Manfred Schmitt

Challenges to the Construct Validity of Belief in a Just World Scales

ISSN 1430-1148
ABSTRACT
Nine critical issues regarding the construct validity of belief in a just world (BJW) questionnaires as measures of the justice motive according to Melvin Lerner are discussed. The analysis begins with a proposal of a general criterion for the construct validity of BJW scales in studies employing a mixed organismic-experimental design. Studies using this criterion for the Rubin and Peplau scale are reviewed. Based on theoretical, conceptual, methodological, and empirical analyses, it is argued that: (1) BJW is an indirect indicator of the justice motive, (2) BJW contains knowledge and justification components in addition to justice motive components, (3) the discriminant validity of BJW scales vis à vis measures for belief in control and need for control has not been demonstrated adequately, (4) the psychological meaning of belief in an unjust world is unclear, (5) the psychological meaning of low BJW is unclear, (6) implicit justice criteria contained in BJW have to differentiated, (7) implicit time perspectives contained in BJW have to be separated, (8) little is know on the changes in BJW as a consequence of observing or experiencing injustice (BJW as a dependent variable), (9) ambiguous validation criteria such as well-being and political conservatism have been used in several studies. Possible solutions to some of these issues are suggested.

ZUSAMMENFASSUNG
<table>
<thead>
<tr>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>History and Aims of this Contribution</td>
</tr>
<tr>
<td>Construct Validity</td>
</tr>
<tr>
<td>Justice Motive Theory, Individual Differences in the Justice Motive, and</td>
</tr>
<tr>
<td>a General Criterion for the Construct Validity of BJW Scales</td>
</tr>
<tr>
<td>Experimental Evidence on the Construct Validity of the Rubin &amp; Peplau</td>
</tr>
<tr>
<td>Scale</td>
</tr>
<tr>
<td>Belief in a Just World as an Indirect Indicator of the Justice Motive</td>
</tr>
<tr>
<td>Delusion, Knowledge, and Justification Components of Belief in a Just</td>
</tr>
<tr>
<td>World</td>
</tr>
<tr>
<td>Empirical Distribution of Belief in a Just World</td>
</tr>
<tr>
<td>Group Differences in Belief in a Just World</td>
</tr>
<tr>
<td>Discriminant Validity of Belief in a Just World Scales</td>
</tr>
<tr>
<td>Correlations among Measures with Discriminant Content Validity</td>
</tr>
<tr>
<td>Experimental Investigation of Discriminant Construct Validity</td>
</tr>
<tr>
<td>Belief in an Unjust World</td>
</tr>
<tr>
<td>Low Belief in a Just World</td>
</tr>
<tr>
<td>Implicit Justice Criteria in Belief in a Just World</td>
</tr>
<tr>
<td>Implicit Time Perspectives in Belief in a Just World</td>
</tr>
<tr>
<td>Stability and Occasion Specificity of Belief in a Just World</td>
</tr>
<tr>
<td>Ambiguous Validation Criteria for Belief in a Just World Scales</td>
</tr>
<tr>
<td>Well-Being and Emotions</td>
</tr>
<tr>
<td>Political Conservatism</td>
</tr>
<tr>
<td>Possible Resolutions</td>
</tr>
<tr>
<td>References</td>
</tr>
<tr>
<td>Bisher erschienene Arbeiten dieser Reihe</td>
</tr>
<tr>
<td>Andernorts publizierte Arbeiten aus dieser Arbeitsgruppe</td>
</tr>
</tbody>
</table>
History and Aims of this Contribution

When I first came across Justice Motive Theory (JMT) and related research (Lerner, 1970; Lerner & Simmons, 1966), I was initially confused, then puzzled, and eventually fascinated by the idea that certain social judgments and behaviors, such as blaming victims for their misfortune, might be motivated by the exact opposite of what these judgments and behaviors seemed to reflect on first sight: a need for justice. Much of my theoretical thinking and research interest has been influenced by and devoted to this intriguing phenomenon. Given my background in personality and individual differences and my conviction that traits add to our understanding of behavior if included properly in a theory, measured appropriately, and included correctly in research designs and data analyses, I was attracted to Rubin and Peplau's (1973, 1975) suggestion that the justice motive (JM) might differ between individuals just like other motives and needs do (power, achievement, approval, etc.).

In 1980, Leo Montada invited Claudia Dalbert and me to participate in his work on relative privilegation (Montada, Schmitt, & Dalbert, 1986). Our model for predicting reactions to underprivileged groups, such as reactions to citizens from the developing countries, was inspired by Hoffmann's (1976) concept of existential guilt, Ryan's (1971) Blaming the Victims, Schwartz's (1977) theory of prosocial behavior and responsibility denial, Shaver's (1985) model of responsibility attribution, and, last but not least, Lerner's (1980) JMT. Based on the core premise of JMT, we predicted that Belief in a Just World (BJW) as an indicator of the JM would affect the way in which people perceive, explain, evaluate, and react emotionally and in action to the fate of disadvantaged groups and suffering victims. Accordingly, we assumed that BJW would influence the way in which privileged individuals deal with their own advantages, for example, whether they justify their privileges as deserved or feel guilty about them.

The methodological paradigm we chose for our research required measuring individual differences in the JM. Following Rubin & Peplau’s (1973, 1975) reasoning, we chose BJW as the best available indicator of the JM. Dalbert (1982) translated Rubin and Peplau's (1975) Just World Scale into German for this purpose. However, the psychometric qualities of the German version of that scale were poor in two German samples with the greatest problem being factorial heterogeneity (Dalbert, 1982). In addition to these problems, we had suspected conceptual problems with some of the items confounding BJW with distributive justice criteria (mostly equity). For these reasons, we decided to develop a new self-report questionnaire (Dalbert, Montada, & Schmitt, 1987). Our BJW inventory consists of two scales. The general BJW scale (GBJW) contains items whose content is most general. Our domain specific BJW scales address justice regarding privileges which distinguish our subjects from disadvantaged comparison groups (Dalbert et al., 1987). Our GBJW scale contains six items. Item 1 is identical with Item 2 of the Rubin & Peplau (1975) scale, while Item 2 of our GBJW scale is similar to Item 11 of the Rubin & Peplau (1975) scale.

1. Basically, the world is a just place.
2. By and large, people get what they deserve.
3. Justice always prevails over injustice.
4. In the long run, people will be compensated for injustices.
5. Injustices in all areas of life (e.g., profession, family, politics) are the exception rather than the rule.
6. People try to be fair when making important decisions.

These six items are very general in three regards: (1) Unlike our domain-specific scale, the Furnham & Procter (1992) questionnaire, and the Maes (1992) scales, they do not specify domains or spheres of justice. (2) Unlike some of the Rubin and Peplau items, the items of our scale refer neither to specific rules of distributive justice (such as equity, equality, or need) nor to specific criteria.
of procedural justice (such as process control, outcome control, consistency, nonpartiality, accuracy, lack of bias, correctability, representativeness). (3) Unlike some of the Rubin and Peplau items, no agents are mentioned who violate or protect justice. This general format was chosen to make it possible for subjects to select their own domains, justice criteria, agents, and means when expressing their BJW. This is feasible because JMT does not imply what justice is and how it can be attained. More specifically, a general format does not confound belief in justice with domains, criteria, agents, and so forth. Statements in which justice is related specifically to domains, criteria, agents, means, etc., introduce additional sources of variance and may harm construct validity. For example, answers to items in which justice is linked to a certain criterion such as equality are ambiguous. Consider an item claiming that men and women have equal career chances. Agreeing to such an assertion indicates BJW only if equality of career chances is accepted as just. A second argument for the general format was that unspecific items allow individuals to associate their personal themes of justice, i.e., justice issues which are important for them. It can be assumed that the JM is most powerful in domains which people care about. A third argument for the general format was parsimony.

Several studies have been conducted to explore the psychometric quality of our GBJW scale. Its internal consistency is satisfactory in absolute terms (alpha ranging between .70 and .80 in several independent studies) and good in relation to the small number of items. Dalbert & Schneider (1995) report a peer rating validity of .41 when correlating the scale with the average scale scores of three acquaintances. This validity coefficient is not high in absolute terms but substantial given that belief in a just world cannot be linked as directly with observable behavior as other personality variables (such as sociability, dominance, anxiety, intelligence, etc.). Finally, the pattern of correlations between our GBJW scale and measures for other constructs appears to be coherent, theoretically conclusive and thus affirming the scale's construct validity (for reviews see Dalbert, 1996; Montada, in press; Reichle, Schneider, & Montada, in press).

Given the widespread scepticism of experimental social psychologists toward survey research, I decided to additionally investigate the construct validity of our scale experimentally, treating BJW as an independent organismic variable and moderator of experimental factors -- as others have done before (e.g., Miller, 1977). In one of these experiments which will be described in more detail later, subjects witnessed another person winning or losing money by chance (Schmitt et al., 1991). Afterwards, they were asked to rate the stimulus person on a number of favorable and unfavorable traits. It was predicted that subjects would upgrade the winner and downgrade the loser, and that the extent of this motivated bias would be larger for subjects with a strong BJW than for subjects with a weak BJW. The obtained pattern of results ran directly counter to these expectations: Subjects scoring high on our scale tended to upgrade the loser and downgrade the winner. Although a reasonable psychological interpretation for this result is possible (see below), it instigated, together with results from three other experimental studies (Schmitt, 1991, 1992; Schmitt & Herbst, 1993, see below), a process of reflecting about the construct validity of self-reported BJW. During this process, which is still in progress, several issues came up that seem crucial to me for understanding the psychological meaning of self-reported BJW. In this chapter, I will address these issues which have, in part, been raised by other authors before (Lerner, 1980, Chapter 10; Furnham & Proctor, 1989; Maes, 1994a, in press).

My reflections and the data that I present will challenge the construct validity of self-reported BJW. It is important to note that it is not my intention to dismiss research on individual differences in self-reported BJW as useless per se. Also, I am not suggesting that BJW scales may generally lack construct validity. Furthermore, my objections are not objections against JMT itself. On the contrary,
I start out assuming that JMT validly describes a set of psychological processes and phenomena that can be explained best by people's desire for justice. Without relying on this assumption, it would be impossible to derive empirical criteria for testing the construct validity of BJW scales. The purpose of this chapter is to share some conceptual problems and some puzzling data regarding self-reported BJW with the hope that a discussion of these issues and data will lead (1) toward a better understanding of the psychological meanings, origins, and consequences of BJW, (2) toward an improvement of instruments for measuring individual differences in the JM, and (3) toward a better knowledge of the conditions under which self-reported BJW is a valid indicator of the JM. For a variety of reasons, limited space being one of them, I will be unable to provide satisfactory solutions to all issues raised and to offer comprehensive and conclusive explanations for all data presented. Nevertheless, it is my hope that BJW theory and research will not suffer but instead profit from a discussion of controversial issues and data.

**Construct Validity**

The issues that will be addressed are issues of construct validity. There is consensus among most scholars in the social sciences that construct validity can only be defined and explored empirically on the basis of a theory (Cronbach & Meehl, 1955). Further, testing a theory and testing the construct validity of measures for elements of that theory cannot occur independently from each other. A simple but important implication of that conjunction is that at least two conclusions are logically possible whenever empirical observations contradict expectations: (1) The theory, from which the expectations were derived, is wrong. (2) The method by which the observations were made, is inappropriate. Many more interpretations are possible, of course. Let me point to one that will become relevant later in this paper. In personality research, the construct validity of measures for dispositional constructs has often been investigated within the multitrait-multimethod framework proposed by Campbell and Fiske (1959). It is obvious that this method requires the measures for theoretically similar and dissimilar constructs to be valid themselves if they are to be used as criteria for the validity of a new construct or a new method. Similarly, the predictive validity of a measurement instrument can only be judged fairly if the criterion itself is valid, a prerequisite which has not always been taken seriously (cf., Fishbein & Ajzen, 1974).

