

Appendix A

Features and the Model for Predicting Expected Workload

A.1 Features

In the following complete list of features extracted and used for building the expect workload prediction model (cf. Sect. 5.5) is given.

A.1.1 *Syntax Based Features*

Feature	Description
characters	Number of characters visible to user by first time loading the HTML
fileSize	Size of HTML file
a_count_external	Number of external <a> tags
a_count_internal	Number of internal <a> tags i.e. links to other elements in the same page
button	Number of <button> tags
div	Number of <div> tags
span	Number of tags
p	Number of <p> tags
header_h1	Number of <h1> tags
header_h2	Number of <h2> tags
header_h3	Number of <h3> tags
header_h4	Number of <h4> tags
selection	Number of <selection> tags
input_count	Number of <input> tags, considering unique “name”
input_type_hidden	Number of <input type=“hidden”> tags (not visible to user)

input_type_button	Number of <input type="button"> tags
input_type_number	Number of <input type="number"> tags
input_type_checkbox	Number of <input type="checkbox"> tags, aggregated by "name"
input_type_radio	Number of <input type="radio"> tags, aggregated by "name"
input_type_submit	Number of <input type="submit"> tags
input_type_text	Number of <input type="text"> tags
input_type_url	Number of <input type="url"> tags
textarea	Number of <textarea> tags
sum_inputs	sum of all inputs which required user's action
img_count	Number of tags
img_under100px	Number of tags with width smaller than 100 pixel
audio	Number of <audio> tags
video	Number of <video> tags

A.1.2 *Semantic Based Features*

Feature	Description
sentences	Number of sentences visible to the user
subclauses	Number of subclauses visible to the user
words	Number of words in the text visible to the user
avgWordLength	Average word length
numUniqueStems	Number of unique stems
avgUniqueStems	Ratio of unique stems to the number of words

A.1.3 *Visual Features*

All of visual features were calculated based on an intermediate feature *Text density* which refers to ratio of the white pixels (text in the images) to the size of the window.

Feature	Description
threshold	Mean (over all windows) text density in the image
quartiles 25	25% of windows have less text density than or equal to this value
quartiles 50	50% of windows have less text density than or equal to this value

quartiles 75	75% of windows have less text density than or equal to this value
quartiles 95	95% of windows have less text density than or equal to this value
quartiles 98	98% of windows have less text density than or equal to this value
quartiles 100	100% of windows have less text density than or equal to this value
count_dense_windows	The number of dense windows
max_conn_region_w	Number of windows in the largest connected region
max_conn_region_p	Number of pixels in the largest connected region
avg_conn_region_w	The average size of connected regions (windows)
quar_conn_region 75	75% of connected regions contains less less windows than or equal to this value
quar_conn_region 98	98% of connected regions contains less less windows than or equal to this value

A.2 Prediction Model

The ensemble model, created for predicting the expected workload given the HIT design, consists of ten component models each with different predictors and corresponding coefficient. The result of the ensemble model is calculated by getting mean over results of all ten component models.

In the following the over all relative importance of top ten predictors, and coefficients of predictors on each component model are presented.

Predictor	Relative importance
HIT type	0.2489
quartiles 100	0.0685
quartiles 98	0.0595
quartiles 75	0.0554
quartiles 95	0.0537
Subclauses	0.0529
Reward	0.0452
numUniqueStems	0.0406
Words	0.0299
Threshold	0.0261

