Loneliness</td>
<td>-22**</td>
<td>-08</td>
<td>-26**</td>
<td>-24**</td>
<td></td>
</tr>
<tr>
<td>Shyness</td>
<td>-24**</td>
<td>-08</td>
<td>-21**</td>
<td>-23**</td>
<td></td>
</tr>
</tbody>
</table>

Note. \( N = 413 \).
\( * p < .01 \). \( ** p < .001 \).

ences in complex social settings. The two studies that have done this follow. Both adopted the same basic rationale. Some subjects were preselected on the basis of their attributional responses to hypothetical situations of the same type (i.e., interpersonal failure) as the target task situation. In both experiments, the preselected groups had controllable versus uncontrollable AS for interpersonal failures. Prior to performing the target task, some subjects received attribution manipulations paralleling the preselection differences.

If AS is cross-situationally consistent (at a moderate level of specificity, at least), then the AS factor should have important effects on performance at the target task. Specifically, those with a controllable AS for interpersonal failure should respond more adaptively to initial failures on the target task. Adaptive responding to interpersonal persuasion failures (the target task) includes maintaining high success expectancies and high motivation, making strategic changes to improve performance, and actually improving the persuasion performance. These effects of AS should be paralleled by comparable effects of the parallel attribution manipulations (AMs). Finally, we expected that AS effects would be of an important magnitude. Specifically, we expected AS effects to be about as large as the parallel effects of AM.

The first study we considered was Jennings (1980). This study has not been published previously, and so the method is presented in considerable detail. Note that our analyses are somewhat different than those originally conducted by Jennings.

Study 1: Jennings (1980)

Method

Subjects. The subjects were 32 male and 28 female undergraduates enrolled in introductory psychology. For their participation, they received either $3.00 or credit toward a class requirement. Thirty of these subjects (17 men and 13 women) were randomly assigned to the AM conditions.

The remaining 30 subjects (15 men and 15 women) were recruited and assigned to the AS conditions on the basis of their responses to an item embedded in a large questionnaire completed 1 month prior to the study. This item had students imagine themselves giving an important talk in front of an audience that reacts negatively. Subjects were to specify the major cause for this failure. Subjects then used 7-point scales to rate the locus (external or internal) and the stability (unstable or stable) of the cause. From this pool of respondents, two groups of 15 students each were recruited. One group had rated the cause as uncontrollable (defined as internal and stable); the second had rated the cause as controllable (defined as internal and unstable).

Procedure. On arrival, each subject was presented with a brief introduction to a study on persuasion techniques being conducted by the Red Cross in preparation for an upcoming blood donor drive on campus. Specifically, each subject was told that the Red Cross was planning to conduct a radio campaign on campus in which student volunteers were to make short pitches on the campus radio station attempting to persuade their peers to donate blood. The experimenter explained that past radio campaigns had not been as effective as the Red Cross would like. Although some volunteers were successful in getting a large number of people to donate blood, others were unsuccessful in getting anyone to donate.

Attribution manipulation. In both conditions, the experimenter first said, "We're conducting this experiment because no one in the Red Cross really knows why some volunteers are more successful than others." At this point, the type of attributions primed for subjects in the AM conditions were manipulated. The experimenter commented that in the opinion of a Red Cross staff member, persuading people to donate blood is a task in which either abilities (uncontrollable condition) or strategies (controllable condition) determine a volunteer's success or failure.

For the subjects in the AS conditions, the experimenter made no attempt to influence subjects' attributions for failure. He merely indicated that "We're conducting this experiment because no one in the Red Cross really knows why some volunteers are more successful than others." However, recall that these subjects had been selected for their preexisting tendency to attribute failure at a similar communication task to either uncontrollable or controllable factors. Thus, these predispositions should have produced a difference in attributions that parallels the difference in the AM conditions.

Perturbation task. In all conditions, the experimenter then explained that a radio campaign was being simulated in order to discover what factors influence a person's success or failure at persuading others over the radio to donate blood. The experimenter indicated that the subject would have three presentation opportunities in the experiment, and that the subject would have 5 min immediately prior to each presentation to prepare. The experimenter explained that he would provide the subject with presentation effectiveness feedback on each of the trials. A different expert judge was to listen to each presentation broadcast on a loudspeaker in an adjoining room (in order to simulate a radio broadcast) and then estimate the percentage of students that might donate if they heard the presentation over the radio. The experimenter emphasized that each judge would hear and evaluate only one of the subject's presentations. Thus, each evaluation was independent of the others. The experimenter then gave the subject information about the blood dona-
tion procedure, and told the subject to prepare for the first persuasive
attempt.

