Online Supplement for
More than Words in Medical Question-and-AnswerSites:
A Content-Context Congruence Perspective

Appendix A: Literature Review on Antecedents of the Perceived Value of Answers

Authors and Year	Answer Characteristics	Question Characteristics	Source Characteristics	Q&A Site Characteristics
Gazan (2006)			answerer expertise	
Jeon et al. (2006)	answer length, the number of times the answer is clicked, the number of times the answer is recommended by other users, the number of times that users print the answer		answerer expertise, answerer activity	
Kim et al. (2007)	content value, cognitive value, and socio-emotional value			
Adamic et al. (2008)	answer length		answerer expertise	
Harper et al. (2008)		question topic		paid-based vs. free-based sites
Kim and Oh (2009)	content, cognition, utility, socio-emotion		answerer expertise, answerer experience	
Chen et al. (2010)			answerer reputation	paid-based vs. free-based sites
Jeon et al. (2010)				paid-based vs. free-based sites
Shah and Pomerantz (2010)	answer length	question length, number of answers for the question, number of comments for the question	answerer profile	
Fichman (2011)				site popularity
Shah (2011)	answer time lapse			
Edelman (2012)	answer length, answer time lapse		answerer experience	
Lou et al. (2013)			internal motivation, external motivation	
Oh and Worrall (2013)	answer length		answerer expertise, answerer experience	
Zhang and Wang (2016)		question length		
Lee et al. (2019)	answer politeness			

Appendix B: Robustness Checks

We conducted several additional tests to examine the robustness of our results. First, a potential sample selection bias exists in the data, as not all answers have received helpfulness votes. More importantly, the likelihood of an answer being voted on may be correlated with the explanatory variables that predict answer helpfulness. To account for this potential bias, we employed Heckman's (1979) two-stage selection model as a robustness check. The first stage is a Probit "selection" equation that identifies the determinants of whether an answer was voted on or not. A vivid answer that draws readers' attention is more likely to receive votes, while a pallid answer that fails to draw attention is less likely to receive votes. We included author and answer characteristics that are related to information vividness in this stage, including author expertise, author credibility, answer length, answer readability, answer concreteness and answer emotional intensity (see Kuan et al. 2015). We also included the number of days since the answer was posted. In the second stage, the determinants of an answer's helpfulness are estimated using only voted answers, conditional on the first stage. As shown in Table B1, the inverse Mills ratio was significant ($p< 0.01$) and its inclusion in the stage 2 model alleviates potential bias due to sample selection and endogeneity (Shaver 1998). As shown in Model 2 of Table B1, the results of the second stage of Heckman's model did not qualitatively differ from the results reported earlier.

Table B1: Robustness Check with Heckman's Two-Stage Selection Model

	Model 1 First-stage selection equation	**Model 2** Second-stage outcome equation
Author Expertise	0.511***	0.034***
	(0.052)	(0.003)
Author Credibility	5.523***	-0.001
	(0.089)	(0.001)
Answer Length	0.263***	0.022***
	(0.015)	(0.001)
Answer Reading Difficulty	-0.050***	0.006***
	(0.015)	(0.001)
Answer Total Votes		0.002**
		(0.001)
Answer Days	-0.886***	0.118
	(0.020)	(0.133)

Answer Concreteness	-0.082***	0.001
	(0.012)	(0.001)
Answer Emotional Intensity	-0.081***	0.004***
	(0.011)	(0.001)
Answer Sequence	-0.162***	0.007***
	(0.011)	(0.001)
Other Helpful Answers	0.624***	-0.002
	(0.021)	(0.001)
Question Keywords	0.531***	-0.022***
	(0.022)	(0.002)
Question Length	-0.128***	0.0003
	(0.021)	(0.002)
Question-Answer Concreteness Congruence		0.003**
		(0.001)
Question-Answer Emotion Congruence		0.002**
		(0.001)
Disease-Answer Concreteness Congruence		0.006***
		(0.001)
Disease-Answer Emotion Congruence		0.004***
		(0.001)
Inverse Mills Ratio		-0.056***
		(0.005)
Constant	0.347***	0.338
	(0.016)	(0.394)
N	34894	16726
Chi2		4,612.25

