Ipek Ensari, PhD
Entertainment to Affect Change (E2AC) Research Team
Nick Hudson, Michael Gabriel, Bene Cipolla, Sunita Uthra
August, 2023

IAS Development Analyses Appendix: Qualitative interviews

This is a brief overview of the methods and results for the analysis of the responses from the qualitative interviews to guide the design and development of the IAS.

Methods

Data Collection. As a part of the preliminary research to inform the content of the IAS, we conducted individual interviews with 3 filmmakers as participants. We used open-ended questions (See Table 1) on themes of impact in documentary to guide the conversation, and probed with follow-up questions (e.g., “How” or “Why” based explanations) to ask for further details on their responses as appropriate. All participants provided informed consent and agreed to have their responses recorded for use in aggregate, anonymized analyses. Interviews lasted approximately 45 minutes each. A copy of the interview was shared with each participant after completion of the session.

Table 1. List of probing questions used to guide the qualitative interviews.
1. What are some of the keywords or thoughts that come to your mind when you think of "social impact" of a film on the audience?
2. Can you tell me how you think about the potential social impact of it when you are working on a film project? you can use a specific example/project. 1. Does the potential social impact influence your creative process and how you go about writing/directing a movie?
3. What kind of impact do you hope to evoke in your audience? how do you think about this in the near- and long-term?
4. How do you define success in terms of the impact of your film? are there certain outcomes you hope to achieve regarding the potential social impact of your film? 1. are there certain platforms you care about/look at more to gauge impact/influence? Why? 2. donations- does the amount matter? certain roles that volunteers play - does this matter? 3. any awards that matter to you? why?
5. Can you tell me about a time/experience where you observed the downstream influence/impact of one of your movies?
6. How do you think of impact at the more immediate level, i.e., after an audience sees your movie?
7. What are some of the factors you think influence the social impact of a film?
8. How do you choose which documentaries to make/direct/produce? what is your decision-making process?

Data Analysis. Data were processed and analyzed using R/RStudio and tidytext,[1] lda,[2] stm,[3] quanteda,[4] topicmodels[5] packages. We first processed the data by stemming (reducing words to their root form) and removing punctuation and stop words. For the descriptive and correlational analyses, we tokenized the responses into bigrams (2-word terms). We used the *phi* correlation coefficient, which is a commonly used metric of binary correlation analogous to the Pearson’s *r* for continuous variables.[6] *Phi* coefficient is calculated based on the co-occurrence of 2 words (i.e., both words will occur OR neither will occur) in a document,[1] similarly to the point mutual information metric. Given our hypotheses and focus of the analyses, we conducted all analyses by using a list of words (in their stem forms) pertaining to our a priori determined list of target themes (i.e., impact, change, entertain, fun, like, behavior, influence, agree/disagree, influence, donate, share, volunteer, know/knowledge, learn, understand, thought).

To identify major underlying “latent” topics, we tested several topic modeling approaches, primarily relying on latent dirilecht allocation (LDA). LDA is a well-established approach for topic modeling, which is an unsupervised classification method for clustering text data.[7] LDA treats each document in the data as a mixture of topics and each topic as a mixture of words.[8] The LDA model looks for words that co-occur frequently in the documents (e.g., participant responses) to learn and identify the latent topics. The model simultaneously estimates the likelihood of each word being generated from a topic to identify the words that represent each topic[8] (i.e., those with highest likelihood). As such, each document is represented by a mixture of k topics and each topic is represented by a set of semantically related words.

To determine the optimal number of topics (k), we evaluated topic quality. This is done through the metrics of **semantic coherence** (i.e., *are the topics meaningful and making sense?*), which has been indicated to strongly correlate with human evaluation,[9-11] and **exclusivity** (i.e., *how unique/specific are the topic words to a given topic?*). Best model fit was further determined through residual dispersion, where lower values (<0) suggest potential spuriousness (i.e., *are the same themes being repeated across topics?*). Accordingly, values that are significantly above 1 might suggest existence of additional topics in the corpus and evaluation of models with higher *k*.

Results

Topic Correlation and Topic/Word Prevalence. Figure 1 depicts the associations between the most frequently occurring bigrams in the interview responses.