Predictor	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Constant	2.195	1.744	1.545	1.747	1.984	0.512	2.329	1.895	2.107	0.409
[HIT Type] (SV)	-0.283	-0.374	-0.351	-0.346	-0.338	-0.100	-0.255	-0.263	-0.376	-0.218
[HIT Type] (CC, IF)	0.356	0.338	0.341	0.303	0.382	0.384	0.399	0.328	0.349	0.434
[HIT Type] (CA, IA, VV)	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
reward	1.266	1.217	1.536	1.547	1.330	1.405	1.407	1.251	1.458	1.557
Syntax based features										
a_count_external								-0.004		
button	-0.060		-0.041	-0.066	-0.103	-0.089	-0.080	-0.092	-0.052	-0.085
span					0.006					0.005
p			0.007		0.005				0.007	
div		-0.002		-0.003	-0.002		-0.003			
header_h1	-0.231			-0.122					-0.175	
header_h2		-0.134				-0.109		0.111		
header_h3		0.081	0.073	0.102	0.083	0.068		0.098	0.058	
header_h4			-0.066		-0.033	-0.053		0.035	-0.047	0.049
input_count	-0.006				-0.008					-0.004
[input_type_file=0]	-1.580	-1.355	-1.224	-1.462	-1.404		-1.664	-1.424	-1.482	
[input_type_file=2]	0.000	0.000	0.000	0.000	0.000		0.000	0.000	0.000	
[input_type_submit=0]	0.153	0.132	0.268	0.276	0.247	0.243	0.369	0.216	0.233	
[input_type_submit=1]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
input_type_button	0.606				0.283			0.694		0.335
input_type_checkbox				0.022	0.034			0.040		0.026
input_type_hidden			-0.009	-0.016			-0.015	-0.023	-0.010	
input_type_number						-0.290	-0.138		-0.167	
input_type_radio					0.052	-0.046				
input_type_text		0.007	0.028	0.017	0.020			0.018		0.012
input_type_url	-0.384	-0.529		-0.531	-0.249	-0.597	-0.311	-0.256	-0.499	-0.525
textarea_count						0.045	0.044	0.050	0.038	
selection							0.024	0.023		
sum_inputs			-0.005					-0.007		
img_count								-0.013		
img_under100px	0.050	0.097		0.073			0.100		0.061	
video	0.335	0.343	0.277	0.453	0.399	0.351	0.530	0.447	0.723	0.320
Semantic based features										
avgUniqueStems	2.257		1.238			2.452	2.285		2.089	1.745
avgWordLength		0.077	0.084	0.103						
numUniqueStems		0.004		0.006	0.004			0.007		
sentences							-0.002		-0.002	-0.002
Subclauses	0.009				0.003	0.004	0.010		0.004	0.006
words				0.000						
Visual features										
quartiles 25		12.996	14.922			8.736	13.050			
quartiles 50			-2.499							
quartiles 75			0.906		-0.927		-1.240	-1.057		0.775
quartiles 95					0.987					
quartiles 98				0.336						
quartiles 100	1.031	0.983				0.674	0.480	0.744	0.489	0.905
max_conn_region_w									-0.285	
avg_conn_region_w			0.083			0.075		0.086	0.349	
threshold	-2.275	-1.646	-1.181			-2.046			-1.325	-1.450

References

1. Amitabh Mall Rachit Mathur, N.P.: Decoding digital @ retail: Winning the omnichannel consumer with digital in retail (2017)
2. Antin, J., Shaw, A.: Social desirability bias and self-reports of motivation: a study of Amazon mechanical turk in the U.S and India. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2925–2934, CHI 2012 (2012). <https://doi.org/10.1145/2207676.2208699>
3. Baard, P.P., Deci, E.L., Ryan, R.M.: Intrinsic need satisfaction: a motivational basis of performance and well-being in two work settings. *J. Appl. Soc. Psychol.* **34**(10), 2045–2068 (2004)
4. Bederson, B.B., Quinn, A.J.: Web workers unite! addressing challenges of online laborers. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems, pp. 97–106. ACM (2011)
5. Berg, J.: Income security in the on-demand economy: findings and policy lessons from a survey of crowdworkers. *Comparative Labor Law Policy J.* **37**(3) (2016)
6. Borst, I.: Understanding crowdsourcing: effects of motivation and rewards on participation and performance in voluntary online activities. EPS-2010-221-LIS (2010)
7. Brawley, A.M., Pury, C.L.: Work experiences on MTurk: job satisfaction, turnover, and information sharing. *Comput. Hum. Behav.* **54**, 531–546 (2016)
8. Breiman, L.: Bagging predictors. *Machine learning* **24**(2), 123–140 (1996)
9. Buhrmester, M., Kwang, T., Gosling, S.D.: Amazon’s mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives Psychological Sci.* **6**(1), 3–5 (2011)
10. Byrne, B.M.: *Structural Equation Modeling with Amos Basic Concepts, Applications, and Programming*. Routledge (2010)
11. Casner, S.M., Gore, B.F.: Measuring and evaluating workload: a primer. NASA Technical Memorandum (2010-216395) (2010)
12. Chandler, D., Kapelner, A.: Breaking monotony with meaning: motivation in crowdsourcing markets. *J. Economic Behav. Organization* **90**, 123–133 (2013)
13. Chandler, J., Mueller, P., Paolacci, G.: Nonnaïveté among amazon mechanical turk workers: consequences and solutions for behavioral researchers. *Behav. Res. Methods* **46**(1), 112–130 (2014)
14. Chapman, C.N., Alford, J.L., Johnson, C., Weidemann, R., Lahav, M.: CBC versus ACBC: comparing results with real product selection. In: 2009 Sawtooth Software Conference Proceedings, pp. 25–27 (2009)
15. Chen, J.: Enhanced recursive feature elimination. In: *Machine Learning and Applications, ICMLA 2007* (2007)
16. Cumming, G., Finch, S.: Inference by eye: confidence intervals and how to read pictures of data. *Am. Psychol.* **60**(2), 170 (2005)

