**Study Hypotheses and Research Questions**

Hypothesis 1: Leader vision will positively relate to LMX scores (both closed- and open-ended responses on LMX)

Hypothesis 2: LMX (both closed- and open-ended responses) will positively relate to employee (a) empowerment, (b) perceived organizational support, (c) satisfaction with supervisor, and (d) turnover intentions

Research Question 1: To what extent are followers’ descriptions of LMX in their open-ended responses contingent upon the leader’s level in the organization?

Research Question 2: To what extent does LMX in their open-ended responses depend on country, gender (both leader and follower) and industry?

Research Question 3: To what extent do closed-ended LMX scores relate to open-ended LMX scores?

**Developing Scale Scores and Hypothesis Testing**

Although not appropriate in all circumstances, researchers often wish to quantify text in meaningful ways in order to run descriptive and inferential statistics. In order to minimize the loss of information in the qualitative to quantitative transformation, we first recommend that researchers consider identifying construct definitions from the extant literature or developing working definitions of the topics to facilitate the selection of words for a word list (Short, Broberg, Cogliser, & Brigham, 2010).

This moves us into the creation of the word lists that are used to generate the scale scores. We recommend that researchers select words from the word frequency lists that are consistent with the construct definition. In the current study, we first began by conducting topic modeling on half of the dataset. Through discussion we practiced selecting words for the word list. We then conducted analyses on the full dataset, established the number of topics, and calculated inter-rater reliability by having two raters independently select words for the individual relationship term. Inter-rater reliability can be established for chosen words using Cohen’s Kappa which is advantageous over the simple use of percent agreement as it takes into account the potential for agreement to occur by chance (Cohen, 1960). In the current study, across 100 coding decisions, percent agreement was 89% and Cohen’s Kappa was .78. As a heuristic, Cohen’s Kappa below .40 is considered to be problematic, .40 to .75 is acceptable, and .75 or higher may be interpreted as outstanding (Fleiss, 1981). The remaining word lists were also established independently and any discrepancies were resolved through discussion. Next, the word lists should be refined and finalized to ensure they are representative of the construct definition. In this step, we recommend that researchers ask at least two subject matter experts to independently review the word lists and provide any suggestions which we did in this study.

In order to calculate the scale score for the open-ended LMX variable (LMX-O), we calculated in DICTION (Wordstat is another user friendly option) the frequency with which these words occurred in each participant response (Short et al., 2010). We then divided the word frequency by the total number of words used. As participants also used negative words to describe a lack of a quality relationship with their supervisor (e.g., confused, friction, fault), we also developed a second word list by inductively coding about twenty percent of the open-ended responses. A second score was then developed using the second list and the negative score was then subtracted from the positive score to calculate the final LMX-O score. We conducted analyses in DICTION with un-stemmed and stemmed word list. We found that the conclusions of statistical significance did not change and that the general magnitude of the correlations largely did not change. What we ultimately reported were scales scores derived from the stemmed word list.

There are a few additional steps that researchers can conduct if the study context permits it. First, when possible, researchers could evaluate external validity by replicating one’s work with an independent sample. While we were not able to do this with the current study, we do recommend that researchers conduct such comparisons when possible. Other steps that have been recommended include considering transient error (via test-retest), specific factor error (using parallel forms), and algorithm error (Kripepndorff’s alpha); for a review see McKenny et al. (2016). It may also be valuable to consider the dimensionality of the emerging topics. One means to do this is by comparing the magnitude of correlations. Small magnitude correlations may be an indication that dimensions are unique; large magnitude correlations could suggest that the dimensions should be simply reduced to an overarching construct. We did not reach this point in the current study because of the large amount of overlapping words across the three dimensions identified due to our focus on only one construct (i.e., LMX).

