**Hinkin, 1998**

• The Scale Development Process

Step 1 Item generation

1. Well-articulated theoretic foundation.

2. 2 approaches:

1) Deductive: (1) Logical partitioning/ classification from above.

(2) Theoretical foundation provides enough information to generate the initial set of items.

(3) Requires understanding of the phenomenon and a thorough understanding of the literature.

(4) Advantage: if properly conducted, will help to assure content validity.

Disadvantage: time-consuming; high working knowledge requirement.

2) Inductive: (1) Grouping/ classification from below.

(2) Appropriate when conceptual basis for a construct may not result in easily identifiable dimensions.

(3) Asking respondents to provide descriptions🡪 classify answers into categories by content analysis/ Q-sorting technique🡪items derived for subsequent factor analysis.

(4) Advantage: Useful for exploratory research and when it’s hard to generate items representing an abstract construct.

Disadvantage: No guarantee that derived items measure the same theoretical construct or come from the same sampling domain.

3. Item development: 1) Statements: simple and as short as possible.

2) Language: easy/ familiar to target respondents.

3) Consistency among items; address 1 issue per item.

4) Reverse-scored items: yes or no?

4. Content validity assessment:

1) Schriesheim et. al., (1993)

(1) Administer a set of items developed + their

definitions.

(2) Respondents rate on item-definition

correspondence.

(3) Q-correlation matrix🡪PCA🡪extract the

number of factors corresponding to the

theoretical dimensions.

(4) Factor loading 0.4+ / greater on the   
 appropriate factor with no major cross

loadings🡪 meaningful and representative items.

2) Anderson and Gerbing (1991)

(1) The proportion of respondents who assign an item to its intended construct.

(2) The degree to which each rater assigned an item to its intended construct.

3) Ask naïve respondents to match items with their corresponding definition; an acceptable agreement index must be determined in advance.

1) Keep it short to minimize biases caused by boredom/ fatigue.

2) At least 4 items per scale to test for homogeneity.

3) With 5 items, adding new ones are not likely to improve the internal consistency relibilities.

4) Adequate internal consistency reliabilities can be obtained with 3 items.

5) Assure that the domain has been adequately sampled.

6) Generate at least twice as many items as will be needed in the final scales.

\* Likert-type scales:

(1) The most frequently used in survey questionnaire research; the most useful in behavioral research; the most suitable for use in factor analysis.

(2) 5 equal appearing intervals with a neutral midpoint.

(3) Coefficient alpha reliability increases up to the use of 5 point, and then levels off.

(4) Researchers should accurately benchmark the response range to maximize the obtained variance.

5. Number of items:

6. Item scaling:

Step 2 Questionnaire administration

1. Collect info. From sources other than the respondents too ameliorate the common source/ common method concerns.

2. Sample size:

1) A large sample size🡪helps assure that factor loadings are accurate reflections of the true population values.

(2) Recommendation for item-to-response ratios: 1:4 or at least 1:10.

Step 3 Initial item reduction

1. Exploratory factor analysis

1) Factor analysis: allows the reduction of a set of variables to a smaller set.

2) Before conducting the factor analysis, researcher needs to examine the interitem correlations; delete any variable that correlates at less than .4.

3) The number of factors to be retained depends on both underlying theory and quantitative results.

4) Parsimony + simple structure; retain only items that clearly load on a single appropriate factor.

2. Internal consistency assessment

1) Reliability: the accuracy or precision of a measuring instrument; necessary condition for validity.

2) Cronbach’s alpha: the most commonly used measure.

3) A large coefficient alpha: strong item covariance + sampling domain has been captured adequately.

4) Alpha is very sensitive to the number of items in a measure.

5) 0.7 – minimum for newly developed measures; with proper factor analysis, reliability should be considerably higher than 0.7.

Step 4 Confirmatory factor analysis

1. A confirmation that the prior analyses have been conducted thoroughly and appropriately.

2. Conduct the analysis using the item variance-covariance matrix computed from data collected from an independent sample.

3. Purpose 1: to assess the goodness of fit of the measurement model comparing a single common factor model with a multitrait model with the number of factors equal to the number of constructs in the new measure.

4. Purpose 2: to examine the fit of individual items within the specified model using the modification indices and *t* values.

5. Chi-square: the smaller the chi-square the better the fit; a nonsignificant chi-square is desirable.

6. CFI and RNI may be more appropriate to determine the quality of fit.

7. The quality of the model can be further assessed by the item *t* value and modification indices.

8. If all appropriate loadings are significant at *p*<0.05 or less, and the magnitude of any inappropriate cross loadings as indicated by modification indices are relatively small, the researcher can be assured that the data fit the model quite well.

Step 5 Convergent/ Discriminant validity

1. Convergent validity: the extent to which the scales correlate with other measures designed to assess similar constructs.

2. Discriminant validity: the extent to which the scales do not correlate with dissimilar measures.

3. Multitrait-Multimethod Matrix (MTMM)

1) Is most commonly used to examine two types of validity.

2) Convergent validity is achieved when the correlation between measures of similar constructs using different methods are significantly different from zero and sufficiently large.

3) Discriminant validity is achieved when

(1) corr. monotrait-heteromethod > heterotrait-heteromethod;

(2) corr. monotrait-heteromethod > heterotrait-monomethod;

(3) corr. heterotrait-monomethod > heterotrait-heteromethod.

4. Alternative method: confirmatory factor analysis

5. Criterion-related validity

Step 6 Replication

1. The use of an independent sample—enhance the generalizability of the new measure.

2. When items are added or deleted, the “new” scale should be administered to another independent sample.

3. Collect data from sources other than the respondents and repeat the scale-testing process with the new scale.

4. The replication should include confirmatory factor analysis, assessment of internal consistency reliability, and convergent, discriminant, and criterion-related validity assessement.

