

Discussion of Reported Causes on Communication Problem Frequency.

- Causes Correlation Analysis
- Causes Regression Analysis

We note that all the cause correlation with the communication problems response are negative since the response is ordered more severe with response value, and the causes are ordered most significant as value increases.

The general 95% margin of error for correlation coefficients is around +/- 30%. For a rule of thumb, estimated correlation coefficients greater than 30% in absolute value are most likely significant.

From the correlation analysis we see that none of the CI or II variable sets have any significant correlation with the frequency of communication problem response. CI and II do have variables within their sets showing significant correlation amongst themselves, but just not with the response. For example, all the “stupid writing mistakes” are highly correlated with misspelled words.

For LI we have significant correlation with questions b2a (use of jargon) and b2b (misspelled words), although the misspelled words are correlated significantly only at the 13% significance level.

For CD questions b3c (company req'd language resented) and b3d (cultural behaviors offended other party) are marginally significant at about 13%. These both indicate a hostile environment so we expect these to contribute to communication problems.

Causes Correlation Analysis

Problem Frequency Correlation - LI: LANGUAGE INTERPRETATION

	A19	B2a	B2b	B2c	B2d	B2e	B2f	B2g	B2h	B2i
A19	1	-0.4	-0.22	-0.1	-0.17	-0.12	-0.09	0.05	-0.01	-0.05
B2a	-0.4	1	0.4	0.28	0.31	0.31	0.32	0.27	0.25	0.38
B2b	-0.22	0.4	1	0.46	0.6	0.63	0.62	0.59	0.64	0.61
B2c	-0.1	0.28	0.46	1	0.53	0.6	0.68	0.43	0.52	0.52
B2d	-0.17	0.31	0.6	0.53	1	0.6	0.67	0.57	0.62	0.57
B2e	-0.12	0.31	0.63	0.6	0.6	1	0.89	0.67	0.72	0.61
B2f	-0.09	0.32	0.62	0.68	0.67	0.89	1	0.74	0.82	0.71
B2g	0.05	0.27	0.59	0.43	0.57	0.67	0.74	1	0.9	0.66
B2h	-0.01	0.25	0.64	0.52	0.62	0.72	0.82	0.9	1	0.76
B2i	-0.05	0.38	0.61	0.52	0.57	0.61	0.71	0.66	0.76	1

Variables:

- B2a Cause of comm breakdown – LI – use of jargon
- B2b Cause of comm breakdown – LI – misspelled words
- B2c Cause of comm breakdown – LI – translation problems
- B2d Cause of comm breakdown – LI – lack/improper use of accent marks
- B2e Cause of comm breakdown – LI – improper syntax/word order
- B2f Cause of comm breakdown – LI – improper use of words
- B2g Cause of comm breakdown – LI – punctuation errors

B2h Cause of comm breakdown – LI – grammatical errors

B2i Cause of comm breakdown – LI – homonyms

Problem Frequency Correlation - CD: CULTURAL DIFFERENCES

A19	B3a	B3b	B3c	B3d	B3e	B3f	
A19	1	-0.19	-0.09	-0.22	-0.22	-0.18	-0.15
B3a	-0.19	1	0.26	0.37	0.24	0.47	-0.08
B3b	-0.09	0.26	1	0.69	0.6	0.25	0.32
B3c	-0.22	0.37	0.69	1	0.73	0.3	0.18
B3d	-0.22	0.24	0.6	0.73	1	0.32	0.15
B3e	-0.18	0.47	0.25	0.3	0.32	1	0.27
B3f	-0.15	-0.08	0.32	0.18	0.15	0.27	1

Variables:

B3a Cause of comm breakdown – CD – didn't understand urgency

B3b Cause of comm breakdown – CD – co. req'd language not appropriate

B3c Cause of comm breakdown – CD – co. req'd language resented

B3d Cause of comm breakdown – CD – cultural behaviors offended other party

B3e Cause of comm breakdown – CD – power/status of request misunderstood

B3f Cause of comm breakdown – CD – personality problems

Problem Frequency Correlation - CI: CONTEXT ISSUES

A19	B4a	B4b	B4c	B4d	B4e	B4f	B4g	
A19	1	-0.2	0.09	-0.06	-0.1	-0.12	-0.03	-0.06
B4a	-0.2	1	0.3	0.39	0.55	0.1	0.38	0.33
B4b	0.09	0.3	1	0.43	0.17	0.49	0.44	0.52
B4c	-0.06	0.39	0.43	1	0.43	0.18	0.38	0.21
B4d	-0.1	0.55	0.17	0.43	1	0.24	0.14	0.29
B4e	-0.12	0.1	0.49	0.18	0.24	1	0.18	0.6
B4f	-0.03	0.38	0.44	0.38	0.14	0.18	1	0.51
B4g	-0.06	0.33	0.52	0.21	0.29	0.6	0.51	1