Despite this general ambiguity of empirical research regarding the validity of a theory or a methodology, the conclusions that are drawn from data depend, to a great deal, on the status quo of both. If the theory has been successfully applied many times for predicting observations implied by the theory, we will tend to attribute a false prediction with a new methodology to that methodology. Conversely, we will more likely dismiss a theory or revise it if contradicting observations were made using a well-tried method. Both conclusions may be wrong, but they are reasonable and common in the social sciences (Cattell, 1966).

What is the validity status quo of the JMT and BJW scales? JMT has been supported by a considerable number of experimental studies (for reviews see Lerner, 1980; Lerner & Miller, 1978; Lerner, Miller, & Holmes, 1976). At the same time, results from correlational studies including BJW scales have been interpreted in favor of JMT (e.g., Dalbert, 1996; Reichle, Schneider, & Montada, in press). Strictly speaking, this latter survey research may not count as evidence supporting JMT, however. This would only be possible under the assumption that BJW scales are valid. Since this assumption is at issue here, we can only rely on experimental or other general evidence when judging the validity of JMT. If we consider the available reviews of experimental research on JMT (Lerner, 1980; Lerner & Miller, 1978; Lerner, Miller, & Holmes, 1976), it seems safe to conclude that this theory is indeed valid. Therefore, our inquiry of the construct validity of self-reported BJW starts out from the assumption that the central premises of
JMT are true.

**Justice Motive Theory, Individual Differences in the Justice Motive, and a General Criterion for the Construct Validity of BJW Scales**

Given that the construct validity of any BJW scale must be judged with regard to the core premises of JMT, it may be useful to recapitulate these premises before considering any data. The core assumptions of JMT, according to the available writings (e.g., Lerner, 1970, 1975, 1977, 1980; Lerner & Miller, 1978; Lerner, Miller, & Holmes, 1976), may be summarized by the following statements, which are partially redundant:

1. Justice is a central and pervasive concern for individuals.
2. What is considered just in a particular situation depends on the individual's sense of deserving (personal contract), on social norms and conventions, on social contracts, and on the social relations among those involved in a social transaction.
3. People want to be sure that they get what they are entitled to.
4. More generally, they are committed to live in just world - a world in which all get what they deserve and deserve what they get.
5. The JM generally leads to helping innocent victims if that is possible and if helping does not interfere with the person's own sense of deserving.
6. If helping innocent victims is impossible or possible only at the cost of the person's own entitlements, the person will solve the justice conflict by distorting the situation cognitively in order to make it appear just. Distortions can take on various forms such as denying innocent suffering, blaming victims for their fate, or derogating them.
7. Such paradoxical effects serve the same purpose as helping innocent victims: They are functional for defending the person's belief in a just world.
8. Accordingly, the need for justice may delude the person's justice belief.
9. BJW can also be preserved by escaping the situation and by dissociating the subject's world from the victim's world ascertaining that the laws in the victim's world do not apply to the subject's world. Consequently, unjust events in the victim's world do not imply that the subject's world may be unjust as well.
10. People are generally not aware of their commitment to living in a just world and of the ways in which the JM affects their emotions, thoughts, and actions. They are not able to predict subject's behavior in the prototypical JW experiments and they would not come up with the JM as an explanation for the paradoxical reactions of bystanders toward innocent victims.

Each of these assumption is formulated in a general manner, i.e., it is imputed to all human beings. To begin with, it is assumed that every individual has a JM. Furthermore, it is assumed that justice for all individuals motivates prosocial behavior under a first set of conditions, leads to unconsidered behavior under a second set of conditions, and causes (seemingly) heartless behavior under a third set of conditions. From an individual differences perspective, each of these assumptions may be an overgeneralization across individuals because individuals may differ in their JM. This possibility was acknowledged first by Rubin & Peplau (1973, 1975) who developed a well-known and often used self-report questionnaire for measuring BJW.

Getting back to construct validity, how can the central postulates of JMT be transformed into validity criteria for an individual difference measure of the JM? The general answer is that such a measure should moderate every general effect of the JM. Every general effect that follows from JMT
should be stronger for subjects with a high need for justice and weaker for subjects with a low need for justice. In other words, valid measures of the JM should interact in a synergetic fashion with situational factors that generate a general JM effect.

**Experimental Evidence on the Construct Validity of the Rubin & Peplau Scale**

Several empirical studies have implemented this general methodological rationale for the Rubin & Peplau scale. An initial study was conducted by Rubin & Peplau (1973). In this study, groups of 19-year-old men listened to a radio transmission of the 1971 national draft lottery in which they and their age peers were assigned the likelihood of being drafted as soldiers with the possible consequence of being sent to Vietnam. Most subjects reacted with greater sympathy, greater liking, and less resentment toward individuals with high draft priorities. Among subjects with high BJW scores (upper third of the distribution), however, the compassionate pattern was neutralized and even reversed regarding four of seven measures for liking, with fortunate targets being rated slightly more favorably than unfortunate targets. This moderator effect corresponds to the proposed criterion for the construct validity of BJW scales.

Adopting the Lerner & Simmons (1966) paradigm, Zuckerman, Gerbasi, Kravitz, & Wheeler (1975) showed subjects a videotape of a person who appeared to be receiving electric shocks in a supposed learning experiment. Subjects with high BJW derogated the victim more than subjects with low BJW. No moderator effect of BJW on observational instructions were found with observational instructions guiding subjects' attention to (1) indicators of the target's suffering (emotional and physical arousal), (2) taking the target's perspective (imagining one's self in the situation), and (3) the target's obedience for the experimenter's instructions (Lerner, 1980, p. 148).

Miller, Smith, Ferree, & Taylor (1976) showed subjects three video tapes displaying verbal interactions between a medical doctor and a female patient-intern. The patient described her injuries which were caused by three incidents: (1) a car accident in which the woman was the driver, (2) a forcible rape, (3) and a car accident in which the victim was a pedestrian. Pretest data and manipulation control data suggested that responsibility decreased in the order of these causes. A main effect of responsibility according to JMT was expected with responsible victims being less derogated than innocent victims. More importantly in the present context, it was expected that this effect would be stronger for believers in a just world than for nonbelievers. This interaction effect was not significant. Only a main effect of BJW on derogation was found.

Miller (1977) gave subjects the opportunity to donate some of their compensation received for participating in his experiments to needy families. Threat to BJW was varied in the first experiment by describing the family either as an isolated case of neediness or as an example for many similar cases. It was expected that more help would be given to a single victim because only in this case does help seem effective for restoring justice. In the second experiment, need was displayed either as temporary or as permanent. It was expected that more help would be given to temporarily needy families than to families with enduring needs because donating money cannot relieve permanent suffering. More importantly for our present discussion, it was predicted that BJW would moderate these effects. In comparison with low believers, high believers were expected to donate more money in the Isolated and Temporary conditions and less money in the Group and Permanent conditions. Both interaction effects were confirmed.

Subjects in an experiment by Dion & Dion (1987) were shown pictures of attractive and unattractive stimulus persons. Subjects were asked to rate the targets on 17 personality trait dimensions (social desirability index) and nine life objectives (life-outcome index). A significant two-way interaction effect of BJW x attractiveness x gender of target on the social desirability index was
found. For male targets, BJW moderated the effect of attractiveness in the way that was expected: While believers in a just world attributed a more favorable personality to attractive than to unattractive targets, nonbelievers displayed no such difference. For female targets, the same effect was not found. Rather, there was even a noticeable, albeit insignificant, tendency toward the opposite effect with the difference in social desirability between attractive and unattractive female targets being larger for nonbelievers than for believers. Regarding the life-outcome index as a dependent variable, no interaction between BJW x attractiveness was found. To summarize, a moderator effect of BJW was found in one of four cases.

In two experiments, Hafer & Olson (1989) treated BJW as a moderator of the effects of personal deprivation on perceived fairness and resentment. The experiments shall be described in more detail than the previous ones because they provide a good opportunity for introducing some of the issues that will be discussed in more detail later. In the first experiment by Hafer & Olson (1989), subjects were either randomly assigned to or they could choose one of two computer tasks in which desirable bonus points could be earned (Factor Choice). The second factor was Referent Outcome. In the High-Referent-Outcome condition, it was pretended that the subject's task was more difficult than the alternative task. In the Low-Referent-Outcome condition, no such difference in task difficulty was suggested. No subject received bonus points. Fairness and Resentment were among the dependent variables. A significant interaction between BJW and Choice on both dependent variables was found. In the Choose condition, BJW correlated positively with Fairness and negatively with Resentment. In the Assigned condition, these correlations were nonsignificant. The authors interpret the latter result as evidence that the random assignment was not perceived as unfair and did not threaten BJW. It might also be argued, however, that the Assign condition created a more severe threat to the person's sense of justice because it may be harder to maintain a sense of deserving if the favorability of an outcome is a matter of randomness (Lerner, 1965; Lerner & Matthews, 1967). Furthermore, one might argue from Folger's (1986) theory, that the High-Referent-Outcome condition provided a threat to the person's need for justice while the Low-Referent-Outcome condition did not. In fact, this was the authors' intention. However, no two-way interaction of BJW x Referent Outcome on the dependent variables was found.

In the second experiment, an implicit suggestion was wiven to the subjects that they did not earn points because either they had decided not to practice the task or because the experimenter had forgotten to let them practice. Several items served as dependent variables including fairness, resentment, and attributing responsibility and blame for the outcome to one's self and to the experimenter. Regression analyses revealed a significant two-way interaction effect of BJW x Experimental Condition on Experimenter Responsibility with BJW having a stronger negative effect on Experimenter Responsibility in the Experimenter Error condition than in the Personal Decision condition. The authors interpret the effect as reflecting believers' confidence in and respect for authorities. If authorities are viewed as safeguards for justice, a strong need for justice makes it difficult to blame an authority for an unfavorable outcome. Furthermore, the effect on Experimenter Responsibility is consistent with the positive effect of BJW on Fairness (p < .12): If no one can be blamed for an unfavorable outcome, it may be difficult, if not impossible, to judge the outcome as unfair. Although this reasoning is plausible, I would like to challenge it and submit that the opposite interaction effects would have been been consistent with JMT as well. Rubin & Peplau (1975) conclude from two unpublished studies (Hanback, 1974; Izzett, 1974) that in cases of obvious unfairness, need for justice motivates blaming and punishing the victimizer rather than blaming or derogating the victim. Accordingly, a negative effect of BJW on Fairness and a positive effect of BJW on Experimenter Responsibility would have made sense in the Experimenter Error condition of the present experiment because the experimenter left no doubt that the unfavorable outcome was her...
fault.

To summarize the studies reviewed so far, the pattern of results does provide some, but not equivocal, evidence for the construct validity of the Rubin & Peplau scale. Two experiments revealed clear support for the interaction hypothesis (Miller, 1977; Rubin & Peplau, 1973). The remaining studies provided no or only partial support for the construct validity of the scale according to the moderator criterion.

It could be argued in defense of the Just World Scale that experiments can attest to its construct validity only to the extent that they are internally valid from the perspective of JMT. Internal validity might be questioned in cases where main effects of experimental factors according to JMT were not found. In the Miller et al. (1976) study, for example, a main effect was found for Responsibility, but it ran counter to what JMT would predict: low responsible victims were not derogated more than high responsible victims but less. Note, however, that this objection would throw the results from the Draft Lottery Study (Rubin & Peplau, 1973) into doubt because the prevailing reaction in this study was sympathy with the least fortunate peers.