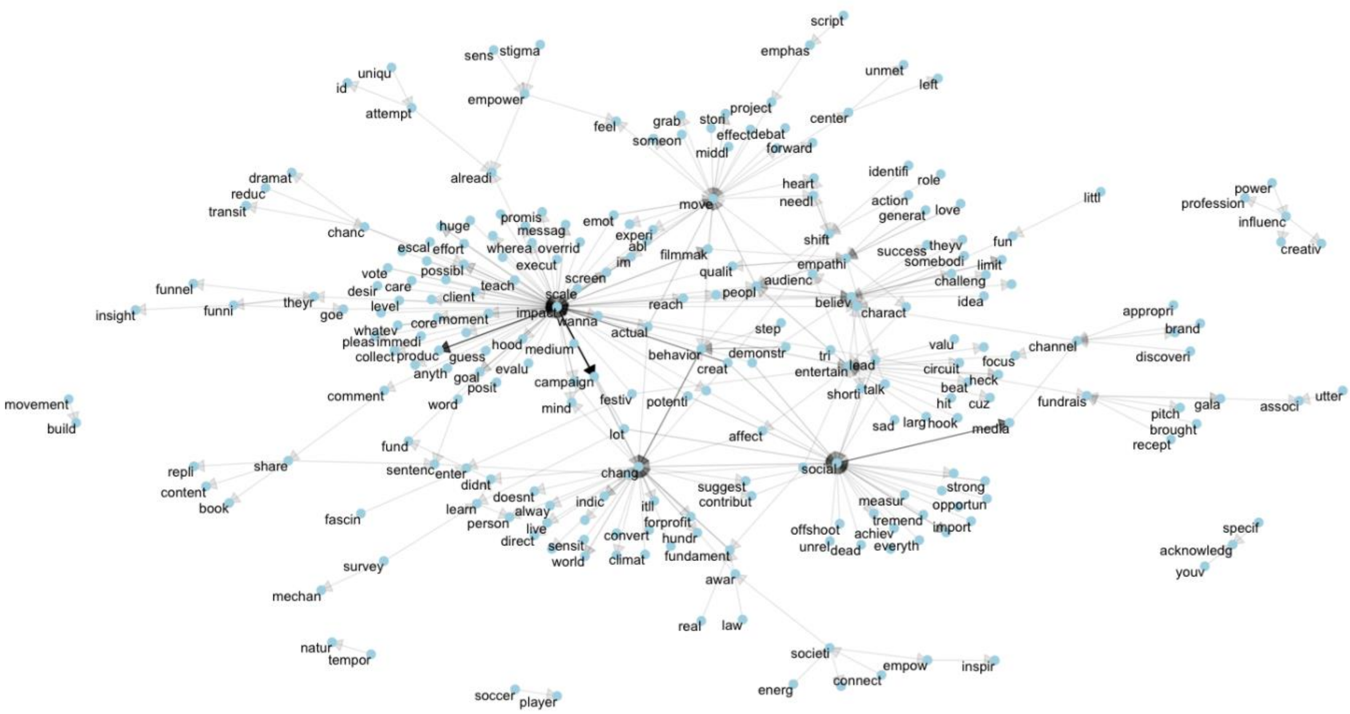


Figure 1. Association between words in the top occurring bigrams in the corpus. Blue circles denote the nodes, and the edges (lines with arrows) are in proportion to their relevance and frequency of co-occurrence with other words. Nodes with the most arrows indicate those that are the most commonly appearing words and have the widest range of associations with other words. The more distant a node (or bigram) is, the more exclusive they are in theme, as well as less frequent.

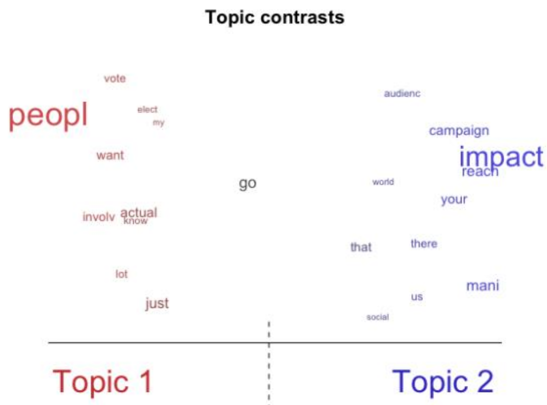
Topic modeling: The best fitting model indicated k=4 topics (residuals= 1.124, $p<0.001$, df=12,752). Table 2 provides the top words for each topic estimated by the LDA according to four different statistics: 1) highest probability (β), 2) “FREX”, 3) “Lift”, and 4) “Score.”

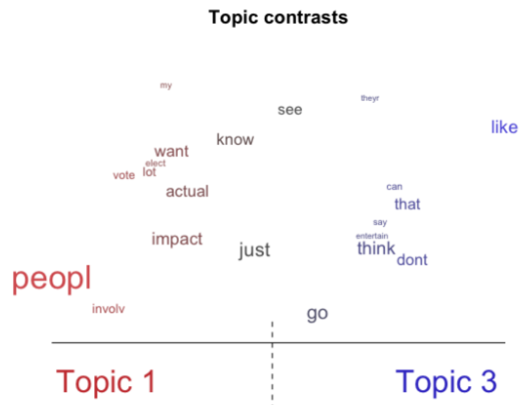
Table 2. Top words for the topics identified from the LDA model.