**Colquitt, 2001**

1. The notion of justice has become increasingly visible construct in the social science over the last 30 years.

2. Initially—distributive justice: the justice of decision outcomes; is fostered where outcomes are consistent with implicit norms for allocation.

3. More recently—procedural justice: the justice of the processes that lead to decision outcomes; is fostered through voice during a decision-making process/ influence over the outcome/ by adherence to fair process criteria.

4. Research integrating procedural and distributive justice has found consistent support for a 2-factor conceptualization of organizational justice.

5. Interactional justice: interpersonal treatment people receive as procedures are enacted; is fostered when decision makers treat people with respect and sensitivity and explain the rationale for decisions thoroughly; clouded the clarity of 2-factor model.

6. Greenberg: 4-factor model for organizational justice.

7. Extremely high correlations between the 2 constructs suggests that some individuals may view justice from a 1-factor perspective.

**The Measurement of Organizational Justice**

1. Greenberg: many measurement efforts are plagued by items that attempt to measure one type of justice but that seem more applicable to another. Such problems are most common where interactional justice is the focus.

2. Perception of Fair Interpersonal Treatment Scale may artificially inflate the relationships among the different types of justice, and create theoretical and practical problems.

3. The practice of merging interactional and procedural justice prevents researchers from uncovering important differences between the constructs.

4. Procedural and interactional justice affected other variables through different intervening mechanism—agent system model.

5. Interactional justice affected 3 agent variables more than procedural justice, while procedural justice had a greater effect on a system variable.

6. The measure in this study is indirect measure, which assesses fairness criteria, instead of asking directly how fair something is.

**Study 1**

1. Criteria that guide the choice of outcome variables:

1) The outcomes had to be relevant to the study setting.

2) The outcomes had to represent both heavily researched and recently introduced outcomes.

3) The outcomes had to be applicable to both the instrumental and relational models of justice.

**Study Hypotheses**

1. Distributive justice will be positively related to outcome satisfaction.

2. Procedural justice will be positively related to rule compliance.

3. Interpersonal justice will be positively related to leader evaluation.

4. Informational justice will be positively related to collective esteem.

**Method**

1. All justice items were tailored to an education context by altering the parenthetical parts of the measure.

2. The organizational justice and outcome variables were administered as part of a survey given to students the week before the final exam. This timing ensured that the outcome (i.e., the current grade) was meaningful, given that the majority of the course's points had already been allocated (the final exam was 25% of the grade).

3. All outcome measures were assessed with 5-point Likert scales with anchors of 1 = *strongly disagree* and 5 = *strongly agree.*

4. Confirmatory factor analysis of the organizational justice and outcome variable items using EQS; maximum likelihood estimation; combined the measurement model with a structure model.

**Results and Discussion**

1. The best fitting model is the 4-factor model; the worst is the 1-factor model.

2. The results support the construct validity of the justice measure; adequate discriminant validity and predictive validity; the 4 organizational justice factors—distinct constructs.

**Study 2**

**Study Hypotheses**

1. Distributive justice will be positively related to instrumentality.

2. Procedural justice will be positively related to group commitment.

3. Interpersonal justice will be positively related to helping behavior.

4. Informational justice will be positively related to collective esteem.

**Method**

1. All justice items were tailored to apply to a field setting by altering the parenthetical parts of the measure.

2. The organizational justice and outcome variables were administered as part of a survey given to employees on company time, during their monthly group meetings.

3. All outcome measures were assessed with 5-point Likert scales with anchors of 1 = *strongly disagree* and 5 = *strongly agree.*

**Results and Discussion**

1. The best fitting model is the 4-factor model; the worst is the 1-factor model.

2. The justice measures possess a good degree of predictive validity on a zero-order basis.

3. All predicted path coefficients were significant; so all four hypotheses were supported.

4. Overall, the results of Study Two again support the construct validity of the organizational justice measure; the good fit of the structural model, together with the statistical significance of its paths, suggests adequate predictive validity. In addition, the fact that the four organizational justice factors predicted four different outcomes supports treating them as distinct constructs.

**General Discussion**

1. The results of both of the studies suggest that organizational justice is best conceptualized as four distinct dimensions: procedural justice, distributive justice, interpersonal justice, and informational justice.

2. Collapsing procedural and interactional justice together would mask important differences.

3. Interactional justice should be broken down into its interpersonal and informational justice components, as they too had differential effects.

4. The measure was able to discriminate among four different organizational justice factors. Thus, this measure provides the ability to statistically separate constructs that have always seemed to be distinct but have usually been combined because of high intercorrelations.

5. The correlations in Tables 3 and 4 suggest that it is possible to discriminate among the justice factors whether the researcher uses scale composites or latent variables.

6. The justice measure predicts an intentionally diverse set of outcomes taken from the existing literature.

The results shown here demonstrate that the justice measure is capable of being used with either model.

The results of both studies provide support for the recent body of research on the agent-system model.

**Limitations**

1. Further research is needed to refine construct validity, as scale development is an iterative process.

2. A further limitation of both studies was a reliance on self-report variables measured from the same source. This limitation raises the potential concern of effect size inflation due to same source bias-- further research is needed to validate the measure in a manner less susceptible to same source bias.

3. The distributive justice items in Table 1 do not specify a comparison other with whom the respondent can compare outcome-contribution relationships.

**Suggestions for Future Research**

1. Future research is needed that begins to separate the effects of the justice content from the effects of the justice source.

2. What are the effects of system versus agent-originating procedural, informational, or interpersonal justice?