Furthermore, the discussion of the Hafer & Olson (1989) study shows how assumption-laden experimental arrangements for testing the construct validity of BJW scales are (and perhaps have to be). To the extent that these assumptions are uncertain and cannot be tested directly, i.e., independent from the main hypotheses, empirical observations remain ambiguous. In addition, it is important to recognize that experimental situations may convey a multitude of messages, some of which may not be intended but harm the internal validity of the experiment. In the Experimenter Error condition from Hafer & Olson's (1989) second experiment, for example, "the experimenter said in an embarrassed and apologetic manner, that she had forgotten part of the procedure." (p. 814). This behavior may communicate various information to the subjects: The apology may have made the subjects aware that their outcomes were unjust. It may have suggested who was guilty of the injustice. The apology may also have operated as a full or partial restoration of justice. These and additional messages may have different effects on the person's sense of justice and evoke different strategies for defending the BJW. Without knowing how the subjects construed the situation, it is difficult to draw safe conclusions regarding the meaning of main and moderator effects of BJW scales.

**Belief in a Just World as an Indirect Indicator of the Justice Motive**

BJW is a rather indirect indicator of the JM. No item appears in available BJW questionnaires which addresses the person's desire for justice directly (Dalbert et al., 1987; Lipkus, 1991; Furnham & Procter, 1992; Maes, 1992; Rubin & Peplau, 1973, 1975; Schmitt, Maes, & Schmal, 1995). This is remarkable for two reasons: First, BJW is not the core element of JMT but only one of several consequences or signs of the JM. Second, indirect indicators are not typical for motive questionnaires. Burger & Coopers's (1979) Desirability of Control Scale, for example, contains items which address the control motive more directly (e.g., "I enjoy having control over my own destiny.").

What justification can be given for choosing BJW as an indicator of the JM? The answer which follows most directly from JMT is provided by Assumption 10 (above). According to this assumption, people are generally unaware of their need for justice. If this assumption were safe, it would not make sense to ask people for the strength of their need for justice. One of several possible empirical investigations of this conclusion would be to correlate BJW scales with more direct measures of the JM. According to my knowledge, the only available instrument which comes close to such a direct measure of the JM is a questionnaire developed in our group for measuring what we
have called Centrality of Justice (Dalbert et al., 1987). Our Justice Centrality scale consists of six items which have only one factor in common and which have a sufficient internal consistency (alpha = .73):

1. There is hardly anything which infuriates me as much as the observation of injustice.
2. I could not be friends with someone who is insensitive to justice issues.
3. I feel that injustices in our society have to be pointed out again and again.
4. I believe that the observation of injustice makes me more upset than most other people.
5. I feel guilty for a long time after I have done something unjust or not stopped an injustice.
6. People who don't care for justice make me furious.

This scale and our GBJW scale were administered together in two studies: in our first Existential Guilt Study (Montada, Schmitt, & Dalbert, 1986) and in a study by Herbst, Montada, & Schmitt (Herbst, 1992; Schmitt & Herbst, 1993). The latter research will be described in more detail later. In both studies, the items from the Justice Centrality scale and the Justice Belief scale (GBJW) loaded on separate orthogonal factors in a simultaneous principle axes analysis. Consistent with this result, a small correlation of .20 between the scales was obtained in the Herbst, Montada, & Schmitt Study.

These results leave open the question, of course, which of the two scales is a more construct valid measure of the JM. It only shows that our measure for justice centrality does not correlate with our measure for BJW. Obviously, the wording of the centrality items captures the core of JMT better than the wording of the BJW items. According to Assumption 10, however, content validity cannot substitute construct validity in the present case. According to Assumption 10, our centrality measure cannot be a valid measure of the JM. One might even consider the small correlation between both scales as supporting Assumption 10. Since the validity of our GBJW scale is uncertain, however, such a conclusion would be unsafe.

What should be done with these ambiguities? The only solution I see is to use our centrality scale or any other more direct measure of the JM together with BJW scales in the same validation studies and compare their effects. Besides providing a comparative investigation of the construct validity of both types of measures for the JM, such studies may add to our knowledge of whether and when people are aware of their need for justice. Such a study was conducted by Herbst, Montada, & Schmitt (Herbst, 1992; Schmitt & Herbst, 1993) and will be reported later.

**Delusion, Knowledge, and Justification Components of Belief in a Just World**

Challenging the construct validity of BJW as an indicator of the JM implies presuming that a person may believe in a just world for other reasons than the JM. What other origins besides the JM may exist? Let's consider the wording of typical items from BJW scales. What does it mean if a person agrees to assertions like "Men who keep in shape have little chance of suffering a heart attack" (Rubin & Peplau, 1975) or "People try to be fair when making important decisions" (Dalbert et al., 1987)? One possibility is that the person wants to believe what these statements express in order to feel safe. This interpretation is consistent with JMT. A second possibility is that the person agrees to these items on the basis of experience or knowledge. A third possibility is that the person acknowledges injustice but in public appeals to BJW assertions in a strategic manner, for example, in order to justify own advantages.

How can we empirically explore the extent to which these interpretations are appropriate? This question is perhaps posed too simply because different answers may be true for different individuals.
and for different domains. Yet it may be useful to begin with the question in its general form. As I see it, at least two methodological strategies can be employed for investigating the question empirically.

**Empirical Distribution of Belief in a Just World**

An initial and straightforward strategy is to compare subjects’ answers to items from BJW scales with "objective reality", or at least with what we think is an appropriate description thereof.

“All of us know that there are great numbers of people in our midst whose lives are filled with suffering pain, emotional anguish, deprivation of the body and spirit. Many people live under conditions of devastating poverty. Many others spend the greater part of their lives in mental hospitals with budgets barely adequate to provide minimal care, let alone the kind of treatment which could give them back their lives. And many children spend their most vulnerable and formative years in overcrowded and potentially crippling institutions.” (Lerner, 1970, p. 205)

If we agree that this statement is a fairly appropriate portrayal of the world in which we live and that most subjects of our studies have come across similar instances of inequality, undeserved suffering, and unjust treatment, what answers can we expect them to give to BJW items? According to JMT, the observation of injustice motivates people to restore justice either by actions or by distorting their observations. Both strategies help to preserve BJW. Consequently, one would expect a fair amount of subjects endorsing items which claim a just world and objecting to negatively keyed items. The reverse pattern should emerge if BJW items measured an unbiased aggregation of the subjects' knowledge.

Rubin & Peplau (1975) report an average item mean of 3.08 for college students when using a six-point rating scale ranging from 1/total rejection of the just world premise to 6/total acceptance of the just world premise. Hence, there was slight average tendency to reject the notion of a just world. Rubin & Peplau (1975) cite an unpublished study (by Merrifield & Timpe) in which an average item mean of 3.79 was found for a student sample. Both values are rather close to the midpoint of the scale (3.5), and they do not provide compelling evidence either for the delusion assumption or the knowledge assumption.

Our GBJW scale also uses six-point rating scales. If keyed in the same way as the Rubin & Peplau items, the item means for the six items of our GBJW scale amounted to the following values in a demographically heterogeneous sample of more than 2,000 subjects: 1.22, 1.65, 1.90, .95, 1.80, 1.91. In student samples, the item distributions were skewed even more toward the no-BJW end of the rating scales. For example, in a study by Mohiyeddini (1995), the following item means were obtained in a sample of 281 students: .79, 1.46, 1.24, 1.38, 1.11, 1.73. These values show that most individuals tend to reject the notion of a just world. This result is more in line with the knowledge hypothesis than with the delusion hypothesis.

**Group Differences in Belief in a Just World**

As a second strategy for empirically clarifying the nature of self-reported BJW, one could compare groups for whom different scores on BJW scales are to be expected under the delusion and knowledge assumptions. More specifically, one would expect a lower BJW under the knowledge hypothesis, but not under the motivation hypothesis, for groups whose members have experienced many or severe instances of injustice (e.g., refugees, victims of torture) in comparison to groups who have experienced fewer or less severe instances of injustice. Supporting the knowledge hypothesis, several studies have found that objectively disadvantaged groups have lower BJW compared to more privileged groups. For example, Smith & Green (1984) found lower BJW for blacks than for whites living in the same society. Using the Rubin & Peplau scale, Furnham (1993) compared BJW
in twelve countries. For the negatively keyed items from the scale, he found a significant correlation between BJW and gross domestic product with BJW being lower in poorer countries. There are other studies, however, which have not found a correlation between objective disadvantages and BJW. For example, Rubin & Peplau (1975) report a correlation of .03 between social class and BJW. Finally, there is at least one data set which supports the delusion hypothesis. Lerner (1980, p. 165-171) describes results from an unpublished study by Lerner & Elkinton suggesting a negative correlation between social class and BJW.

The problem with such group comparisons is that group differences on the crucial independent variable (here: amount of disadvantages or number and severity of unjust treatments) are likely to be confounded with differences in other variables that might also affect BJW. If these variables are unknown, unmeasured, or remain uncontrolled for statistically, they harm the internal validity of the design. It is therefore important to find groups who are as similar as possible in as many irrelevant variables as possible. Fortunately (from the perspective of internal validity), recent historical events have provided us with such groups. After the German reunification, East and West Germans have been compared in several social science research projects aimed at finding out the effects of living in different political and economic environments for almost two generations. It is pretty safe to attribute any mean differences between East and West Germans to this treatment (and its aftermaths) because the treatment comes very close to the ideal of a random experiment. In fact, the separation of East and West Germany after World War II may be among the best ecological experiments (Bronfenbrenner, 1979) that ever took place. At least six studies have included indicators for BJW for comparing East and West Germans after the reunification.

Using our GBJW scale, Limbach (1992) investigated a demographically heterogeneous sample of 106 jobless adults. Sixty-seven subjects were West German, 37 East German. Data were collected from January through March 1992 (Germany was reunified on October 3, 1990). The author reports a lower BJW for East Germans (mean item score = 1.49) than for West Germans (mean item score = 1.90). The difference was marginally significant (p = .06).

Dalbert (1993a) administered our GBJW scale to a sample of 91 West German and 88 East German students in March 1991. She expected a higher BJW for East Germans and offers two reasons for this expectation: (1) The German Democratic Republic was just by definition. (2) In the Federal Republic, issues of social injustice were discussed openly. Mean item scores differed significantly in the predicted direction (East Germany: 1.80; West Germany: .94).

Schmitt & Janetzko (1993) administered our GBJW scale to a demographically heterogeneous sample of 122 West German and 122 East German adults. Data were collected in March 1991. Based on the knowledge assumption, we suspected two counteracting mechanisms that may cause East-West differences in BJW: (1) East Germans may have a higher BJW than West Germans due to the revolution and the reunification as two recent and salient experiences of justice. This assumption implies that knowledge based BJW is short-lived, i.e., it can change with changing experiences. (2) East Germans may have a lower BJW than West Germans due to living in a basically unjust society for 40 years. This assumption implies that BJW is a trait which develops in a long process of experiences and changes slowly. Given evidence on the stability of BJW (see below) and starting out from the knowledge hypothesis, we predicted that East Germans would have a lower BJW than West Germans. The difference between mean item scores accorded to this speculation (East: 4.35; West: 4.10), but it was not significant.

Braun (1993) analyzed data from a microcensus (ALLBUS) conducted in Germany between May and July 1991. The survey contained one interview question which can be considered an indicator for BJW: "By and large, social differences are just in our country." Subjects could answer
this assertion on a four-point rating scale ranging from 1/fully agree to 4/do not agree at all. The sample was representative and consisted of 1514 West Germans and 1544 East Germans. The mean scores for East and West Germans were 3.2 and 2.6, respectively. This difference is significant (p < .01). Accordingly, West Germans had a higher BJW than East Germans.