All continuous independent variables standardized; Disease topic dummies, days of week dummies, and month-year dummies included; Standard errors in parentheses; ** $p < 0.05$, *** $p < 0.01$

Second, we useda number of alternative models to see if our findings would still hold. One alternative model was a multilevel regression model, chosen because most disease topics in WebMD Answers have more than one posted question. It is expected that answers addressing the same topic will be more similar to each other than to answers not addressing the same topic. Such similarity can cause intra-class correlation (ICC), which results in the standard errors of regression coefficients being underestimated (Klein and Kozlowski 2000; Kreft and Leeuw 1998; Raudenbush and Bryk 2002). Therefore, we useda random coefficient multilevel model to account for the interdependence of individual answers within the same topic (Model 1 of Table B2).Another chosen alternative was a fractional logit model. This approach was deemed appropriate because our dependent variable is a ratio bounded in the range of 0 to 1(Baum 2008). We usedfractional logit regressions to analyze the data. As shown in Model 1-2 of Table B2, we found consistent results by using these alternative models.

Table B2: Additional Robustness Checks

	Model 1 Multilevel	Model 2 Fractional Logit	Model 3 Tobit #votes >= 3	Model 4 Tobit >= 5	Model 5 Tobit >= 10
Author Expertise	0.041***	0.170***	0.042***	0.042***	0.045***
	(0.003)	(0.016)	(0.003)	(0.003)	(0.004)
Author Credibility	-0.001	-0.004	-0.0002	-0.0001	-0.001
	(0.001)	(0.007)	(0.001)	(0.001)	(0.001)
Answer Length	0.024***	0.108***	0.024***	0.024***	0.022***
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Answer Reading Difficulty	0.006***	0.026***	0.007***	0.008***	0.006***
	(0.001)	(0.006)	(0.001)	(0.001)	(0.001)
Answer Total Votes	0.002*	0.011	0.002**	0.002**	0.002*
	(0.001)	(0.008)	(0.001)	(0.001)	(0.001)
Answer Days	-0.004	0.046	-0.115	-0.118	0.036
	(0.133)	(0.557)	(0.130)	(0.133)	(0.155)
Answer Concreteness	0.0003	0.001	0.0002	0.001	0.0002
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Answer Emotional Intensity	0.003**	0.012***	0.003**	0.002*	0.002
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Answer Sequence	0.006***	0.023***	0.004***	0.002**	0.004***
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Other Helpful Answers	0.003**	0.011**	0.005***	0.006***	0.005***
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Question Keywords	-0.017***	-0.074***	-0.012***	-0.012***	-0.016***
	(0.002)	(0.008)	(0.002)	(0.002)	(0.002)
Question Length	-0.001	-0.004	-0.001	-0.001	0.001
	(0.001)	(0.006)	(0.001)	(0.001)	(0.002)
Question-Answer Concreteness Congruence	0.003***	0.013**	0.002*	0.002*	0.004**
	(0.001)	(0.005)	(0.001)	(0.001)	(0.002)
Question-Answer Emotion Congruence	0.003**	0.011**	0.004***	0.004***	0.003**
	(0.001)	(0.005)	(0.001)	(0.001)	(0.001)
Disease-Answer Concreteness Congruence	0.006***	0.023***	0.006***	0.007***	0.006***
	(0.001)	(0.005)	(0.001)	(0.001)	(0.002)
Disease-Answer Emotion Congruence	0.004***	0.017***	0.004***	0.004***	0.004***

	(0.001)	(0.005)	(0.001)	(0.001)	(0.002)
Disease Topic Dummies	N.A.	Included	Included	Included	Included
Days of Week Dummies	Included	Included	Included	Included	Included
*Month*Year Dummies*	Included	Included	Included	Included	Included
Constant	0.648	0.362	0.946**	0.955**	0.597
	(0.395)	(1.601)	(0.384)	(0.394)	(0.367)
N	16726	16726	15362	13757	9960
Log-likelihood	7,350.61	-7,646.56	7,374.56	7,170.59	5,504.65

All continuous independent variables standardized; Disease topic dummies included in all models except Model 1; Days of week dummies and month-year dummies included in all models; Standard errors in parentheses; * p < 0.10, ** p < 0.05, *** p < 0.01

Third, in the main analyses, we measured our dependent variable by dividing the number of helpful votes by the total number of votes for a given answer. Although this measure was commonly adopted in prior research, its accuracy may suffer when measuring extreme values. For example, assume that two answers both get zero helpful votes, while one answer has ten not-helpful votes, and the other answer has one not-helpful vote. These two answers received the same score for answer helpfulness, but the latter answer may be perceived as more helpful. To reduce this concern, we conducted a set of additional analyses, in which we included only the answers whose total number of votes were at least 3, 5, or 10. As shown in Model 3-5 of Table B2, the results were again consistent with those reported earlier.