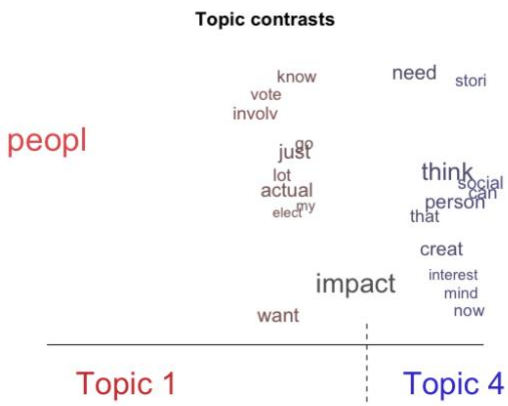
Topic 1	Probability: peopl, impact, involv, want, just, actual, vote, lot, know, go, elect, my, peac, feel, well FREX: involv, vote, elect, lot, march, want, my, peac, realiz, hood, actual, peopl, heard, cours, know Lift: involv, march, vote, cours, elect, realiz, hood, heard, peac, my, lot, came, measur, turn, somebodi Score: cours, vote, elect, involv, realiz, march, hood, peac, my, somebodi, heard, sure, came, peopl, guess
Topic 2	Probability: impact, peopl, mani, reach, campaign, your, there, us, audienc, that, world, social, produc, go, tri FREX: reach, your, mani, campaign, audienc, weve, produc, us, there, possibl, nonprofit, abl, potenti, thousand, impact Lift: your, reach, nonprofit, weve, thousand, possibl, campaign, mani, potenti, produc, media, abl, audienc, partner, hope Score: your, reach, campaign, possibl, mani, nonprofit, festiv, there, weve, audienc, thousand, abl, us, sure, potenti
Topic 3	Probability: peopl, like, think, dont, that, go, just, can, see, say, entertain, theyr, thing, said, know FREX: dont, like, entertain, theyr, love, part, said, ive, oh, say, talk, see, watch, power, festiv Lift: laugh, ive, love, part, reason, entertain, cuz, oh, theyr, talk, said, dont, doesnt, seen, someon Score: laugh, entertain, theyr, festiv, said, around, say, show, everi, start, power, reason, ive, love, can
Topic 4	Probability: think, can, social, person, stori, impact, now, creat, mind, need, interest, that, chang, way, import FREX: stori, mind, now, person, interest, immedi, creat, identifi, can, follow, ask, chang, need, least, right Lift: specif, identifi, immedi, mind, least, ask, interest, stori, follow, now, yes, take, person, creat, question Score: specif, identifi, immedi, now, follow, stori, ask, mind, can, tell, social, goal, say, interest, chang

Panel figure 2 depicts the bivariate comparisons (topic contrasts) between the 4 topics based on the marginal distribution of words they are represented by. Darker colors indicating higher probability of a word given a topic. The distance between the words is proportional to the uniqueness of the word to the given topic.

Ipek Ensari, PhD
Entertainment to Affect Change (E2AC) Research Team
Nick Hudson, Michael Gabriel, Bene Cipolla, Sunita Uthra
August, 2023







References

1. Silge J, Robinson D. tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *Journal of Open Source Software*. 2016;1:37. doi: doi:10.21105/joss.00037
2. Chang J. *lda: Collapsed Gibbs Sampling Methods for Topic Models; 1.4.2.* 2015.
3. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, Albertson B, Rand DG. Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science*. 2014;58:1064-1082. doi: https://doi.org/10.1111/ajps.12103
4. Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, Matsuo A, Lowe W, Council ER. *quanteda: Quantitative Analysis of Textual Data; 3.3.1*. 2023.
5. Grün B, Hornik K. topicmodels: An R Package for Fitting Topic Models. *Journal of Statistical Software*. 2011;40:1 - 30. doi: 10.18637/jss.v040.i13
6. Grant MJ, Button CM, Snook B. An Evaluation of Interrater Reliability Measures on Binary Tasks Using d-Prime. *Appl Psychol Meas*. 2017;41:264-276. doi: 10.1177/0146621616684584
7. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res*. 2003;3:993–1022.
8. Silge J, Robinson D. *Text Mining with R: A Tidy Approach*. O'Reilly Media; 2017.
9. Rijcken E, Mosteiro P, Zervanou K, Spruit M, Scheepers F, Kaymak U. FuzzyTM: a Software Package for Fuzzy Topic Modeling. Paper/Poster presented at: 2022 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE); 18-23 July 2022, 2022;
10. Nanda G, A. Douglas K, R. Waller D, E. Merzdorf H, Goldwasser D. Analyzing Large Collections of Open-Ended Feedback From MOOC Learners Using LDA Topic Modeling and Qualitative Analysis. *IEEE transactions on learning technologies*. 2021;14:146-160. doi: 10.1109/TLT.2021.3064798
11. Röder M, Both A, Hinneburg A. Exploring the Space of Topic Coherence Measures. In: *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*. Shanghai, China: Association for Computing Machinery; 2015:399–408.