Wegener & Liebig (1994) report data from the German part of the International Social Justice Project. Their sample was representative and included 1839 West Germans and 1017 East Germans. Data were collected in April and May 1991. The survey contained two items for measuring Distributional Fatalism. The items may also be considered indicators for BJW: (1) "Prejudice and maltreatment of certain groups in our society are the cause for poverty in unified Germany." (2) "Lack of equality of chances is the cause for poverty in unified Germany." Items were answered on five-point frequency scales. The correlation between group membership and the sum of both items was significant and amounted to .13 with East Germans having a lower BJW than West Germans.

Finally, in our project on Justice in Unified Germany, we administered our GBJW scale to more than 2500 East and West Germans (Schmitt, Maes, & Schmal, 1997a). The sample was not representative but demographically homogeneous. Data were collected in January 1996. East Germans displayed a significantly (p < .01) lower BJW than West Germans. Mean item scores were 1.40 and 1.57, respectively.

Taken together, the data show that East Germans had a lower BJW shortly after the reunification than West Germans. And this difference endured until six years after the reunification (Schmitt et al., 1997a). Only Dalbert (1993a) found the opposite pattern. Although the data accord to her reasoning, the reasoning itself may be questioned. Dalbert (1993a) argues that East Germans have a higher BJW because they were told to live in a much juster society than West Germans. If they believed this, why would they have reasons for wanting to overthrow their regime and strive so urgently for the reunification?

Although the empirical evidence is rather clear, it is not unambiguous regarding the delusion versus knowledge hypothesis. The pattern of results is more favorable toward the knowledge hypothesis than toward the wishful thinking hypothesis, but only under the assumption that individual differences in BJW are fairly stable across time. As will be shown later, there is empirical evidence for the stability assumption, but from a sample which did not undergo as severe changes in living conditions as East Germans did during the revolution and the reunification. If we assume that BJW can change either due to changes in the knowledge basis, or due to changes in the need for justice, or in the amount of threat to the BJW, the data remain ambiguous regarding the present issue. One could argue, for example, that East Germans discovered after the reunification that the new system also creates many injustices. Even if the number and severity of injustices in the new system were less than those in the old system, the discovery may have been in great contrast to the hopes and expectations East Germans had developed before the revolution. This contrast may have depressed their BJW below the level of West Germans. While this reasoning would still support the knowledge hypothesis, it is not the only explanation for the data. Still assuming that BJW may change considerably if external conditions change, one might argue in favor of the delusion assumption that East Germans felt so safe about justice shortly after the reunification that they had no more need for defensive illusions. Empirically, these competing speculations could only be tested longitudinally by including data on the BJW of East Germans before the unification. Unfortunately, such data are not available. Prospective studies with individuals, groups, or nations for which dramatic changes in justice can be expected (Korea, Middle East, South Africa, Northern Ireland?) might add considerably to our knowledge about the dynamics of the JM and BJW.
Discriminant Validity of Belief in a Just World Scales

Discriminant validity is a crucial criterion for the construct validity of psychological measures (Campbell & Fiske, 1959). The issue of discriminant validity of BJW scales needs to be addressed in the present context for several related reasons. First, alternative interpretations have been advanced for some of the phenomena that JMT seeks to explain (e.g., Miller et al., 1976; Sauer, 1984). Perhaps most importantly, some authors (e.g., Shaver, 1985; Walster, 1966) have argued that blaming the victim may be caused by a need for control. In order to test both explanations against each other, need for justice and need for control have to be conceptually and empirically separated (cf. Maes, 1994a). Regarding discriminant construct validity of self-reported BJW, it has to be explored whether, and to what extent, BJW reflects a need for control in addition to reflecting a need for justice.

The second reason is simply that correlations between measures for BJW and measures for Internal Locus of Control have been repeatedly reported in the literature (e.g., Clayton, 1992; Collins, 1974; Hafer & Olson, 1989; Lerner, 1978; Maes, 1994a; Rubin & Peplau, 1973, 1975; Zuckermann & Gerbasi, 1977). Whether this correlation is a problem of insufficient discriminant validity depends on the true correlation between belief in justice/need for justice and belief in control/need for control. If the correlation between the measures would match the true correlation, there would be no validity problem. Rather, depending on how close the true correlation is, there might be a problem with scientific parsimony. The correlations between the BJW scales and Internal Locus of Control scales reported in the literature amount to an average value of about .40. Given attenuation due to unreliability and the fact that belief in control may not be the best indicator for need for control (see the indirect measurement issue addressed earlier for BJW), the true correlation between the justice motive and the control motive may be considerably higher.

A substantial true correlation between need for justice and need for control is not surprising if we consider the possibility that justice and control are confounded due to bidirectional causal influences. Under certain conditions, justice may imply control and control may imply justice. A just world is a predictable world, and to the extent that the rules are known, individuals can control their outcomes by choosing appropriate actions. Conversely, those who have power are capable of setting up a just world. This intrinsic relation between justice and control can be transferred to the need for control and the need for justice. A need for justice may reflect a more fundamental need for control. In this case, justice may not be desirable per se but as a means for control. Conversely, a need for control may reflect a more fundamental need for justice. In this case, control may not be desirable per se but as a means to obtain justice. An additional reason for a true correlation between need for justice and need for control may be that both serve the same underlying basic need for security and well-being.

The correlations between BJW scales and Locus of Control scales do not necessarily reflect a true correlation. Another explanation is shared method variance. Consider assertions like "heart attacks can be prevented by keeping in shape" or "success can be achieved by hard work." Both assertions seem appropriate indicators for both, belief in justice and belief in control. In fact, items like these can be found in both, BJW scales and Locus of Control scales. How can we empirically determine which of these alternative explanations accounts for the correlation at issue? At least two general strategies may be employed.

1. The first strategy is to devise measures for both constructs with the highest possible discriminant content or face validity (Maes, 1992). This means using only the most prototypical items possible and avoiding items that may also measure the reference construct. The correlations among such pure and prototypical indicators provide less biased estimates of the true correlations between the constructs (Borkenau, 1986).
2. The second strategy is to design experimental studies for investigating the discriminant construct validity of measures for both constructs according to the general rationale outlined earlier. This strategy requires experimental situations which stimulate only one of both motives but not both simultaneously. Under this condition, only the measure for the motive that was aroused should act as a moderator while the measure for the other motive must not.

**Correlations among Measures with Discriminant Content Validity**

Both strategies were implemented in a study by Herbst, Montada, & Schmitt (Herbst, 1992; Schmitt & Herbst, 1993), in which our Justice Belief (GBJW) scale and our Justice Centrality scale (see above) were used. The items from both questionnaires do not address issues of control, at least not explicitly. For this reason, they fulfill the prototypicality criterion better than the Rubin & Peplau scale. In addition to these justice scales, two control scales were developed, a General Belief in Internal or Personal Control scale and a Centrality of Control scale. Both scales were designed as similar as possible in level and scope to the justice scales. Also, the items should be prototypical and not overlap with the justice domain.

**Items of the Control Centrality scale:**

1. I don't like others making decisions concerning myself.
2. I prefer to know what will happen to me.
3. I rather rely on myself than on others.
4. It is important for me to take matters into my own hands.
5. I hate leaving something to chance.
6. I prefer doing things myself.

**Items of the Control Belief scale:**

1. I believe that I can determine most things in my life.
2. I can achieve much if I want to.
3. Chance does not play an important role in my life.
4. If it really matters, I can accomplish my goals.
5. Other people have less influence on my life than me.
6. I can accomplish most of my ideas of how I want to live.

Note that some items of our Centrality scale are similar with some items of the Burger & Cooper (1979) Desirability of Control Scale. This supports the idea that our centrality scales (both justice and control) tap the need aspect better than the belief scales, at least on the level of literal item wording. Note that unlike JMT, Burger & Cooper (1979) make no unawareness assumption regarding the control motive. Although this issue is not addressed explicitly by these authors or by Burger (1992), it can be inferred from their scale.

The Control Centrality scale and the Control Belief scale were submitted separately to principle axes analyses (Herbst, 1992). The results show that both sets of items have only one factor in common. Internal consistency reliability estimates alpha for the Centrality scale and the Belief scale were .76 and .67, respectively. While the coefficient for the Centrality scale is acceptable, the coefficient for the Belief scale is insufficient. Therefore, correlations between the scales should be corrected for attenuation (alphas for the GBJW and Centrality of Justice scale were .80 and .75, respectively, in the present study).

All four scales contain items with high discriminant content validity on the level of explicit item
wording. No justice item refers to control issues, no control item refers to justice, no belief item refers to centrality issues, and no centrality item refers to belief issues. It seems as safe as possible, therefore, that the correlations between the four scales are not biased by shared method variance. The manifest correlations which were obtained from a demographically heterogenous sample of 182 adults are given in Table 1 (first value in a cell). Table 1 also contains the corrected coefficients (second value in a cell) and the probability of type I errors for the manifest correlations (third value in a cell).

### TABLE 1

<table>
<thead>
<tr>
<th></th>
<th>Justice Belief</th>
<th>Justice Centrality</th>
<th>Control Belief</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justice Centrality</td>
<td>.20</td>
<td>.26</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Control Belief</td>
<td>.18</td>
<td>.25</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Control Centrality</td>
<td>.12</td>
<td>.15</td>
<td>&gt;.05</td>
</tr>
</tbody>
</table>

The correlation between Justice Belief and Justice Centrality was discussed earlier already (BJW as an indirect indicator of the JM). The correlation between Control Belief and Control Centrality is not crucial in the present context except for to show that wanting to have control and believing to have control are more closely related than wanting justice and believing in justice. The correlation between Justice Centrality and Control Centrality is important in the present context. This correlation is substantial on the true score level (.57). If both scales were valid need measures, the correlation would show that need for justice and need for control are truly correlated. This would be in line with the foregoing reasoning (a) that justice can be a means for control, (b) that control can be a means for justice, and (c) that justice and control can both be means for security and well-being.

The correlation between Justice Belief and Control Belief is also important in the present context. This correlation is considerably lower, even when corrected for attenuation, than the average correlation of .40 reported in the literature for BJW scales and Internal Locus of Control scales. This difference suggests that the latter correlation is partly due to shared method variance (same or similar items in both measures) while the correlation obtained here is not. Furthermore, the correlation is considerably lower than the correlation between the two Centrality scales. This is not in accordance with assuming an intrinsic connection between justice and control. In other words, the perception of personal control does not imply the perception of justice and vice versa. This may be due, at least partly, to the fact that level and scope are less similar for the belief scales than for the centrality scales. While the items of both Centrality scales and of the Control Belief scale refer to the person, this is not true for the Justice Belief scale. This scale is more general in the sense that the subject is not mentioned explicitly in the items and does not appear as a salient part of the scenario that the items portray. Although this difference in level and scope is a necessary consequence of how the two constructs are conceptualized, it adds to the ambiguity of empirical findings regarding the discriminant validity issue.

Nevertheless, the coefficients in Table 1 fulfill the classic criterion of discriminant validity with regard to our GBJW scale. The correlations of this scale with the three remaining scales is small enough to maintain the assumption that BJW reflects something unique - whatever that may be.
**Experimental Investigation of Discriminant Construct Validity**

The second strategy mentioned above for exploring discriminant validity was also employed in the Herbst, Montada, & Schmitt study. The study was a vignette study containing four scenarios in which an unfavorable outcome occurred to the protagonist:

1. A bystander (protagonist) was hit by a rock during a rightist riot.
2. A car accident occurred to someone (protagonist) due to faulty repairs performed by a car repair shop.
3. A car accident occurred to a person (protagonist) while trying to avoid running over a child.
4. A car accident occurred to someone (protagonist) who ignored another's right of way.