Variables:

B4a Cause of comm breakdown – CI – no access to tech. equipment

B4b Cause of comm breakdown – CI – understood urgency but no time

B4c Cause of comm breakdown – CI – transmission media not appropriate

B4d Cause of comm breakdown – CI – historical background missing

B4e Cause of comm breakdown – CI – overworked/overwhelmed

B4f Cause of comm breakdown – CI – request unreasonable

B4g Cause of comm breakdown – CI – bad timing

Problem Frequency Correlation - II: INFORMATION ISSUES

	A19	B5a	B5b	B5c	B5d	B5e	B5f
A19	1	0.11	0.07	0.06	0.04	0.12	0.14
B5a	0.11	1	0.6	0.54	0.49	0.47	0.47
B5b	0.07	0.6	1	0.64	0.77	0.55	0.58
B5c	0.06	0.54	0.64	1	0.54	0.44	0.4
B5d	0.04	0.49	0.77	0.54	1	0.73	0.69
B5e	0.12	0.47	0.55	0.44	0.73	1	0.66
B5f	0.14	0.47	0.58	0.4	0.69	0.66	1

Variables:

B5a	Cause of comm breakdown – II – asked wrong person
B5b	Cause of comm breakdown – II – Receiver distrusts Sender
B5c	Cause of comm breakdown – II – proprietary info requested by non-company personnel
B5d	Cause of comm breakdown – II – loss of power when giving up info
B5e	Cause of comm breakdown – II – outside scope of work
B5f	Cause of comm breakdown – II – didn't have clearance to access info

Causes Regression Analysis

Ordinary Least Squares and the Parsimonious Model. We performed an ordinary least squares multiple regression fit of a parsimonious model subset of our 28 predictors on the response variable. Missing data were excluded by observation, resulting in about 37 out of 47 observations being used in the full model. Larger degrees of freedom were obtained by regressing certain predictor classes by themselves, but the resulting model fit errors were larger.

To obtain an appropriate subset of predictors, we used both the modern stepwise modified-forward inclusion, and exhaustive search techniques. Appropriate in our context is finding smaller number of predictors which minimize regression error and which are as significant as possible. This reduction is necessary in view of our small sample size but nevertheless unique dataset (see data discussion). The stepwise technique is widely discussed; see Agostinelli (2002), Breaux (1968), or Efroymsen (1960)

The matrix of coefficients (response intercept and predictor coefficients) was estimated along with their standard errors. The t value is the estimate divided by its standard error. This would be an appropriate test statistic to test the null hypothesis that the value of the corresponding regression parameter is 0. The reported p-value indicates evidence against a null hypothesis that the parameter is zero, which is known as significance.

For the model itself, the RSE is the mean regression residual sum of square error, or $\sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$, allowing a hypothesis (F) test of the model coefficients being significantly different than zero, and reported p-value on the F -test is reported. An insignificant p -value and low R^2 indicates the fitted model does not explain the response.

All Predictors. The regression coefficients were first estimated separately by full-rank on each cause (e.g., LI, CD, CI and II); very few were significant. In fact, only LI.B2a (use of jargon) had a significant coefficient, which confirms the validity of the exploratory correlation analysis. We note that none of the residuals for either cause resulted in Normal residuals, which is an important sign that the linear model assumptions are valid.

Reduced Predictors. The stepwise approach signals its usefulness in extracting 9 predictors which make a contribution to the response under the linear model assumption. The resulting model has Normal residuals and an R^2 of 57%. This R^2 is not as high as we would like it, but it is relatively high for work in the social sciences.