Failure manipulation. During the experiment, each subject made a
total of three presentations. About 1 min after each of these presenta-
tions, a voice over the intercom indicated that the subject’s presentation
was weak and would have failed to persuade even 30% (supposedly an
average level of effectiveness) of students to donate. Although subjects
were led to believe that each voice belonged to a different expert judge
in an adjoining room, the voices were actually tape-recorded statements
heard by all subjects. This procedure guaranteed that all subjects re-
ceived identical feedback that was in no way contingent on their actual
performance.

Measures of strategy use and change. To determine whether subjects’
attributions affected the actual content of their persuasive attempts,
each presentation was recorded and later rated by three independent
judges (one graduate and two undergraduate psychology majors) who
were blind to the experimental condition of the subjects.

Judges were given a list of 14 persuasive tactics and strategies and
asked to indicate which one(s) each subject used on each presentation.
These strategies were using direct commands or direct requests; repeat-
ing previous arguments; emphasizing the personal benefits, ease and
safety, or fun of donating blood; inducing fear, guilt, duty and obliga-
tion, compassion, or pride; claiming personal expertise; using peer pres-
Sure; and emphasizing the critical shortage of blood. These strategies
were either power strategies described by Falbo (1977) or strategies that
subjects mentioned in a postexperimental questionnaire that asked
them to list “all the strategies and tactics you could have used to per-
suade students to donate blood.” A subject was given credit for using
a strategy only if at least two of the three judges independently
checked it.

Self-ratings of preparedness and success expectancy. To determine if
subjects’ attributions affected their predictions of success, subjects were
asked to predict how effective their next presentation would be. Specif-
ically, immediately prior to giving each presentation, subjects were asked
to predict the percentage of students that their presentation would suc-
cessfully persuade if it were to be broadcast over the radio. This may be
regarded as a measure of perceived preparedness.

Following the final presentation, subjects were asked to predict the
percentage of students they would successfully persuade over the radio
in an upcoming blood drive on campus, assuming they had a chance to
practice. This may be regarded as a measure of future success expectan-
cies.

Presentation effectiveness. Finally, to determine if subjects’ attribu-
tions affected the quality of their presentations, the judges rated the per-
suasive effectiveness of each presentation. For subject’s first presenta-
tion, judges estimated the initial level of effectiveness using a scale an-
chored at Not at all effective (coded as 1) and Extremely effective (coded
as 9). For subject’s later presentations, judges estimated how much these
initial levels had increased or decreased from the immediately preced-
ing presentations using scales anchored at Much less effective (1) and
Much more effective (9). The interjudge reliabilities as measured by
Cronbach’s alpha for the three presentations were .67, .64, and .66. The
judges’ ratings were then averaged to form a single composite index of
effectiveness for each subject for each presentation.

Debriefing. After all dependent measures had been obtained for a
given subject, the experimenter asked a series of questions designed to
assess the subject’s suspicions regarding the various experimental ma-
ipulations and deceptions used (Carlsmith, Ellsworth, & Aronson,
1976). No subjects detected that the judges’ evaluations during the ex-
periment were actually prerecorded messages heard by all subjects, and
no subjects guessed the hypotheses regarding the use and change of
strategies following failure. The experimenter then explained the true
purpose of the experiment, concentrating on the impact the outcome
manipulation might have had on the subjects’ self-impressions. The
“process debriefing” procedure (cf. Ross, Lepper, & Hubbard, 1975)
has proven effective in eliminating the persistent impact of false feed
back.

Results and Discussion

Three types of measures are of primary interest: success expectan-
tcies, strategy use, and performance quality. All dependent
variables were assessed after the attributional sets or primes (AS or AM) were in operation. Each type of measure
was assessed both before and after subjects’ failure feedback. A
series of 2 (attribute type: controllable vs. uncontrollable) X
2 (attribute source: AS or AM) analyses of variance was con-
ducted on these dependent variables and on derived composite
scores.