Type of Scenario was varied within subjects, i.e., all scenarios were presented to all subjects. Two factors were varied between subjects and were fully crossed: (1) Need for Justice (evoked, fulfilled) and (2) Need for Control (evoked, not evoked).

Need for Justice was evoked by leaving the victim without compensation and letting the victimizer get away without punishment. This should provide a threat to the subject's sense of justice. In the Need-for-Justice-Fulfilled condition, the victim received full compensation and the victimizer was punished. This should affirm the subject's sense of justice. To avoid confounding justice and control, compensation depended on factors that were out of the victim's control. In the riot scenario, for example, compensation was possible because another bystander photographed the event allowing the victimizer to be later identified.

<table>
<thead>
<tr>
<th>Need for control</th>
<th>evoked</th>
<th>kept silent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need for justice</td>
<td>victim similar to subject receives no</td>
<td>victim dissimilar to subject</td>
</tr>
<tr>
<td></td>
<td>compensation</td>
<td>receives no compensation</td>
</tr>
<tr>
<td>fulfilld</td>
<td>victim similar to subject receives</td>
<td>victim dissimilar to subject</td>
</tr>
<tr>
<td></td>
<td>compensation by mere chance</td>
<td>receives compensation by mere</td>
</tr>
<tr>
<td></td>
<td>chance</td>
<td>chance</td>
</tr>
</tbody>
</table>

Figure 1. Between Subjects Part of the Experimental Design of the Herbst, Montada, & Schmitt Vignette Study

Adopting Shaver's (1985) concepts of personal and situational relevance, Need for Control was varied by describing the protagonist as similar (Need for Control evoked) or dissimilar (Need for Control not evoked) to the subject. Objectively, control was impossible in all cases. The control motive was kept silent by portraying the victim as dissimilar to the subject suggesting that the same misfortune could hardly happen to the subject. Need for control was evoked by letting the victim appear similar to the subject. In the first scenario, dissimilarity was operationalized by ethnicity. The victim was an African foreigner at whom the rock might have been thrown by one of the rightist rioters on purpose. Although this was not stated explicitly, it was a likely inference from the complete story. In the Need-for-Control-Evoked condition, the victim was German. In the second scenario, the dissimilar protagonist was a physically handicapped person whose car had complicated special mechanics suggesting that the car repair shop would not make a similar mistake when working on a regular car. In the third scenario, dissimilarity was suggested by age. In the Need-for-Control-Not-Evoked condition, the driver was a very old man suggesting that he may have wrecked
his car due to slow reactions. In the fourth scenario, the dissimilar protagonist was from Great Britain suggesting that the fault may be due to unfamiliarity with continental traffic rules. Figure 1 summarizes the between subjects part of the design and the operationalizations of the two need factors.

Several dependent variables and control variables were assessed. In the present context, only the indicators for blaming the victim are of interest. Six-point rating scales were used for measuring (1) the amount of the victim’s causal contribution to the damage, (2) the extent of the victim’s responsibility for the damage, and (2) the extent of the victim’s moral blameworthiness (guilt).

What experimental (situational) effects on these dependent variables can be expected from JMT and from Control Motive Theory? The former would predict main effects of the Need for Justice Factor on all dependent variables because the subject cannot, in his role, compensate the victim. In order to make the complete event appear just, they have to justify it. This can be achieved by attributing a causal contribution, responsibility, or guilt to the victim. Control Motive Theory would predict main effects of the Need for Control Factor on the same dependent variables because guilt conceptually implies responsibility and responsibility conceptually implies control (Montada, 1992; Reichle, 1994; Shaver, 1985).

What effects can be expected from the organismic factors, i.e., from dispositional need for justice and from dispositional need for control? According to the general validation criterion exposed earlier, the main effects of the experimental factors (situational need) should be moderated by the corresponding dispositional need factors. More specifically, the main effects of the Need for Control Factor should be stronger for subjects with a high dispositional need for control than for subjects with a low dispositional need for control. Correspondingly, the main effects of the Need for Justice Factor should be stronger for subjects with a high dispositional need for justice than for subjects with a low dispositional need for justice. As a consequence of such ordinal moderator or interactions effects, main effects of dispositional need for justice and dispositional need for control can also be expected.

Discriminant construct validity of the measures for both needs requires that the moderator effects are specific or unique: Dispositional need for justice should only moderate the effects of situational need for justice but not the effects of situational need for control. Correspondingly, dispositional need for control should only moderate the effects of situational need for control but not not the effects of situational need for justice.

Using multiple regression analyses with dummy-coded experimental factors and product terms for moderator effects (cf., Aiken & West, 1991), these expectations were tested separately for each scenario. The dependent variables ( attribution of cause, responsibility, and guilt) were summed because the correlations between them were high within situations. Correspondingly, when principle axes analyses with all dependent variables were performed, the three attributions always loaded highly on the same common factor. Note, however, that the attributions did not generalize across situations. Therefore, separate regression analyses had to be performed for each situation. The dependent variable may be considered the best available indicator for blaming the victim. All significant effects are summarized in Table 2.

Regarding the effects of situational need for justice and control, JMT was supported by the results for two of the four scenarios (2, 3). If justice was reestablished, the victim was blamed less than when the victim was not compensated. No defensive attributions according to Control Motive Theory were found. Two main effects of situational need for control were significant, but they were in the wrong direction. In other words, victims similar to the subject were blamed less than nonsimilar victims. According to Shaver (1985), this may indicate a strong identification with the victim and an
attempt to avoid blame instead of harm.

In the first scenario, main effects were found for the two belief scales. Subjects with strong beliefs in justice tended to blame the victim more than subjects with low scores on our GBJW scale. Accordingly, subjects with a strong belief in personal/internal control tended to blame the victim more than subjects with low scores on the Control Belief scale. In Scenario 3, a negative main effect of Justice Centrality on blaming the victim was found. This effect contradicts theoretical expectations.

TABLE 2
Significant Effects (p < .05) of Situational and Dispositional Needs for Justice and Control on Blaming the Victim in the Herbst, Montada, & Schmitt Vignette Study

<table>
<thead>
<tr>
<th>Effect</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main and Interaction Effects of Experimental Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Justice</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Control</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Justice x Need for Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main Effects of Organismic Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice Belief</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Justice Centrality</td>
<td></td>
<td>–</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control Belief</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Control Centrality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderator Effects of Organismic Factors on Experimental Effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Justice x Justice Belief</td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Need for Control x Control Belief</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Control x Control Centrality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Justice x Control Belief</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Control x Control Centrality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Control x Justice Belief</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Need for Control x Justice Centrality</td>
<td></td>
<td></td>
<td></td>
<td>–</td>
</tr>
</tbody>
</table>

Note. A "+" for main effects indicates a positive effect of situational or dispositional need/belief on blaming. A "−" for main effects indicates a negative effect of situational or dispositional need/belief on blaming. A "+" for moderator effects means that dispositional need/belief amplified a positive situation effect. A "−" for moderator effects means that dispositional need/belief amplified a negative situation effect.

Considering the moderator effects of interest regarding discriminant construct validity, only 1 in 16 effects corresponded to theoretical expectations. In Scenario 3, Justice Centrality amplified the effect of situational need for justice on blaming the victim. A second moderator effect appeared in Scenario 1. Belief in justice amplified the negative effect of situational need for control on blaming. This effect contradicts expectations.

Taken together, the pattern of results is inconsistent, nonconclusive, and does not assert discriminant construct validity for any of the scales used. One might object that vignette studies of this
type are inappropriate for investigating JMT and Control Motive Theory or that the design and the procedure of the present study were inappropriate for testing discriminant construct validity. Regarding the control part, this objection seems justified because the main effects of situational need for control contradicted theoretical expectations. Regarding the justice part, however, such objections seem less convincing since, at least for two scenarios, the general blaming effect was found. At least these two scenarios seem internally valid from the perspective of JMT. Note that in one of these two situations, one of the general blaming effects was amplified by Justice Centrality. If we had realized only this scenario, it would be tempting to conclude that the study adds weight to the discriminant construct validity, yet of the Justice Centrality scale, but not of the BJW scale.

Belief in an Unjust World

Several authors have reported orthogonal or fairly independent factors for positively keyed BJW items and for negatively keyed items (e.g., Dalbert, 1982; Maes, 1992; Schmitt et al., 1997a). Accordingly, lower internal consistency coefficients were found for the Rubin & Peplau scale, which contains positively and negatively keyed items, than for subscales containing only positively or only negatively keyed items (Furnham & Gunter, 1984; Furnham, 1993; Heaven & Connors, 1988). The most popular interpretation for this result is that BJW and belief in an unjust world (BUW) are separate dispositions. Accepting this interpretation requires one to consider and to rule out alternative interpretations.

Separate factors for positively and negatively keyed items that were designed for measuring the same latent variable are a common phenomenon. This result has been found in the domain of social attitudes (e.g., Dalbert, Montada, Schmitt, & Schneider, 1984), in the domain of attitudes toward political ideologies (e.g., Weisberg, 1980), in the domain of moods (Emmons & Diener, 1985), in the domain of personality (e.g., dominance vs. submissiveness; Buss & Craik, 1981), in the domain of motives (e.g., the approval motive; Amelang & Bartussek, 1970), in the domain of self-esteem (Schmitt, Maes, & Schmal, 1997b), and for locus of control with internal and external locus of control appearing as separate factors in virtually every study (Skinner, 1995). The phenomenon at issue has intrigued many scholars, simply because many positively keyed items are so obvious contradictions of negatively keyed items that one would not expect anyone to agree (or disagree) to both types of items at the same time. Several explanations have been suggested for this seemingly logical inconsistency: (1) careless responding (Schmitt & Stults, 1985), (2) response sets such as acquiescence (Vickers & Hervig, 1987), (3) nonlinear item characteristics (van Schuur & Kiers, 1994), (4) asymmetric response scales (Meddis, 1972), (5) and cognitive processes (Carver & Scheier, 1990).

(1) The first interpretation assumes that some subjects from the sample do not detect that some items are negatively keyed and answer them as if they were positively keyed. This may happen due to carelessness, to complicated wording, to double negations in the item, or to double negations that result when the content of the item is combined with the meaning of the response scale. Schmitt & Stults (1985) have shown that 10% of subjects making such mistakes suffice to generate orthogonal factors for positively and negatively keyed items. It is difficult to estimate the likelihood of such errors from the wording of the Rubin & Peplau items. Dalbert (1982) investigated the assumption by comparing factor solutions for subsamples with different education, speculating that poor verbal comprehension may lead to more errors for subjects with little education. However, the data did not support this idea.

(2) The acquiescence interpretation assumes that some subjects agree (or disagree) to items independent of content, assuming that scientists would not pose wrong assertions. According to the Schmitt & Stults (1985) simulation study, a small proportion of subjects displaying this behavior will create orthogonal factors. There is a second version of the response set explanation which needs to be considered carefully. Some BJW items sound like sayings or even are sayings (e.g., "Crime doesn't pay."). Sayings have an intrinsic logic and persuasive power. Some individuals may not reflect about the content of such saying like assertions and tend to agree
without relying on their own experiences or desires. Since for most sayings, a contrary saying also exists, it would be easy to have those individuals agree to contradicting assertions.