The following are the included variables for the multivariate regressions:

Regression - Reduced Model: Frequency on Causes

Residuals: Min 1Q Median 3Q Max
 -1.54 -0.326 0.0449 0.369 0.839

Residual standard error: 0.585 on 37 degrees of freedom

Multiple R-Squared: 0.571

F-statistic: 5.47 on 9 and 37 degrees of freedom, the p-value is 0.0000902

Coefficients:

	Value	Std. Error	t value	Pr(> t)	Signif
(Intercept)	3.884	0.212	18.343	0	***
B2a	-0.253	0.055	-4.605	0	***
B2h	-0.201	0.102	-1.98	0.055	*
B2i	0.432	0.119	3.618	0.001	***
B3d	-0.232	0.076	-3.068	0.004	***
B3e	-0.13	0.066	-1.965	0.057	*
B4a	-0.3	0.086	-3.485	0.001	***
B4f	0.099	0.069	1.433	0.16	
B5e	0.221	0.078	2.824	0.008	***
B5f	0.26	0.086	3.041	0.004	***
					Significant at 1% ***
					Significant at 5% *

- B2a: Cause of comm breakdown – LI – use of jargon
- B2h: Cause of comm breakdown – LI – grammatical errors
- B2i: Cause of comm breakdown – LI – homonyms
- B3d: Cause of comm breakdown – CD – cultural behaviors offended other party
- B3e: Cause of comm breakdown – CD – power/status of request misunderstood
- B4a: Cause of comm breakdown – CI – no access to tech. equipment
- B4f: Cause of comm breakdown – CI – request unreasonable
- B5e: Cause of comm breakdown – II – outside scope of work
- B5f: Cause of comm breakdown – II – didn't have clearance to access info

Interpretation of regression coefficients. For the Frequency of Communication Problems response, we first note that the intercept corresponding to zero coefficients on respondents' ranking of cause importance is 3.88, which corresponds to 9-10 problems per month. I.e., with no weight being assigned to the cause, the response is high on the problem side. This would lead us to conclude that independent variable (cause) weights would be mostly negative since positive weights would increase the response, and this is what we see in 5 of 9 cases. Those causes

adding major positive weight to frequency of problems are b2i (LI - homonyms), b5e (II - Outside scope of work) and b5f (II - no security clearance). We note the the LI variable has twice the weight at the II variables, and that agreement on the importance of the II causes are less prone to solution intervention.

The remaining variables have negative weights, indicating that agreement on the importance of the cause results in a lower frequency of problems, which should be borne out in the solutions analysis.

The following table summarizes the correlation among the causes and with the response:

Problem Frequency Correlation - Reduced Causes Model

	A19	B2a	B2h	B2i	B3d	B3e	B4a	B4f	B5e	B5f
A19	1	-0.4	-0.01	-0.05	-0.22	-0.18	-0.2	-0.03	0.12	0.14
B2a	-0.4	1	0.25	0.38	0.22	0.09	0.33	0.2	0.24	0.25
B2h	-0.01	0.25	1	0.76	0.51	0.3	0.28	0.48	0.33	0.27
B2i	-0.05	0.38	0.76	1	0.49	0.37	0.38	0.46	0.18	0.08
B3d	-0.22	0.22	0.51	0.49	1	0.32	0.43	0.46	0.39	0.42
B3e	-0.18	0.09	0.3	0.37	0.32	1	0.35	0.4	0.31	0.15
B4a	-0.2	0.33	0.28	0.38	0.43	0.35	1	0.38	0.6	0.6
B4f	-0.03	0.2	0.48	0.46	0.46	0.4	0.38	1	0.17	0.31
B5e	0.12	0.24	0.33	0.18	0.39	0.31	0.6	0.17	1	0.66
B5f	0.14	0.25	0.27	0.08	0.42	0.15	0.6	0.31	0.66	1

- B2a: Cause of comm breakdown – LI – use of jargon
- B2h: Cause of comm breakdown – LI – grammatical errors
- B2i: Cause of comm breakdown – LI – homonyms
- B3d: Cause of comm breakdown – CD – cultural behaviors offended other party
- B3e: Cause of comm breakdown – CD – power/status of request misunderstood
- B4a: Cause of comm breakdown – CI – no access to tech. equipment
- B4f: Cause of comm breakdown – CI – request unreasonable
- B5e: Cause of comm breakdown – II – outside scope of work
- B5f: Cause of comm breakdown – II – didn't have clearance to access info

References

Agostinelli, C. , "Robust stepwise regression", *Journal of Applied Statistics*, Aug2002, Vol. 29 Issue 6, p825-840

Kennedy, Peter, (1985). *A Guide to Econometrics*, 2nd Ed., MIT Press, Cambridge, MA.

HAROLD J. BREUX, "A Modification of Efroymsen's Technique for Stepwise Regression Analysis", *Communications of the ACM* 557, Volume 11/Number 8 August, 1968 [United States Army Ballistic Research Laboratories]

EFFROYMSON, M. A., "Multiple regression analysis." In Ralston, A., and Wilf, H. S. (Eds.), *Mathematical Methods for Digital Computers*, Wiley, New York, 1960, pp. 191-203.