Success expectancies. Before each presentation, subjects esti-
minated the percentage of listeners who would be persuaded to
donate blood (recall that they had been told to expect about a
30% success rate). On these measures of perceived preparedness
for an immediate performance, there were no significant main
or interaction effects, all Fs < 2.07, ps > .5. Subjects predicted
mean success rates of 26.2%, 26.1%, and 25.9%, respectively,
for their first, second, and third presentations.

ashed after all three presentations and failure feedback were com-
pleted, subjects were asked how successful they would be in an
upcoming blood drive if they had a chance to practice. These
future success expectations were significantly influenced by at-
tribution type. Subjects in the controllable conditions predicted
a mean success rate of 38.8%, whereas those in the uncontrolla-
ble conditions predicted a mean success rate of only 33.0%. This
main effect was significant, £(1, 56) = 5.06, p < .03. Nei-
ther the attribution source nor the interaction were significant,
Fs < 1.

In sum, subjects in different conditions judged themselves to
be equally prepared for the immediate task. However, those in
the controllable conditions were more optimistic about their fu-
ture performances than were those in the uncontrollable con-
ditions, just as the attribution model predicts. In addition, there
was no hint of an Attribution Source X Attribution Type in-
teraction, indicating that the effects of the two attribution sources
(AS and AM) were of approximately equal size.

Strategy use. The number and type of strategies used in each
presentation were assessed by raters blind to subjects’ condi-
tions. Two ideas are of interest here. First, did the subjects’ con-
ditions influence the number or type of strategies selected for
use in their first presentation? Given that the time for each com-
mercial was tightly controlled, one might expect no differences
in the number of strategies used; one must do something with
that air time. Results indicated no systematic differences. Sub-
jects used 2.9 strategies on average; there were no main or inter-
action effects, Fs < 1. Furthermore, there were no significant
differences in the proportion of subjects in the four conditions
that used each of the 14 strategies examined by the judges.

Second, did subjects’ conditions influence their tendencies to
modify their strategies in subsequent presentations? The at-
tribution model predicts such differences. If one thinks that per-
formance quality is based on controllable factors such as strat-
ey, one should be more willing to modify failed strategies (An-
derson & Jennings, 1980). To test this, we summed the number
of strategic changes from first to second presentation, and from second to third presentation, for each subject. As predicted, controllable-condition subjects shifted strategies significantly more often than did uncontrollable-condition subjects (Ms = 3.4 and 1.5, respectively), F(1, 56) = 10.19, p < .0005. Neither the main effect of attribution source nor the interaction were significant, Fs < 1. In brief, both the AS and AM sources of attributions had powerful effects of approximately equal size on strategy shifts.

Performance effectiveness. Two questions are of primary interest here. First, to what extent did the attributional sets and primes influence the quality of the first presentations? A person who believes that performance quality on an upcoming task depends on controllable factors (such as strategy) will be more likely to consider the specifics of his or her presentation carefully than will one who believes performance quality is uncontrollable. Thus, one might expect quality differences to emerge early in the sequence of events. At the least, controllable subjects should be performing better than uncontrollable subjects by the third presentation. Thus, the second (and more critical) question is whether or not performance differences have emerged by the final presentation.

We examined the effectiveness ratings on the first presentation to answer the first question. Controllable-condition subjects did in fact produce better presentations than uncontrollable-condition subjects (Ms = 4.31 and 3.79, respectively), F(1, 56) = 5.92, p < .02. The main effect of attribution source and the interaction were both nonsignificant, Fs < 1.

To assess the overall quality of the final presentations, we created a composite measure by summing the effectiveness ratings on first presentation with the change in effectiveness ratings on the second and third presentations. These overall performance scores yielded the same pattern as the first presentation data; there was a main effect of attribution type, F(1, 56) = 8.34, p < .01, but neither a main effect of attribution source nor an interaction, Fs < 1. The attribution type effect was that controllable condition subjects performed better than did uncontrollable-condition subjects.

Composite index. To better illustrate the effects of AS and AM on subjects' reactions to this complex social situation, we created an overall index of future expectancies, strategy shifts, and performance effectiveness. We did this by computing z scores on each of these three measures and taking the average of the three as the index score for each subject. At a psychological level, this composite index is best thought of as an indication of how well people respond to the situation. Low scores indicate a withdrawal or depressive response, whereas high scores indicate an approach or optimistic response. At a statistical level, this index is essentially a multivariate dependent variable (as in a multivariate analysis of variance) with the specification that all three components be weighted equally.