(3) The third explanation applies to items whose item characteristics accord to the unfolding model (Coombs, 1964). Such item characteristics often result from combining yes/no or agree/disagree response formats with quantitative statements or frequency statements. Consider Item 13 from the Rubin & Peplau (1975) scale: "Good deeds often go unnoticed and unrewarded." Agreeing to this item is unambiguous, disagreeing not. A subject may disagree either because he believes that good deeds always go unnoticed or because he believes that good deeds never go unnoticed. This double meaning of disagreeing responses leads to a nonlinear relation between the item and the latent variable. Other items, e.g., those without quantitative concepts or frequency concepts, may have linear item characteristics. To the extent that negatively and positively keyed items differ in the proportion of items with nonlinear characteristics, orthogonal factors will appear (van Schuur & Kiers, 1994).

(4) Asymmetric response scales for negatively and positively keyed items is not a problem with available BJW items. But a statistically related problem, systematic differences between positive and negative items in item difficulty may be an explanation for the emergence of independent factors. The size of the maximum correlation between two variables depends on the similarity in distributions (cf. Carroll, 1961). Only variables with exactly the same shape of their distributions can have correlations of 1. A lower correlation will result for variables with unequal distributions even if the rank order of individuals is exactly the same for both variables. If positive and negative items differ in difficulty, they necessarily differ in distribution shapes as well. This may lead to the well known phenomenon of difficulty factors. Dalbert's (1982) studies using her German translation of the Rubin & Peplau (1975) scale and recent data from our project on Justice in Unified Germany (Schmitt et al., 1997a) show systematic differences in the difficulty of BJW and BUW items, the former being considerably more difficult than the latter. This may have contributed to the emergence of independent factors.

(5) Some authors have suggested that separate factors for positively and negatively keyed items may reflect cognitive processes and memory functioning. When faced with judging the truth of an assertion, individuals tend to begin screening their memory for confirmatory evidence (Carver & Scheier, 1990). Memory search is broken off if sufficient data were retrieved for making a judgment. Such nonrepresentative data retrieval leads to biased consents. The same bias can be predicted from semantic network theory. According to network models, semantic elements are connected in memory more closely with episodes that are evaluatively consistent than with inconsistent episodes (Baddeley, 1990). Both mechanisms may explain the general acquiescence phenomenon, the survival of contradictory sayings, and the emergence of separate factors for positive and negative items.

Regarding BJW scales, these explanations have to be considered carefully before concluding from factor analytic results the existence of a BUW independent of a BJW. It is difficult to test these hypotheses and rule them out as alternative explanations of the data, especially since they may contribute independently. What needs to be done, nevertheless, is to demonstrate, based on psychological reasoning and empirical data, that BJW and BUW are differentially related to third variables. In other words, the discriminant validity of both factors has to be demonstrated vis à vis other variables on the basis of a convincing theory and according to the general methodological rationale that was suggested in the last section for separating the JM from the control motive.

**Low Belief in a Just World**

Assuming that BJW and BUW are independent creates an intriguing conceptual and psychological problem: What do low scores on BJW scales and on BUW scales express psychologically? In regard to our GB JW scale which does not contain negatively keyed items, this question came up a while ago when an initial experiment was conducted to explore the construct validity of our GB JW scale experimentally (Schmitt et al., 1991). The results of this experiment initiated the present reflections on the construct validity of BJW scales and advanced the assumption that low scores on our GB JW scale reflect BUW. The design of the study is depicted in Figure 2.

One hundred and forty-five students were recruited as subjects for the experiment and assigned randomly to one of four treatments. Subjects were told that they would participate in a study on the facial expression of emotions. They would observe via live video transmission another subject taking part in a gamble. They should watch carefully how the person reacted emotionally. They would later
be asked to give their impression of the person and his emotions. The stimulus person was always the same male student. Because the outcome of the gamble had to be faked, subjects did not observe a real situation but a video film. This was disguised by having the experimenter communicate via phone with the experimenter of the gamble, pretending that they talked about the timing of the gamble, the transmission and so forth. No subject suspected seeing a film.

Subjects participated in groups of five. After arriving, they were given a questionnaire containing our GBJW scale and a German version of the Crowne-Marlowe (1960) Social Desirability Scale (Lück & Timaeus, 1969). The GBJW items were mixed with the Social Desirability items to keep them as nonsalient as possible. In order to further reduce the salience of the GBJW items, a second questionnaire was handed out that dealt with human emotions in great detail. Emotions were chosen for distracting subjects’ attention from the GBJW items and for increasing the credibility of the cover story.

<table>
<thead>
<tr>
<th>Value of fate</th>
<th>Direction of Fate</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>target wins 2 German Marks in a gamble</td>
</tr>
<tr>
<td></td>
<td>target loses 2 German Marks in a gamble</td>
</tr>
<tr>
<td>high</td>
<td>target wins 20 German Marks in a gamble</td>
</tr>
<tr>
<td></td>
<td>target loses 20 German Marks in a gamble</td>
</tr>
</tbody>
</table>

Figure 2. Experimental Design of the Schmitt et al. (1991) Study

In the two Bad Luck or Losing conditions, the target was told by the gamble experimenter on arrival in the gambling room that he had been assigned by lot to the losing condition and that he could now only lose money. The amount of money he lost would depend on how well he betted in a roulette game. The gamble experimenter then asked the target to put 20 German Marks of his own on the table (20 German Marks were the equivalent of about 13 US Dollars at the time of the experiment). The target was told that he had to bet on red or black in ten roulette trials. Each wrong bet would cost him 2 German Marks. If he betted correctly, he could keep his money. In the Low Value condition, the person betted correctly in 9 out of the 10 trials and lost only 2 German Marks. In the High Value condition, the person always betted correctly and won 20 German Marks. In the two Good Luck or Winning conditions, the target was told that he had been assigned randomly to the winning condition and that he could win money depending on how well he betted in a roulette gamble. The experimenter then put 20 German Marks on the table. In the Low Value condition, the person betted correctly in 1 out of 10 trials, thus winning only 2 German Marks. In the High Value condition, the person always betted correctly and won 20 German Marks.

After subjects had watched the video film, they were given a questionnaire containing items for rating the stimulus person's emotions and personality. The emotion items served to maintain the cover story. Personality was rated on a list of 15 bipolar six-point scales which were adopted partly from the Lerner & Simmonds (1966) study. A principle axes analysis of the 15 personality items revealed two common factors. The first factor loaded substantially (> .50) seven socially desirable traits: interesting-boring, bright-dull, lively-lame, likable-unpleasant (highest loading: .77), attractive-unattractive, warm-unfriendly, open-closed. The second factor was emotional stability (e.g., calm-nervous). The items of the first factor were chosen for measuring derogation as the dependent
variable.

Based on JMT, it was expected that the target would be evaluated less favorably in the Losing conditions than in the Winning condition and that this main effect would be stronger if the person lost/won a lot rather than if the person lost/won a little, respectively. Finally and more importantly in the present context, it was expected that these experimental effects would depend in strength on BJW. Subjects with a high BJW were expected to upgrade the winner more and to downgrade the loser more than subjects scoring low on our GBJW scale. The only significant effect that emerged from moderated regression analysis was this two-way interaction ($F_{1,137} = 4.58; p < .05$), but as can be seen from the fitted regression lines in Figure 3, the effect ran exactly counter to predictions.

![Figure 3. Moderating Effect of BJW in the Schmitt et al. (1991) Study](image)

Contrary to expectations, subjects who scored high on our GBJW scale tended to upgrade the loser and to downgrade the winner, while subjects with low BJW scores tended to upgrade the winner and downgrade the loser. Schmitt et al. (1991) offered the following interpretation: Subjects with a high BJW construe justice by assuming that losers in a gamble were more fortunate in real life due to an attractive personality which provides them with all kinds of desirable social outcomes. This construal corresponds to a popular German saying: "Pech im Spiel, Glück in der Liebe" (Bad luck with gambling, good luck in love). Accordingly, high BJW subjects assume that those who were lucky in the gamble were less fortunate in real life. Subjects with a low BJW construe their observation as proof for their assumption that the world is an unjust place and that there are losers and winners.

Note that this interpretation does not start out from a deserving notion of justice but from a nonconditional equality principle. According to the Schmitt et al. (1991) reasoning, an attractive personality serves as a compensation for bad luck in a gamble, while winning in a gamble is a compensation for a less attractive personality. Accordingly, BJW means to believe in such an equality or compensation principle. Whether this belief is based on experience (knowledge) or wishful thinking is irrelevant for the explanation. Most importantly for the present discussion is that the interpretation implies that low scores on our GBJW scale reflect BUW. Only BUW, but not a low BJW or an indifferent belief can explain why an unfavorable personality is attributed to losers.
and a favorable personality to winners.

**Implicit Justice Criteria in Belief in a Just World**

The interpretation offered by Schmitt et al. (1991) for the results of their experiment is feasible under the assumption that the subjects' sense of justice did not accord to concepts of deserving but to concepts of equality and compensation. Indirect support for this assumption is provided by findings showing that German subjects in general and German students in particular consider the parity principle as more just than the equity or achievement principle (Schmitt & Montada, 1982). Contrary to this pattern, North Americans prefer the equity principle over other principles, especially when material resources are to be distributed (Törnblom & Foa, 1983). This cultural difference may explain why the subjects in the Schmitt et al. (1991) experiment behaved other than subjects in some of the North American studies reviewed earlier.

In a follow-up study by Schmitt (1991), this speculation was tested directly. Subjects' attitudes toward equity and equality were measured in addition to BJW. Following the interpretation offered by Schmitt et al. (1991), it was assumed that subjects who favor the equality principle would behave like the subjects in the Schmitt et al. (1991) study, but in a more pronounced manner. Subjects who favor equity (as the principle which comes closest to the concept of deserving) were expected to show the opposite pattern of attributions. More specifically, it was assumed that subjects with a favorable attitude toward equity and high BJW would tend to upgrade winners and downgrade losers, while subjects with a favorable attitude toward equity and low BJW would downgrade winners and upgrade losers.

For reasons which are not relevant in the present context, a different stimulus situation was chosen by Schmitt (1991) than by Schmitt et al. (1991). Subjects in the Schmitt (1991) study were shown a video film (presumably a live transmission) of an interview with a first-year psychology student. Questions were asked about five topics: Trier University, the housing situation for students in Trier, financial situation, most extraordinary events during the last year, activities during the semester break. A single experimental factor was varied between subjects: Good Luck vs. Bad Luck.

In the Good Luck condition, the student portrayed himself as a winner who had been exceptionally lucky and fortunate. (1) He had wanted to be a student at Trier University and he been admitted based on a lottery (a real possibility in Germany). (2) Given a lucky coincidence, he had found a very nice and cheap apartment. (3) He had just won money in a speculation that seemed hopeless. (4) Someone had run into his old car and the insurance paid him more money than he had payed for the car. (5) He was lucky to have found a perfect job for the semester break. The target stressed that he owed these fortunate events to chance or factors that were outside his control. He mentioned several times that he didn't really deserve so much luck and that he thought he was a "Glückspilz" (lucky dog). These remarks were designed to create a need for justification by making it difficult to attribute the favorable events to willful actions on behalf of the target. According to JMT, this should advance ascriptions of (good or bad) character in order to construe the episode as just.

In the Bad Luck condition, the student portrayed himself as a loser who had been very unlucky. (1) He wanted to be a student in Hamburg but was assigned Trier by the Central Agency for the Distribution of Students (a real institution in Germany). (2) Because of a broken public phone, he missed getting a nice and cheap apartment. (3) He had just lost money in a speculation that seemed safe. (4) Someone had run into his special car, an oldtimer. The insurance would pay him much less than he had payed for the car. (5) Because a roommate told him too late to return a phone call, he missed the perfect job for the semester break. The target stressed that his misfortune was not his fault but due to unfortunate circumstances which were outside his control. He mentioned several
times during the interview that he didn't deserve so much bad luck and that he thought he was a "Pechvogel" (unlucky fellow).