17. Dai, P., Rzeszotarski, J.M., Paritosh, P., Chi, E.H.: And now for something completely different: improving crowdsourcing workflows with micro-diversions. In: Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing, pp. 628–638. ACM (2015)
18. De Waard, D.: The Measurement of Drivers' Mental Workload. Groningen University Traffic Research Center, Netherlands (1996)
19. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-determination in Human Behavior. Plenum (1985)
20. Deci, E.L., Ryan, R.M.: The “what” and “why” of goal pursuits: human needs and the self-determination of behavior. *Psychological Inquiry* **11**(4), 227–268 (2000). https://doi.org/10.1207/S15327965PLI1104_01
21. Difallah, D.E., Catasta, M., Demartini, G., Ipeirotis, P.G., Cudr-Mauroux, P.: The dynamics of micro-task crowdsourcing: the case of Amazon MTurk. In: Proceedings of the 24th International Conference on World Wide Web, WWW 2015, pp. 238–247. ACM (2015). <https://doi.org/10.1145/2736277.2741685>
22. Doan, A., Ramakrishnan, R., Halevy, A.Y.: Crowdsourcing systems on the world-wide web. *Commun. ACM* **54**(4), 86–96 (2011)
23. Doyle, C.: Work and Organizational Psychology: An Introduction with Attitude. Psychology at Work. Taylor & Francis (2004)
24. Efroymsen, M.: Multiple regression analysis. *Math. Methods Digital Comput.* **1**, 191–203 (1960)
25. Egelman, S., Chi, E.H., Dow, S.: Crowdsourcing in HCI research. In: Ways of Knowing in HCI, pp. 267–289. Springer, New York (2014)
26. Estellés-Arolas, E., González-Ladrón-De-Guevara, F.: Towards an integrated crowdsourcing definition. *J. Inf. Sci.* **38**(2), 189–200 (2012)
27. Felstiner, A.: Working the crowd: employment and labor law in the crowdsourcing industry. *Berkeley J. Employ. Labor Law*, pp. 143–203 (2011)
28. Field, A.: Discovering Statistics using IBM SPSS Statistics, 4 edn. MobileStudy, SAGE (2013)
29. Finnerty, A., Kucherbaev, P., Tranquillini, S., Convertino, G.: Keep it simple: reward and task design in crowdsourcing. In: Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI, p. 14. ACM (2013)
30. Gadiraju, U., Kawase, R., Dietze, S.: A taxonomy of microtasks on the web. In: Proceedings of the 25th ACM Conference on Hypertext and Social Media, HT 2014, pp. 218–223. ACM, New York, NY, USA (2014). <https://doi.org/10.1145/2631775.2631819>
31. Gadiraju, U., Kawase, R., Dietze, S., Demartini, G.: Understanding malicious behavior in crowdsourcing platforms: the case of online surveys. In: Proceedings of CHI, vol. 15 (2015)
32. Gagn, M., Deci, E.L.: Self-determination theory and work motivation. *J. Organizational Behav.* **26**(4), 331–362 (2005). <https://doi.org/10.1002/job.322>
33. Gagn, M., Forest, J., Vansteenkiste, M., Crevier-Braud, L., Van den Broeck, A., Aspel, A.K., Bellerose, J., Benabou, C., Chemolli, E., Gntert, S.T.: others: The multidimensional work motivation scale: validation evidence in seven languages and nine countries. *Eur. J. Work Organizational Psychol.* **24**(2), 178–196 (2015)
34. Geiger, D., Seedorf, S., Schulze, T., Nickerson, R.C., Schader, M.: Managing the crowd: towards a taxonomy of crowdsourcing processes. In: AMCIS (2011)
35. Guerin, B.: Social facilitation. In: The Corsini Encyclopedia of Psychology. Wiley, New York (2010)
36. Hart, S.G.: NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of The Human Factors and Ergonomics Society Annual Meeting, vol. 50, pp. 904–908. Sage Publications Sage, Los Angeles (2006)
37. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Adv. Psychology* **52**, 139–183 (1988)
38. Hauser, J., Ding, M., Gaskin, S.P.: Non-compensatory (and compensatory) models of consideration-set decisions. In: 2009 Sawtooth Software Conference Proceedings, Sequim WA (2009)