Fourth, we alleviate further concerns about possible endogeneity by employing the instrumental variable approach proposed by Lewbel (2012). This approach identifies instruments as simple functions of the observed regression variables when external instrumental variables are not available. Although similar to Arellano and Bond (1991)'s approach using panel data estimators, Lewbel (2012)'s method can be implemented in cross-sectional data sets. Identification hinges on finding regressors that are uncorrelated with the product of heteroskedastic errors. Lewbel (2012)'s approach has been used in many fields such as information systems, marketing, and accounting (e.g., Anderson and Core 2018; Hong et al. 2018; Kashyap and Murtha 2017).

We followed the approach in Lewbel (2012) to generate instruments for our four congruencevariables ($Question - Answer\ \widehat{Concreteness}\ Congruence$,

$\widehat{Question - Answer\ Emotion\ Congruence}$, $\widehat{Disease - Answer\ Concreteness\ Congruence}$, $\widehat{Disease - Answer\ Emotion\ Congruence}$). We used Hansen's J-statistic of over-identifying restrictions to examine the exogeneity of our four generated instruments and were unable to reject the null hypothesis that our instruments are uncorrelated with the errors (p-value > 0.10). This result suggests that our generated instrumental variables are valid and appropriate. As shown in Table B3, we found consistent results by using these generated instrument variables. Therefore, our earlier results do not appear to be driven by reverse causality or omitted variables.

Table B3: Robustness Check with Lewbel (2012)'s Instrument Variable Approach

	Model 1
Author Expertise	0.041^{***}
	(0.003)
Author Credibility	-0.001
	(0.001)
Answer Length	0.024^{***}
	(0.001)
Answer Reading Difficulty	0.006^{***}
	(0.001)
Answer Total Votes	0.002^{*}
	(0.001)
Answer Days	0.007
	(0.133)
Answer Concreteness	0.0002
	(0.001)
Answer Emotional Intensity	0.003^{**}
	(0.001)
Answer Sequence	0.006^{***}
	(0.001)
Other Helpful Answers	0.003^{**}
	(0.001)
Question Keywords	-0.018^{***}
	(0.002)
Question Length	-0.001
	(0.001)
$\widehat{Question - Answer\ Concreteness\ Congruence}$	0.003^{**}
	(0.001)
$\widehat{Question - Answer\ Emotion\ Congruence}$	0.003^{**}

	(0.001)
$Disease - Answer\ \widehat{Concreteness}\ Congruence$	0.006***
	(0.001)
$Disease - Answer\ \widehat{Emotion}\ Congruence$	0.004***
	(0.001)
Constant	0.599
	(0.395)
N	16726
Log-likelihood	7496.44

All continuous independent variables standardized; Disease topic dummies, days of week dummies, and month-year dummies included; Standard errors in parentheses; ** p < 0.05, *** p < 0.01

References

Adamic, L. A., Zhang, J., Bakshy, E., and Ackerman, M. S. 2008. "Knowledge Sharing and Yahoo Answers: Everyone Knows Something," *Proceedings of the 17th international conference on World Wide Web*: ACM, pp. 665-674.

Anderson, J. D., and Core, J. E. 2018. "Managerial Incentives to Increase Risk Provided by Debt, Stock, and Options," *Management Science* (64:9), pp. 4408-4432.

Arellano, M., and Bond, S. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *The Review of Economic Studies* (58:2), pp. 277-297.

Baum, C. F. 2008. "Stata Tip 63: Modeling Proportions," *Stata Journal* (8:2), pp. 299-303.

Chen, Y., Ho, T.-H., and Kim, Y.-M. 2010. "Knowledge Market Design: A Field Experiment at Google Answers," *Journal of Public Economic Theory* (12:4), pp. 641-664.