The results are presented in Figure 1. As can be seen, the attribution type main effect was quite strong, F(1, 56) = 26.54, p < .0001. That is, subjects in the controllable conditions responded better to the failure situation than did those in the uncontrollable conditions. The main effect of attribution source and the interaction were both nonsignificant, Fs < 1. In other words, both attribution sources (AS and AM) reliably affected responses in this situation, and did so to an approximately equal extent. Indeed, AS had a slightly larger effect than did AM.

Study 2: Anderson (1983b)

The Jennings (1980) study showed powerful effects of attribution type (controllable vs. uncontrollable) on a complex social behavior. Preselection based on attributions for one hypothetical situation 1 month prior to task engagement resulted in reliable attributional effects on success expectancies, strategy use during task performance, and quality of task performance. These same attribution type effects were produced by a subtle manipulation consisting of a casual statement made by the experimenter about possible causes of success and failure in the upcoming task. Of additional importance was the lack of any Attribution Source × Attribution Type interaction. Essentially, the attribution type effects of both attribution sources (AS and AM) were equal in size. Indeed, the direction of the trivial differences that did emerge suggests that AS had a slightly greater impact. Thus, those who would abandon the individual difference approach because of the small amount of variance explained by such variables would also have to abandon the situational approach.

The second study to examine simultaneously AS and AM effects in complex social settings has been previously reported (Anderson, 1983b). It differed from Jennings's work in several important respects. First, preselection was based on a 5-item scale designed to measure controllability (changeability) AS for interpersonal failures in general, rather than on the basis of a

---

2 We also examined the change in overall effectiveness ratings for the second and third presentations. Because the change ratings were correlated with ratings of the first presentation, covariance analyses were conducted. These analyses revealed no systematic effects on the second presentation. On the third presentation, controllable-condition subjects improved more than did uncontrollable-condition subjects, F(1, 55) = 4.47, p < .04.
single item that was very similar to the eventual task situation. Second, the AS assessment techniques were very different. Anderson's subjects responded to items in a forced-choice format (ASAT-I), whereas Jennings's (1980) subjects generated their own cause for the item and rated it on two causal dimensions. Third, the subjects' tasks were quite different. In Anderson's study the task was to call people on the telephone and try to convince them to donate blood, whereas in Jennings's study subjects prepared and presented radio appeals. The major advantage of examining studies having these different features is that if similar results occurred in both studies, conclusions about attribution effects would be more generalizable and would be seen as contributing to the convergent validity of the AS construct.

One other difference to be noted concerns the design of the studies. In Jennings's (1980) study, subjects' attributions were based on either their AS or on the experimenter's manipulations. This allowed a simple test of which source had the greatest impact on subsequently assessed variables. However, it did not pit the two attribution sources against each other. Will AS override a situational manipulation of attributions, or vice versa? Assessment of this question was made possible in Anderson (1983b) by factorially crossing AS with parallel AMs. When the two sources are in conflict, three possible effects may occur. First, they may combine additively. Second, one source may completely dominate the other. Third, one source may dominate in one combination but not in another.

The goals of this study demanded a situation that had many features of difficult social situations commonly faced by college students, but also that was not a well-practiced one for most college students. The prototypicality of the situation should prime use of the desired AS, but its unfamiliarity should allow AMs to have some impact. The telephone solicitation situation seemed to have the desired features. Because subjects had never been in the situation before, we expected AM to override AS when in direct conflict. Otherwise, we expected AS to determine subjects' reactions to the situation.

Method

College student subjects were preselected on the basis of their responses to the ASAT-I. Those scoring in the upper and lower thirds of the controllability (changeability) index of AS for interpersonal failure were asked to participate; 63 agreed to do so. Participation in the study took place several weeks to several months after the questionnaire had been returned. Subjects participated in an interpersonal persuasion study, which consisted of two sessions. In the first session, subjects were given information about telephone blood drives and then randomly assigned to one of three attribution manipulations. In the controllable conditions, the experimenter mentioned that strategies or effort might be important determinants of success. In the uncontrollable conditions he mentioned ability and personality traits as possible determinants. In the no-manipulation conditions no mention was made of possible determinants of success and failure. Thus, the design was a $2 \times 3$ (AS × AM) factorial. After the manipulation, subjects made a practice call (a rigged failure), responded to a number of expectancy measures, and received a list of names to call during the following week prior to the second session. At the second session, data on subjects' success expectancies, motivation, and success rates were obtained. Because the results for these three types of measures were essentially the same, we present here only an analysis of a composite index of the three measures. For more details, see Anderson (1983b).