39. Herzberg, F.: One more time: how do you motivate employees? *Harvard Bus. Rev.* **46**(1), 53–62 (1968)
40. Hill, S.G., Iavecchia, H.P., Bittner Jr., A.C., Byers, J.C., Zaklad, A.L., Christ, R.E.: Comparison of four subjective workload rating scales. *Hum. Factors* **34**(4), 429–439 (1992)
41. Hirth, M., Hoßfeld, T., Tran-Gia, P.: Anatomy of a crowdsourcing platform-using the example of microworkers.com. In: 2011 Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), pp. 322–329. IEEE (2011)
42. Homburg, C., Giering, A.: Konzeptualisierung und operationalisierung komplexer konstrukte. ein leitfaden für die marketingforschung. *Marketing-Zeitschrift für Forschung und. Praxis* **18**(1), 5–24 (1996)
43. Hoßfeld, T., Heegaard, P.E., Varela, M., Möller, S.: QoE beyond the MOS: an in-depth look at QoE via better metrics and their relation to MOS. *Quality User Experience* **1**(1), 2 (2016)
44. Hoßfeld, T., Hirth, M., Korshunov, P., Hanhart, P., Gardlo, B., Keimel, C., Timmerer, C.: Survey of web-based crowdsourcing frameworks for subjective quality assessment. In: 2014 IEEE 16th International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6. IEEE (2014)
45. Hoßfeld, T., Hirth, M., Redi, J., Mazza, F., Korshunov, P., Naderi, B., Seufert, M., Gardlo, B., Egger, S., Keimel, C.: Best practices and recommendations for crowdsourced QoE - lessons learned from the qualinet task force crowdsourcing (2014)
46. Hoßfeld, T., Keimel, C.: Crowdsourcing in QoE evaluation. In: *Quality of Experience*, pp. 315–327. Springer, Cham (2014)
47. Hoßfeld, T., Seufert, M., Hirth, M., Zinner, T., Tran-Gia, P., Schatz, R.: Quantification of YouTube QoE via crowdsourcing. In: *Multimedia (ISM), 2011 IEEE International Symposium* pp. 494–499 (2011)
48. House, R.J., Wigdor, L.A.: Herzberg’s dual-factor theory of job satisfaction and motivation: a review of the evidence and a criticism. *Pers. Psychol.* **20**(4), 369–390 (1967)
49. Howe, J.: The rise of crowdsourcing. *Wired Magazine* **14**(6), 1–4 (2006)
50. Hu, L.t., Bentler, P.M.: Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equ. Modeling Multidisciplinary J.* **6**(1), 1–55 (1999)
51. Hurvich, C.M., Tsai, C.L.: Regression and time series model selection in small samples. *Biometrika*, pp. 297–307 (1989)
52. Iacobucci, D.: Structural equations modeling: fit indices, sample size, and advanced topics. *J. Consumer Psychology* **20**(1), 90–98 (2010). <https://doi.org/10.1016/j.jcps.2009.09.003>
53. Ipeirotis, P.: Demographics of mechanical turk: now live! (April 2015 edition) (2015). <http://www.behind-the-enemy-lines.com/2015/04/demographics-of-mechanical-turk-now.html>
54. Ipeirotis, P.G.: Demographics of mechanical turk (2010)
55. Ipeirotis, P.G.: Analyzing the Amazon mechanical turk marketplace. *XRDS* **17**(2), 16–21 (2010–12). <https://doi.org/10.1145/1869086.1869094>
56. Irani, L.C., Silberman, M.S.: Turkoicon: interrupting worker invisibility in Amazon mechanical turk. In: *Proceedings of CHI 2013*, pp. 611–620. ACM (2013). <https://doi.org/10.1145/2470654.2470742>
57. Kahneman, D., Diener, E., Schwarz, N.: *Well-Being: foundations of hedonic psychology*. Russell Sage Foundation (1999)
58. Kaufmann, N., Schulze, T., Veit, D.: More than fun and money. Worker motivation in crowdsourcing - a study on mechanical turk. *AMCIS* (2011)
59. Kazai, G., Kamps, J., Milic-Frayling, N.: The face of quality in crowdsourcing relevance labels: demographics, personality and labeling accuracy. In: *Proceedings of the 21st ACM International Conference on Information and Knowledge Management*, pp. 2583–2586. ACM (2012)
60. Keimel, C., Habigt, J., Horch, C., Diepold, K.: Qualitycrowd—a framework for crowd-based quality evaluation. In: *Picture Coding Symposium (PCS)*, pp. 245–248. IEEE (2012)
61. Kinnaird, P., Dabbish, L., Kiesler, S., Faste, H.: Co-worker transparency in a microtask marketplace. In: *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, pp. 1285–1290. ACM (2013)