Finally, researchers could evaluate the predictive validity variables (i.e., the scale scores). The predictive validity of the variables can be considered using traditional techniques, such as correlation and regression analyses as well as structural equation modeling. Using the scale scores that we developed from the topics, we tested the proposed example hypotheses and provide answers to the research questions. We illustrate the results of the correlation analyses in

[Insert Table 4 about here]

For Hypothesis 1, we posited that leader vision will positively relate to LMX scores (both closed- (LMX-C) and open-ended responses (LMX-O)). We found support for this hypothesis as vision was positively correlated with both LMX-C and LMX-O (*r* = .62, *p* < .01 and *r* =.15, *p* < .01, respectively). For Hypothesis 2, we proposed that LMX-C and LMX-O will positively relate to employee (a) empowerment (*r* = .42, *p* < .01 and *r* =.07, *p* > .05, respectively), (b) perceived organizational support (POS; *r* = .69, *p* < .01 and *r* =.21, *p* < .01, respectively), (c) satisfaction with supervisor (*r* = .74, *p* < .01 and *r* =.27, *p* < .01, respectively), and (d) turnover intentions (*r* = -.59, *p* < .01 and *r* = -.20, *p* < .01, respectively). These results illustrate support for Hypothesis 2 with the exception of the POS-LMX-O relation.

Next, we suggested several research questions. First, we asked to what extent are followers’ descriptions of LMX in their open-ended responses (LMX-O) contingent upon the leader’s level in the organization? We completed independent samples *t*-tests to consider the mean difference LMX-O reported by employees with their leaders at the executive-level and leaders at the mid-level manager level. The results illustrated that there was not a statistically significant difference in LMX-O (*t* (557) = .40, *p* > .05). We then asked to what extent does LMX-O depend on country, gender (both leader and follower), and industry? When comparing the LMX-O scores of employees in the U.S. versus abroad, employees abroad tended to have higher LMX-O scores (*t* (582) = -3.4, *p* < .01). We compared employees in professional settings relative to those in an industrial/manufacturing setting and did not find significant differences (*t* (486) = 1.42, *p* > .05). Results of the analysis also demonstrated that there was no mean-difference between LMX-O scores based on either the gender of the follower or gender of the leader. Finally, we asked to what extent do closed-ended LMX scores relate to open-ended LMX scores? Our analyses showed these two measures were correlated at *r* = .25 (*p* < .01). While the results of these hypotheses and research questions may be of minimal interest to leadership scholars, the primary aim of conducting such examinations was to illustrate to readers the potential means through which one might use text in more traditional quantitative analyses.

**Table 4. Means, standard deviations, and correlations of all variables**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | Mean | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1. Vision | 3.66 | 0.95 | .91 |  |  |  |  |  |  |
| 2. POS | 4.88 | 1.27 | .62\*\* | .90 |  |  |  |  |  |
| 3. LMX-C | 4.80 | 1.36 | .54\*\* | .69\*\* | .96 |  |  |  |  |
| 4. LMX-O | .02 | 0.14 | .15\*\* | .21\*\* | .25\*\* | - |  |  |  |
| 5. EM | 4.05 | 0.96 | .22\*\* | .41\*\* | .42\*\* | .07 | .91 |  |  |
| 6. SS | 4.00 | 1.18 | .51\*\* | .62\*\* | .74\*\* | .27\*\* | .46\*\* | .92 |  |
| 7. TI | 2.41 | 1.24 | -.50\*\* | -.65\*\* | -.59\*\* | -.20\*\* | -.38\*\* | -.66\*\* | .87 |

*Note. N* = 584~585. Coefficients alpha are listed in the diagonal where appropriate. † *p*<0.1; \* *p*<0.05; \*\* *p*<0.01 (two-tailed tests). POS=Perceived organizational support; LMX-C=Leader-member exchange-closed-ended; LMX-O=Leader-member exchange-open-ended; EM=Empowerment (self-determination); SS=Satisfaction with supervisor; TI=Turnover intentions

References

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement, 20*: 37-46.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd Ed.). New York, NY: John Wiley.

McKenny, A. F., Aguinis, H., Short, J. C., & Anglin, A. H. (2016). What Doesn’t Get Measured Does Exist Improving the Accuracy of Computer-Aided Text Analysis. *Journal of Management*: 0149206316657594.

Short, J. C., Broberg, J. C., Cogliser, C. C., & Brigham, K. H. (2010). Construct validation using computer-aided text analysis (CATA) an illustration using entrepreneurial orientation. *Organizational Research Methods, 13*: 320-347.