Edelman, B. 2012. "Earnings and Ratings at Google Answers," *Economic Inquiry* (50:2), pp. 309-320.

Fichman, P. 2011. "A Comparative Assessment of Answer Quality on Four Question Answering Sites," *Journal of Information Science* (37:5), pp. 476-486.

Gazan, R. 2006. "Specialists and Synthesists in a Question Answering Community," *Proceedings of the American Society for Information Science and Technology* (43:1), pp. 1-10.

Harper, F. M., Raban, D., Rafaeli, S., and Konstan, J. A. 2008. "Predictors of Answer Quality in Online Q&a Sites," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 865-874.

Heckman, J. J. 1979. "Sample Selection Bias as a Specification Error," *Econometrica* (47:1), pp. 153-161.

Hong, Y., Hu, Y., and Burtch, G. 2018. "Embeddedness, Prosociality, and Social Influence: Evidence from Online Crowdfunding," *MIS Quarterly* (42:4), pp. 1211-1224.

Jeon, G. Y., Kim, Y.-M., and Chen, Y. 2010. "Re-Examining Price as a Predictor of Answer Quality in an Online Q&a Site," *Proceedings of the SIGCHI conference on human factors in computing systems*: ACM, pp. 325-328.

Jeon, J., Croft, W. B., Lee, J. H., and Park, S. 2006. "A Framework to Predict the Quality of Answers with Non-Textual Features," *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, Seattle, Washington, USA: ACM, pp. 228-235.

Kashyap, V., and Murtha, B. R. 2017. "The Joint Effects of Ex Ante Contractual Completeness and Ex Post Governance on Compliance in Franchised Marketing Channels," *Journal of Marketing* (81:3), pp. 130-153.

Kim, S., Oh, J. S., and Oh, S. 2007. "Best-Answer Selection Criteria in a Social Q&a Site from the User-Oriented Relevance Perspective," *Proceedings of the 70th annual meeting of the American Society for Information Science and Technology*, pp. 1-15.

Kim, S., and Oh, S. 2009. "Users' Relevance Criteria for Evaluating Answers in a Social Q&a Site," *Journal of the American Society for Information Science & Technology* (60:4), pp. 716-727.

Klein, K. J., and Kozlowski, S. W. J. 2000. *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*. San Francisco, CA: Jossey-Bass.

Kreft, I. G., and Leeuw, J. d. 1998. *Introducing Multilevel Modeling*. Thousand Oaks, CA: Sage.

Kuan, K. K. Y., Hui, K.-L., Prasarnphanich, P., and Lai, H.-Y. 2015. "What Makes a Review Voted? An Empirical Investigation of Review Voting in Online Review Systems," *Journal of the Association for Information Systems* (16:1), pp. 48-71.

Lee, S.-Y., Rui, H., and Whinston, A. B. 2019. "Is Best Answer Really the Best Answer? The Politeness Bias," *MIS Quarterly* (43:2), pp. 579-600.

Lewbel, A. 2012. "Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models," *Journal of Business & Economic Statistics* (30:1), pp. 67-80.

Lou, J., Fang, Y., Lim, K. H., and Peng, J. Z. 2013. "Contributing High Quantity and Quality Knowledge to Online Q&a Communities," *Journal of the American Society for Information Science and Technology* (64:2), pp. 356-371.

Oh, S., and Worrall, A. 2013. "Health Answer Quality Evaluation by Librarians, Nurses, and Users in Social Q&A," *Library & Information Science Research* (35:4), pp. 288-298.
Raudenbush, S. W., and Bryk, A. S. 2002. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage.
Shah, C. 2011. "Measuring Effectiveness and User Satisfaction in Yahoo! Answers," *First Monday* (16:2).
Shah, C., and Pomerantz, J. 2010. "Evaluating and Predicting Answer Quality in Community Qa," in: *Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval*. Geneva, Switzerland: ACM, pp. 411-418.
Shaver, J. M. 1998. "Accounting for Endogeneity When Assessing Strategy Performance: Does Entry Mode Choice Affect Fdi Survival?," *Management Science* (44:4), pp. 571-585.
Zhang, Y., and Wang, P. 2016. "Interactions and User-Perceived Helpfulness in Diet Information Social Questions & Answers," *Health Information & Libraries Journal* (33:4), pp. 295-307.