62. Kittur, A., Chi, E.H., Suh, B.: Crowdsourcing user studies with mechanical turk. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2008, pp. 453–456. ACM (2008). <https://doi.org/10.1145/1357054.1357127>
63. Kittur, A., Nickerson, J.V., Bernstein, M., Gerber, E., Shaw, A., Zimmerman, J., Lease, M., Horton, J.: The future of crowd work. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work, p. 1301. ACM Press (2013). <https://doi.org/10.1145/2441776.2441923>
64. Koestner, R., Losier, G.F., Vallerand, R.J., Carducci, D.: Identified and introjected forms of political internalization: extending self-determination theory. *J. Pers. Soc. Psychol.* **70**(5), 1025 (1996)
65. Lam, C.F., Gurland, S.T.: Self-determined work motivation predicts job outcomes, but what predicts self-determined work motivation? *J. Res. Pers.* **42**(4), 1109–1115 (2008)
66. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. *Biometrics*, pp. 159–174 (1977)
67. Le Moan, S., Mansouri, A., Hardeberg, J.Y., Voisin, Y.: Saliency for spectral image analysis. *IEEE J. Selected Topics Appl. Earth Observations Remote Sensing* **6**(6), 2472–2479 (2013)
68. Loomis, J.B.: Strategies for overcoming hypothetical BIAS in stated preference surveys. *J. Agric. Resour. Economics* **39**(1), 34–46 (2014)
69. Lowry, P.B., Gaskin, J.: Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: when to choose it and how to use it. *IEEE Trans. Professional Commun.* **57**(2), 123–146 (2014)
70. Malone, T.W., Laubacher, R., Dellarocas, C.: Harnessing crowds: mapping the genome of collective intelligence. Massachusetts Institute of Technology (2009)
71. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: Association for Computational Linguistics (ACL) System Demonstrations, pp. 55–60 (2014). <http://www.aclweb.org/anthology/P/P14/P14-5010>
72. Martin, D., Carpendale, S., Gupta, N., Hoßfeld, T., Naderi, B., Redi, J., Siahaan, E., Wechsung, I.: Understanding the crowd: ethical and practical matters in the academic use of crowdsourcing, pp. 27–69. Springer International Publishing (2017). https://doi.org/10.1007/978-3-319-66435-4_3
73. Martin, D., Hanrahan, B.V., O’Neill, J., Gupta, N.: Being a turker. In: Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 224–235. ACM (2014)
74. Mason, W., Watts, D.J.: Financial incentives and the performance of crowds. *ACM SigKDD Explorations Newsl.* **11**(2), 100–108 (2010)
75. McAuley, E., Duncan, T., Tammen, V.V.: Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: a confirmatory factor analysis. *Res. Q. Exercise Sport* **60**(1), 48–58 (1989). <https://doi.org/10.1080/02701367.1989.10607413>
76. Miller, S.: Workload measures. National Advanced Driving Simulator. Iowa City, United States (2001)
77. Möller, S., Hoßfeld, T., Naderi, B.: Comments on P.CROWD (2016)
78. Möller, S., Raake, A.: Quality of Experience: Advanced Concepts. Applications and Methods. T-Labs Series in Telecommunication Services. Springer, New York (2014)
79. Murphy, J.J., Allen, P.G., Stevens, T.H., Weatherhead, D.: A meta-analysis of hypothetical bias in stated preference valuation. *Environ. Resource Econ.* **30**(3), 313–325 (2005)
80. Naderi, B., Möller, S.: Speech quality assessment in crowdsourcing: comparison with laboratory and recommendations on task design. Technical Report COM 12-C 039, Geneva, Switzerland (2017)
81. Naderi, B., Polzehl, T., Beyer, A., Pilz, t., Möller, S.: Crowdee: mobile crowdsourcing micro-task platform—for celebrating the diversity of languages. In: Proceedings of 15th Annual Conference of the International Speech Communication Association (Interspeech 2014). IEEE (2014)