One hundred and twenty eight students participated in this experiment. BJW and attitudes toward equity and equality were measured in an independent study. Subjects were assigned randomly to experimental conditions and watched the interview in groups of five or less. Subjects were told that the experiment was designed to find out how well the personality of a person could be judged from the person's description of current life. After subjects had watched the interview, they were distributed a questionnaire which contained a number of manipulation control items and the same personality ratings as in the Schmitt et al. (1991) study. Principle axes analysis revealed a slightly different factorial structure here than in the Schmitt et al. (1991) study. One factor with good simple structure loaded highly (> .50) three trait ratings that overlapped with items from Factor 1 in the Schmitt et al. (1991) study (see above): interesting-boring, likable-unpleasant, attractive-unattractive. These items are clearly evaluative and cannot be inferred rationally from the target's behavior in the interview or his portrayal of himself. The sum of these items was chosen as the best available measure for the dependent variable.

Two-way interactions between the experimental factor (Good Luck vs. Bad Luck), BJW, and attitude toward equity/equality were expected and tested via moderated regression analyses. In addition, the three-way interaction (Experimental Factor x BJW x Attitude toward Equity x Attitude toward Equality) was tested to explore the possibility that only subjects with negative attitudes toward one principle and positive attitudes toward the other principle would show the expected moderator effect of BJW.

The only significant effect from these analyses was a main effect for the experimental factor. In line with JMT and previous experimental findings, the lucky target was upgraded and the unlucky target was downgraded (mean item scores amounting to 3.7 vs. 4.3, respectively, on six-point rating scales with 1/most likable and 6/most unpleasant personality).

This result adds to the inconsistency and ambiguity of empirical findings regarding the construct validity of BJW scales because (a) a moderator effect of BJW according the original expectation could not be found, (b) the opposite moderator effect of BJW found in the Schmitt et al. (1991) study could not be replicated, and (c) the assumptions on which the interpretations of the Schmitt et al. (1991) results rest could not be confirmed. Given the sample size, lack of statistical power cannot explain the result. Given the general derogation effect, it is difficult to dismiss the experiment as inappropriate for testing the reasoning offered by Schmitt et al. (1991).

**Implicit Time Perspectives in Belief in a Just World**

While the issue of implicit justice criteria has received little attention in the literature, the time perspective on which perceptions of justice are based and which may be implied in BJW have been discussed by several authors (e.g., Lerner, 1980; Maes, 1994a, in press; Rubin & Peplau, 1975) and two different time perspectives have mainly been differentiated: immanent justice and ultimate justice. What seems to be unjust in a narrow time perspective (immanent justice) may turn out to be just eventually (ultimate justice). Accordingly, individuals may be willing to invest costs into their education, a project, a business, or a relationship if they expect delayed rewards. From a short-term perspective, the lack of immediate rewards for such investments appears unjust while in the long run, inputs and outcomes are balanced. It can be expected from research in delay of gratification (e.g., Mischel, 1974) that individuals differ in the time perspective they apply when making such input-outcome comparisons. As a consequence, different individuals may judge the same situation as just or unjust, even if they apply the same distribution principle. Accordingly, BJW may differ between
individuals greatly depending on what time perspective they prefer. Unless we know these preferences and the corresponding beliefs, difficulties similar to the ones described in the last section seem unavoidable.

The issue has been demonstrated convincingly in a study by Maes (1994a, in press) who measured Belief in Immanent Justice and Belief in Ultimate Justice with regard to cancer. Both beliefs emerged as independent factors in factor analyses. More importantly, both types of BJW were related differently to some of the outcome variables included in JMT. More specifically, only subjects who believed in immanent justice tended to derogate cancer victims while subjects who believed in ultimate justice did not.

Stability and Occasion Specificity of Belief in a Just World

Rubin & Peplau (1973) assumed that BJW is a trait with stable individual differences. The temporal stability of BJW may therefore be considered as an additional construct validity criterion. Data on the temporal stability of BJW are rare. In one of our studies (Montada & Schneider, 1989), a six-month retest correlation of .75 was obtained for our GBJW scale. The true stability may be higher because retest correlations confound stability and reliability. However, the true stability may also be lower because retest correlations confound trait stability and method stability. In order to explore both possibilities, a latent state-trait analysis was conducted by Schmitt, Reichle, Schneider, & Steyer (1993) according to the methodology suggested by Steyer, Ferring, & Schmitt (1992).

Latent state trait theory (LSTT) can be considered an extension of Classical Test Theory (CTT; Lord & Novick, 1968). In CTT, a stable true-score is assumed for the person. Any change in the test score is attributed to unreliability. The possibility of systematic change in the person’s true score is not considered. Because of the stability postulate, the true score can be interpreted as the person’s trait score. In LSTT, a true score is assumed for each occasion of measurement. This score is interpreted as the person’s state. Systematic changes in the state from occasion to occasion are acknowledged and attributed to systematic influences between occasions of measurement or at occasions of measurement. The person’s state at an occasion of measurement is considered to be a function of the person’s trait and all systematic factors between or at occasions which lead to deviations of the states from the trait. In addition, the possibility that measurement variables contain some proportion of method variance (systematic measurement error) is considered. Method variance is reliable but invalid. According to these postulates of LSTT, each measurement variable can be decomposed into four independent latent variables:

1. The common trait is perfectly stable by definition across the time interval considered.
2. The occasion-specific factors lead to systematic intra-individual changes in the state from occasion to occasion and thereby to systematic deviations of the states from the trait.
3. The method factors represent systematic and stable but irrelevant proportions of variance of the measurement variables.
4. The latent error variables reflect all unsystematic sources of variance.

The variance of measurement variables can be decomposed accordingly. Since the latent variables are specified as independent, they explain unique and additive proportions of variance of measurement variables. Four coefficients can be defined which correspond to these proportions of variance:

1. Trait stability is the proportion of variance of the measurement variable that can be explained by the stable trait.
2. Occasion specificity is the proportion of variance that is due to occasion-specific factors.
(3) **Method specificity** is the proportion of variance that is due to method factors. 

(4) **Unreliability** is the proportion of variance that is due to unsystematic error variables. Accordingly, **reliability** is the sum of trait stability, occasion specificity, and method specificity.

In order to test a latent state-trait model and estimate these variance components, the construct has to be measured on at least two occasions by using at least two measurement methods. If no different methods are available, such as self-reports and peer reports, test halves have to be taken as methods.

Schmitt, Reichle, Schneider, & Steyer (1993) specified and tested a latent state-trait model for our GBJW scale. Test halves were used as methods by splitting items randomly into two sets of three items each. Occasions of measurement were six months apart. LISREL was used for testing the model and estimating parameters. The fit of the model was excellent ($\chi^2 = 5.92; p = .43$). The coefficients of stability, occasion specificity, method specificity, and reliability amounted to values of .69, .15, .07, and .91, respectively. These coefficients are important and interesting in several regards:

- Coefficient alpha (.79 in this sample) underestimates the reliability of our GBJW scale which is .91 according to the more appropriate latent state-trait analysis. This is an excellent value for a scale which has only six items.
- The method specificity coefficient (.07) reflects a partial nonequivalence (heterogeneity) of the test halves. This is the reason why the internal consistency of the scale is lower than its reliability.
- The coefficient of stability (.69) shows that, in line with Rubin & Peplau's (1973) assumption, BJW is considerably stable, at least across a period of six months.
- However, stability is not perfect. A significant and substantial amount of systematic variance is due to occasion-specific influences (.15). In other words, BJW does fluctuate systematically within an individual.

These coefficients cannot be taken as absolute values. Stability and occasion specificity may depend considerably on the events that occurred to the subjects of the sample during the study. Going back to the knowledge versus delusion issue and the studies comparing the BJW of East and West Germans, stability may go down considerably in times of severe societal change. Accordingly, individual stability versus occasion specificity of BJW may depend greatly on recent experiences of the person.

---

**Figure 4. Four Prototypical Changes in BJW due to the Observation of a Severe Injustice**

We know little, if anything, regarding changes in BJW when subjects are confronted with severe
cases of injustice. In Figure 4, only a few hypothetical reactions to such observations are suggested. Subjects may react to the observation of a severe injustice by lowering their BJW. Such a change would follow from the knowledge hypothesis discussed earlier. This change could be temporary, i.e., the impact of the event declines with temporal distance because the event becomes less salient in comparison to more recent events. Depending on how salient the event was in relation to other observations, the change might also be long term. The observation of injustice may also lead to temporary or enduring paradoxical changes in BJW. Such changes could be interpreted as defensive boosts and would accord to the delusion hypothesis.

Prospective longitudinal studies and systematic comparisons of individuals or groups with different justice experiences can provide the data we need for clarifying these speculations empirically. Using latent state-trait methodology, larger coefficients of occasion specificity and lower coefficients of stability can be expected for samples and for times with substantial societal changes, whereas larger stability coefficients and smaller fluctuations can be expected for samples living in stable environments. In addition to such prospective survey studies, the robustness of BJW could also be tested experimentally by administering BJW scales before and after confronting subjects with events that deviate from the subjects' justice beliefs. Latent state-trait analyses would be suitable for data from such studies as well.

Ambiguous Validation Criteria for Belief in a Just World Scales

A last issue that needs to be addressed is the quality and psychological meaning of variables that are used as criteria for assessing the construct validity of BJW scales. This issue has been discussed briefly in the context of group comparisons (see above). Groups differ on many variables and unless these were measured and controlled for statistically, it is difficult to know which of these variables accounts for group differences in BJW. But even if we know which variable is responsible for group difference in BJW, we do not yet know why. Two examples will be given for illustrating this issue: (1) well-being and emotions and (2) political conservatism as validation criteria for BJW.

Well-Being and Emotions

Dalbert (1993b) conducted a study with three groups of subjects: East German female students (n = 61), East German female teachers awaiting evaluation for competence and political background and facing the possibility of dismissal (n = 31), and East German jobless women (n = 54). BJW was measured with our GBJW scale and correlated positively in all three groups with indicators of well-being (life satisfaction, positive mood trait, positive mood state). Dalbert (1993b) interprets these positive correlations as evidence for the adaptive function of BJW.

Given the salience of joblessness, the result seems puzzling for the group of jobless because in this group, belief in immanent justice would imply that they deserved their fate. This should lower their self-esteem and affect their well-being negatively. The positive correlation would make sense if our GBJW scale measured belief in ultimate justice and if well-being reflected trust and hope in a more positive future.

In Limbach's (1992) study with 37 jobless East Germans and 69 jobless West Germans, the correlation between our GBJW scale and a homogeneous negative emotions scale (shame, rage against others, helplessness, fear of not finding work, guilt, anger about self) was negative for East Germans (-.71) and positive for West Germans (.21). The negative correlation for East Germans is consistent with Dalbert’s (1993b) result. Again, this would makes sense only if BJW meant belief in ultimate justice and if negative emotions indicated hopelessness. The positive correlation for West Germans would make sense if our GBJW scale measured belief in immanent justice and if the
negative emotions of West Germans included self-blame. But why should our GBJW scale measure belief in ultimate justice of East Germans although only one item conveys this notion explicitly (Item 4)? A possible speculation is that the notion of ultimate justice may predominate the notion of immanent justice due to a low individual self-efficacy in an authoritarian society and due to the great delay in consequences of collective behavior. Obviously, this is a very vague speculation, but it demonstrates how open to diverse interpretations correlations between BJW and other psychological variables can be.

**Political Conservatism**

Several authors have reported positive correlations between BJW and political conservatism (Dalbert et al., 1987; Dalbert, 1992; Furnham & Gunter, 1984; Rubin & Peplau, 1973; Schneider, 1988; Wagstaff, 1983). At least three interpretations for this correlation can be considered which differ greatly in the psychological functions of BJW that they imply.