Results and Discussion

The results of the composite index are presented in Figure 2. There are several points to note. First, AS had a significant effect on subjects in this complex social situation, as can be seen in the no-attribute-manipulation conditions, $t(57) = 2.11$, $p < .05$. Second, the AMs also had a significant effect, as shown by the contrast testing the difference between the AM controllable conditions (i.e., strategy or effort) and the AM uncontrollable conditions (i.e., ability or trait), $t(57) = 2.84$, $p < .01$. Third, AM completely overrode AS in this unfamiliar setting. Considering only the conditions in which attributions were manipulated, we found that the main effect of AS did not approach significance, $t(57) < 1$.

However, these results do not mean that AS effects are so weak that they are unimportant determinants of social behavior. This interpretation is strongly contradicted by the results in the no-manipulation conditions.

A final test of the relative magnitude of AS and AM effects was to compare the mean controllable–uncontrollable differences in the no-manipulation conditions (the AS effect) with the corresponding differences in the four manipulation conditions (the AM effect). This specific contrast revealed that the effect sizes were not reliably different, $t(57) < 1$. However, just as we found in Jennings's (1980) study, the mean effect of AS was just slightly larger than the AM effect, as is evident in Figure 2.

Pooled Analyses

We pooled the results of the Jennings (1980) and Anderson (1983b) studies by using the chi-square technique described by Winer (1971). This involves combining the $p$ levels associated with independent tests of the same hypothesis. The three main hypotheses concerned the effects of AS, the effects of parallel AMs, and the relative size of these two effects. We used the composite indexes from both studies (as presented in Figures 1 and 2). The first test involved a comparison of the nonmanipulated groups who differed in AS. The second test involved a comparison of manipulated attribution groups (controllable vs. uncontrollable). Both the AS and the AM effects were highly significant, $\chi^2(4) = 18, ps < .005$. The third and most interesting test involved a comparison of the AS effect (in the absence of an AM) to the AM effect. This was nonsignificant in both studies independently; however, in both cases AS had slightly larger effects. Would the combination of the studies yield a significant difference in effect size? Results indicated that the effects of AS were not reliably different in size from the AM effects, $\chi^2(4) < 1$. ¹

¹ Effect size comparisons between individual difference and situational constructs are always problematic. In the present case, the major interpretational difficulty results from the question of how extreme the preselected groups were versus how strong the situational manipulation was. Obviously, if one used very extreme groups and a weak manipulation, one could increase the effect size of the individual difference variable relative to the situational one. Just as obviously, the opposite strategy would lead to opposite relative effect sizes. In the present case, the
CONCLUSIONS

Overall, the data presented in this article support a number of suggestions made previously in the literature, but contradict extreme positions for or against the AS construct. Attributional style does seem to be a valid construct. It shows both convergent and discriminant validity when assessed at a moderate level of specificity. The different ASAT measures were moderately intercorrelated and showed theoretically predicted variations in correlations with depression, loneliness, and shyness. The various ASAT measures of AS appeared sufficiently reliable to warrant future use in research settings, although improvements would benefit researchers and practitioners alike.

The forced-choice and rating scale versions showed remarkable similarities, both in the intercorrelations of various types of AS and in AS correlations with various problems in living. However, there were dramatic differences between the forced-choice and rating scale versions in AS correlations with problems in living when assessed for success situations.

Finally, the effects of AS on complex social behaviors have been clearly demonstrated. Furthermore, the sizes of AS effects were at least as large as the effects of AM. This is not to say that manipulations cannot override AS effects; in the last study this is precisely what happened. However, in the absence of intentional situational manipulation, AS effects were clear and of a comparable magnitude. Thus, both the personality (i.e., individual differences) and the social (i.e., situational) approaches contribute significantly to our understanding of these attributional phenomena.

In sum, AS does not appear to be as general or cross-situationally consistent as originally thought. Neither is it so situationally specific as to cease being a meaningful individual difference construct. Our guess that a distinction between interpersonal and noninterpersonal situations was needed proved to be correct, but there are many other ways to categorize situations that may be more valuable (e.g., Baumeister & Tice, 1985). We feel that research on naturalistic situational taxonomies may help define the basic level of ASs needed. We also feel that more work on attributional effects in a variety of complex situations is needed, including in clinical and other applied settings. We hope others now feel such work is warranted.

References