82. Naderi, B., Polzehl, T., Wechsung, I., Kster, F., Möller, S.: Effect of trapping questions on the reliability of speech quality judgments in a crowdsourcing paradigm. In: 16th Annual Conference of the International Speech Communication Association (Interspeech 2015), ISCA, pp. 2799–2803 (2015)
83. Naderi, B., Wechsung, I., Möller, S.: Effect of being observed on the reliability of responses in crowdsourcing micro-task platforms. In: 2015 Seventh International Workshop on Quality of Multimedia Experience (QoMEX), pp. 1–2. IEEE (2015). <https://doi.org/10.1109/QoMEX.2015.7148091>
84. Naderi, B., Wechsung, I., Möller, S.: Crowdsourcing work motivation scale: development and validation for crowdsourcing micro-task platforms. In: Prepration (2016)
85. Naderi, B., Wechsung, I., Polzehl, T., Möller, S.: Development and validation of extrinsic motivation scale for crowdsourcing micro-task platforms. In: Proceedings of CrowdMM 2014, pp. 31–36. ACM (2014). <https://doi.org/10.1145/2660114.2660122>
86. Novotorova, N.: A Conjoint Analysis of Consumer Preferences for Product Attributes: The Case of Illinois Apples. University of Illinois at Urbana-Champaign (2007)
87. O’Neill, J., Martin, D.: Relationship-based business process crowdsourcing? In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) Human-Computer Interaction-INTERACT 2013, pp. 429–446. Springer, Heidelberg (2013)
88. Orme, B.K.: The cbc/hb system for hierarchical bayes estimation version 5.0 technical paper. Technical Paper Series, Sawtooth Software, Orem, UT (2009)
89. Orme, B.K.: Getting started with conjoint analysis: strategies for product design and pricing research. Research Publishers (2010)
90. Peer, E., Samat, S., Brandimarte, L., Acquisti, A.: Beyond the turk: an empirical comparison of alternative platforms for crowdsourcing online behavioral research **2594183** (2015). <https://doi.org/10.2139/ssrn>
91. Redi, J., Povaia, I.: Crowdsourcing for rating image aesthetic appeal: Better a paid or a volunteer crowd? In: Proceedings of the 2014 International ACM Workshop on Crowdsourcing for Multimedia, pp. 25–30. ACM (2014). <https://doi.org/10.1145/2660114.2660118>
92. Reid, G.B., Shingledecker, C.A., Eggemeier, F.T.: Application of conjoint measurement to workload scale development. In: Proceedings of the Human Factors Society Annual Meeting, vol. 25, pp. 522–526. Sage Publications Sage, Los Angeles (1981)
93. Reips, U.D.: Standards for internet-based experimenting. *Experimental Psychol.* **49**(4), 243 (2002)
94. Reiter, U., Brunnström, K., De Moor, K., Larabi, M.C., Pereira, M., Pinheiro, A., You, J., Zgank, A.: Factors influencing quality of experience. In: Quality of Experience, pp. 55–72. Springer, Cham (2014)
95. RICE, J., BAKKEN, D.G.: Estimating attribute level utilities from “design your own product” data—chapter 3. In: Sawtooth Software Conference (2006)
96. Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., Vukovic, M.: An assessment of intrinsic and extrinsic motivation on task performance in crowdsourcing markets. In: ICWSM (2011)
97. Ross, J., Irani, L., Silberman, M., Zaldivar, A., Tomlinson, B.: Who are the crowdworkers?: shifting demographics in mechanical turk. In: CHI 2010 Extended Abstracts on Human Factors in Computing Systems, pp. 2863–2872. ACM (2010)
98. Ryan, R.M., Connell, J.P.: Perceived locus of causality and internalization: examining reasons for acting in two domains. *J. Personality Soc. Psychol.* **57**(5), 749–761 (1989). <https://doi.org/10.1037/0022-3514.57.5.749>
99. Ryan, R.M., Deci, E.L.: Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educ. Psychol.* **25**(1), 54–67 (2000). <https://doi.org/10.1006/ceps.1999.1020>
100. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **55**(1), 68–78 (2000). <https://doi.org/10.1037/0003-066X.55.1.68>