1. Conservatism and BJW are both manifestations of a need for law, order, continuity, stability, and predictability. This interpretation is supported by the fact that political conservatism and BJW both correlate with authoritarianism (Dalbert, 1992; Rubin & Peplau, 1973).

2. Conservatism means wanting to preserve things as they are. If things are just, it is good to keep them. Note that this interpretation considers BJW as a causal explanation for conservatism.

3. As can be seen from Figure 5, the order of conservatism, political power, and BJW were the same while the studies who reported a positive correlation between BJW and conservatism were conducted. One interpretation for this consistent pattern is that a high BJW serves as a justification for political decisions made by politicians with political attitudes similar to the subject's attitudes. Accordingly, low BJW may signify a critique of political decisions by subjects who disagree with the leading party. Note that this interpretation implies that the correlation between BJW and conservatism is spurious and will reverse if the leading political parties are liberal.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year of data collection</th>
<th>Order of political power, conservatism, and BJW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dalbert et al. (1987)</td>
<td>1982</td>
<td>CDU/CSU &gt; FDP &gt; SPD &gt; Grüne</td>
</tr>
<tr>
<td>Dalbert (1992)</td>
<td>1990</td>
<td>CDU/CSU &gt; FDP &gt; SPD &gt; Grüne</td>
</tr>
<tr>
<td>Furnham &amp; Gunter (1984)</td>
<td>≈ 1982</td>
<td>Tories &gt; Labor</td>
</tr>
<tr>
<td>Rubin &amp; Peplau (1973)</td>
<td>≈ 1971</td>
<td>Republicans &gt; Democrats</td>
</tr>
<tr>
<td>Wagstaff (1983)</td>
<td>≈ 1981</td>
<td>Tories &gt; Labor</td>
</tr>
</tbody>
</table>

Figure 5. Order of Political Power and Order of BJW of Different Political Parties During the Data Collection of Studies Reporting a Positive Correlation Between BJW and Conservatism

Both examples show how much correlations of BJW with validation criteria are open to alternative interpretations. The same score on a validation criterion may reflect different psychological states and result from various psychological processes. Well-being, emotions, and political attitudes are multifaceted constructs with multiple causes. These causes may vary between individuals and
change within the person. Unless it is known what these causes are for a specific person at a specific moment in time, the co-occurrence of individual scores on the validation criterion and individual BJW cannot tell us safely what BJW measures. In the worst case, the same correlation may be interpreted as evidence in favor or against the construct validity of BJW scales. The foregoing discussion of differences in BJW between East and West Germans provides a good example of this problem.

Possible Resolutions

Human beings have a need for closure (Kruglanski & Webster, 1996) and consistency (Lecky, 1945). Ambiguous, inconsistent, and inconclusive data can be quite annoying. Yet, ignoring such data or discounting arguments that do not fit with the predominant theoretical or methodological Zeitgeist involves the danger of impeding scientific progress (Lane & Dunlap, 1978; Sterling, 1959). Science has often profited from controversy on the meaning of data (Kenrick & Funder, 1988). Inconsistent data call for a resolution and therefore stimulate the refinement of theories and methods (Lakatos, 1978). This chapter was an attempt to contribute to such an advancement in the domain of JMT and BJW by presenting some data and arguments which challenge the validity of BJW as an indicator of the JM. Certainly, not all available evidence was considered, and it cannot be claimed that the reported data are representative of all research that has been done in the field. Nevertheless, I consider the evidence to be sufficient for suggesting that it might be more difficult to measure individual differences in need for justice than has been acknowledged by researchers who have used BJW scales - including myself.

The worst possible reason for this would follow from Assumption 10 of JMT (see above) which states that people are unaware of their need for justice. Accordingly, Melvin Lerner has been sceptical on many occasions about the use of BJW scales for measuring the JM (e.g., Lerner, 1980, Chapter 10). In his recent writings, Lerner draws upon Epstein's level of processing model (Epstein, Lipson, Hostein, & Hub, 1992) to explain why blaming victims differs so greatly from our rational reactions to innocent suffering. Based on this model, Lerner suggests that basic beliefs and motives reside in the preconscious experiential system, they lead to quick and intuitive appraisals of situations, guide our behavior automatically, and cannot be accessed by introspection. The rational system operates logically, makes it possible to construe situations analytically, leads to conscious judgments based on knowledge, and provides a basis for planned behavior. Social norms as a domain of knowledge are part of the rational system. Both systems can operate independently from each other and often do so. The important consequence of this model for the validity of self-report measures for the JM is that the extent to which the JM and the need to believe in a just world are located in the experiential system, they cannot be accessed by introspection and cannot be transformed into answers to questionnaire items. Individuals may simply not know that they have a desire for justice and that they need to believe in a just world. To the extent that this is the case, self-report measures for JM cannot be valid. Yet before throwing away BJW scales, it may be wise to consider other explanations for the validity puzzle.

Perhaps the first step toward a resolution of this puzzle is giving up the assumption or requirement that the construct validity of BJW scales is a constant. It may be more appropriate and more useful for the design of further research to consider construct validity as being a variable which might depend on various factors such as the research setting, the substantive context, and characteristics of the subjects. This way of thinking about the validity of tests and questionnaires is not very common in psychology. However, sufficient research evidence is available suggesting that it is feasible and useful to consider the validity of any measurement instrument as varying across various types of moderators (Schmitt & Borkenau, 1992). This general view calls for psychological reflections on situational and individual conditions on which correlations and effects of BJW may depend. Several lines of
differentiation seem worth consideration.

A first line starts out from the fact that the individual difference perspective regarding the justice motive has been limited so far to its strength. It may be useful to additionally consider the possibility that individuals also differ in how the JM is transformed into specific emotions, cognitions, and actions. For example, individuals with an equally strong JM may differ in the means they prefer to use for defending or restoring their BJW. When observing a case of severe injustice, some individuals may choose to deny or to escape the situation while others may face the situation and try to do something to change it. These differences may be independent of the strength of the JM and exist in addition to situational factors suggesting one or another way of dealing with a threat to BJW. Given an equally strong JM, one person may tend to help an innocent victim while the other may prefer punishing the victimizer. It may be possible to resolve inconsistent findings partly by measuring individual differences in the means that are preferred for defending BJW. These preferences may be stable traits. For example, several studies have shown that individuals differ consistently and in a stable manner in altruism and helping behavior (e.g., Montada, Schmitt, & Dalbert, 1991; Silbereisen, Lamsfuss, Boehnke, & Eisenberg, 1991). If altruism as a trait is measured, it can be used as a moderator to allow tests for conditional effects of BJW. One would expect that for subjects scoring high on trait altruism, BJW has a stronger effect on helping innocent victims than for subjects scoring low on altruism. The opposite moderator effect can be expected for denial of responsibility as a trait (Schmitt, Montada, & Dalbert, 1991; Schwartz, 1977). To give a third example: Maes (1994b) has introduced the construct of draconity and proposed a self-report questionnaire for measuring individual differences in this trait. According to his data, individuals differ considerably in how harshly versus forgivingly they react to others who have made mistakes. It can be reasoned that for individuals with high draconity, BJW has a stronger effect on punishing victimizers than for individuals with low draconity.

A second line of differentiation seems necessary according to conceptual analyses (Lerner, 1980) and empirical evidence from research reported by Maes (1994a, in press). According to these data, the observation of innocent victims has different impacts for individuals whose justice concept contains the notion of long-term compensation (ultimate justice) than for individuals who limit their justice analysis to a limited time frame. While for the latter, the observation of innocent victims is threatening and may lead to derogation, the latter group can more easily maintain their belief in justice by assuming that the victim will be eventually compensated. Consequently, there is no or less need to distort perceptions regarding the causes of the misfortune and no or less need for distorted evaluations of the victim.

A third line of differentiation deals with the justice standards or criteria that individuals prefer when judging distributions or decision-making processes. Independent of the inconsistent results obtained in the Schmitt et al. (1991) study and the Schmitt (1991) study, it seems necessary, on conceptual grounds alone, to control for the justice principles that can be applied in a particular situation. Depending on which justice principle a person prefers or has in mind when making a justice related observation, this observation may affirm or threaten the person's belief in justice. Accordingly, quite different behavioral consequences are likely in the same situation for individuals who compare the situation with different ideals. A similar problem is given when we compare answers from different individuals to BJW items. For example, belief in justice according to equity has a different meaning than belief in justice according to equality. Equal distributions challenge the first type of BJW while equitable distributions challenge the second type of BJW. One reason why we chose a very general format for our GBJW scale was that we wanted to avoid confounding belief in justice with attitudes toward justice principles. Yet without measuring theses attitudes or the
person's specific justice concept in a particular situation, BJW remains ambiguous. Accordingly, it seems difficult to predict how the person will react in a particular situation. Simply asking subjects in the situation or shortly afterwards whether the situation was just or not ("manipulation control") can hardly provide the necessary information unless it were clear at what time defensive mechanisms began to operate. Whatever methodology may be more appropriate, well founded behavioral predictions from BJW according to JMT are only possible to the extent that we know the person's enduring justice concept and to the extent that the person refers to this concept when answering BJW items and when experiencing an experimental or real-life situation.

As a fourth suggestion, we might consider more direct indicators of the JM than BJW. A first possible candidate is justice centrality. Remember that justice centrality was the only organismic factor in the Herbst, Montada, & Schmitt study which interacted synergetically with the equivalent situation factor. This is an isolated result, of course, which should not be given much weight before it has been replicated. But, there is research in the domain of attitudes showing that the consistency of attitudes and values with behavior is a function of attitude/value centrality (e.g., Krosnick, 1988). A second possible candidate may be justice sensitivity, a construct that was recently introduced to the social justice literature by Schmitt, Neumann, & Montada (1995). In several studies, we have been able to show that individuals differ consistently across situations and time in their perceptual threshold for observing cases of injustice, in the intensity of their emotional reactions to injustice, in the mental intrusiveness of such events, and in action tendencies aimed at restoring justice (Schmitt & Mohiyeddini, 1996; Mohiyeddini & Schmitt, 1997; Dörfel & Schmitt, 1997). This sensitivity is generalized across the different roles a person can play in any incidence of injustice (victim, observer, victimizer; Schmitt et al., 1997a). In other words, a person who is more justice sensitive from the victim's perspective is also more sensitive from the remaining two perspectives. This generalized sensitivity and corresponding individual differences in the strength of action tendencies aimed at restoring justice may indicate a desire or need for justice according to JMT. Testing this possibility will require the adoption of the same kind of methodological rational used for exploring the construct validity of BJW scales.
References


Herbst, E. (1992). Zuschreibung von Verantwortlichkeit und Schuld gegenüber Opfern durch...


Fachbereich I - Psychologie (unpubl. diploma thesis).


Bisher erschienene Arbeiten dieser Reihe


Montada, L. & Reichle, B. (1983). Existentielle Schuld: Explikation eines Konzeptes (Berichte aus der Ar-
beitsgruppe "Verantwortung, Gerechtigkeit, Moral" Nr. 18). Trier: Universität Trier, Fachbereich I - Psychologie.


Gehri, U. & Montada, L. (1990). Schutz vor AIDS: Thematisierung in neuen Partnerschaften (Berichte aus der Ar-
beitsgruppe "Verantwortung, Gerechtigkeit, Moral" Nr. 56). Trier: Universität Trier, Fachbereich I - Psychologie.


Moral” Nr. 92). Trier: Universität Trier, Fachbereich I - Psychologie.


**Andernorts publizierte Arbeiten aus dieser Arbeitsgruppe**

Mütter. Regensburg: Roderer.