101. Sauro, J., Dumas, J.S.: Comparison of three one-question, post-task usability questionnaires. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1599–1608. ACM (2009)
102. Shen, C., Zhao, Q.: Webpage saliency. In: European Conference on Computer Vision, pp. 33–46. Springer, Cham (2014)
103. Siddiqui, M.W., Eichelbaum, J., Llombart, V.C., Rosin, P.: Study project: predicting effort of microtask (2017)
104. Silberman, M.S.: What's fair? rational action and its residuals in an electronic market. Unpublished manuscript (2010). <http://www.scribd.com/doc/86592724/Whats-Fair>
105. Silberman, S., Milland, K., LaPlante, R., Ross, J., Irani, L.: Stop citing Ross et al. 2010, who are the crowdworkers? (2015)
106. Sparck Jones, K.: A statistical interpretation of term specificity and its application in retrieval. *J. Documentation* **28**(1), 11–21 (1972)
107. Stevenson, A.: Oxford dictionary of English (2010)
108. Tabachnick, B., Fidell, L.: Using Multivariate Statistics. Always Learning. Pearson Education (2012)
109. Tremblay, M.A., Blanchard, C.M., Taylor, S., Pelletier, L.G., Villeneuve, M.: Work extrinsic and intrinsic motivation scale: Its value for organizational psychology research. *Canadian J. Behav. Science/Revue Canadienne des Sci. du Comport.* **41**(4), 213–226 (2009). <https://doi.org/10.1037/a0015167>
110. Union, I.T.: ITU-t recommendation p. 800: Methods for subjective determination of transmission quality (1996)
111. Urošević, S., Milijić, N., et al.: Influence of demographic factors on employee satisfaction and motivation. *Organizacija* **45**(4), 174–182 (2012)
112. Walthert, D., Koch, C.: Modeling attention to salient proto-objects. *Neural Netw.* **19**(9), 1395–1407 (2006)
113. Warr, P.: Work, Happiness and Unhappiness. Taylor & Francis, New York (2011)
114. Wechsung, I.: An Evaluation Framework for Multimodal Interaction: Determining Quality Aspects and Modality Choice. Springer, Cham (2014)
115. Wechsung, I., Weiß, B., Kühnel, C., Ehrenbrink, P., Möller, S.: Development and validation of the conversational agents scale (CAS). In: Interspeech, pp. 1106–1110 (2013)
116. Winkler, S.: On the properties of subjective ratings in video quality experiments. In: International Workshop on Quality of Multimedia Experience, QoMEX 2009, pp. 139–144. IEEE (2009)
117. Wolters, K.M., Engelbrecht, K.P., Göttsche, F., Möller, S., Naumann, A., Schleicher, R.: Making it easier for older people to talk to smart homes: the effect of early help prompts. *Univ. Access Inf. Soc.* **9**(4), 311–325 (2010)
118. Zhao, Y., Zhu, Q.: Evaluation on crowdsourcing research: current status and future direction. *Inf. Syst. Frontiers* **16**(3), 417–434 (2014)
119. Zhao, Y.C., Zhu, Q.: Effects of extrinsic and intrinsic motivation on participation in crowdsourcing contest. *Online Inf. Rev.* **38**(7), 896 (2014)
120. Zijlstra, F.R.H.: Efficiency in work behaviour: A design approach for modern tools. Ph.D. thesis, TU Delft, Delft University of Technology (1993)
121. Zijlstra, F.R.H., Van Doorn, L.: The construction of a scale to measure perceived effort. University of Technology (1985)