"Analyzing textual data from open-ended survey questions"

Gerard Chung,
Social Service Research Centre

Survey on parents during Circuit-breaker

What has helped you to cope with parenting in the past weeks?

_____

_____

_____

_____

_____

https://sgcovidparenting.mystrikingly.com/

Impact of savings programs on youths' money management

F.16.   If you wanted to buy something valuable but you did not have enough money, what would you do to set aside enough money to buy it later?

https://gsdi.unc.edu/initiatives/youthsave-ghana/

Benefits of open-ended survey questions

- Capture respondents' experiences in their own words

- Gather new information about a topic

- Supplement quantitative findings

Closed-ended questions is more prevalent

- These questions offer a limited set of response options.

- "easier to ask, code, and analyze"[1]

1. Geer,1991, p. 360

It is difficult with large "messy" text data

[1] "Asking his daughter to sign up for one of the options tells the daughter his dad does not truly understand her hobbies and interest. Instead he should say \" Summer lets get you into tennis which I know you like playing with your friends Marcy and Kathy, you seem to enjoy the work out and we can find a league in town that can challenge your skills.\" or he can say \" Well summer you play the guitar well, mom and I would be excited to invest into lessons that work with your homework schedule, giving you time with friends but fills that gap of feeling bored.\""
[2] "Play outside. Give them room to be kids if they have no reason to be watched like hawks, then dont."
[3] "He showed his son they have something in common."
[4] "She showed empathy"
[5] "She is encouraging her child to be social and active. I am assuming she lives in an area where she know her neighbors and would know where Nicole is going."

Credits: Chris Hemsworth as the title character in Thor: The Dark World.
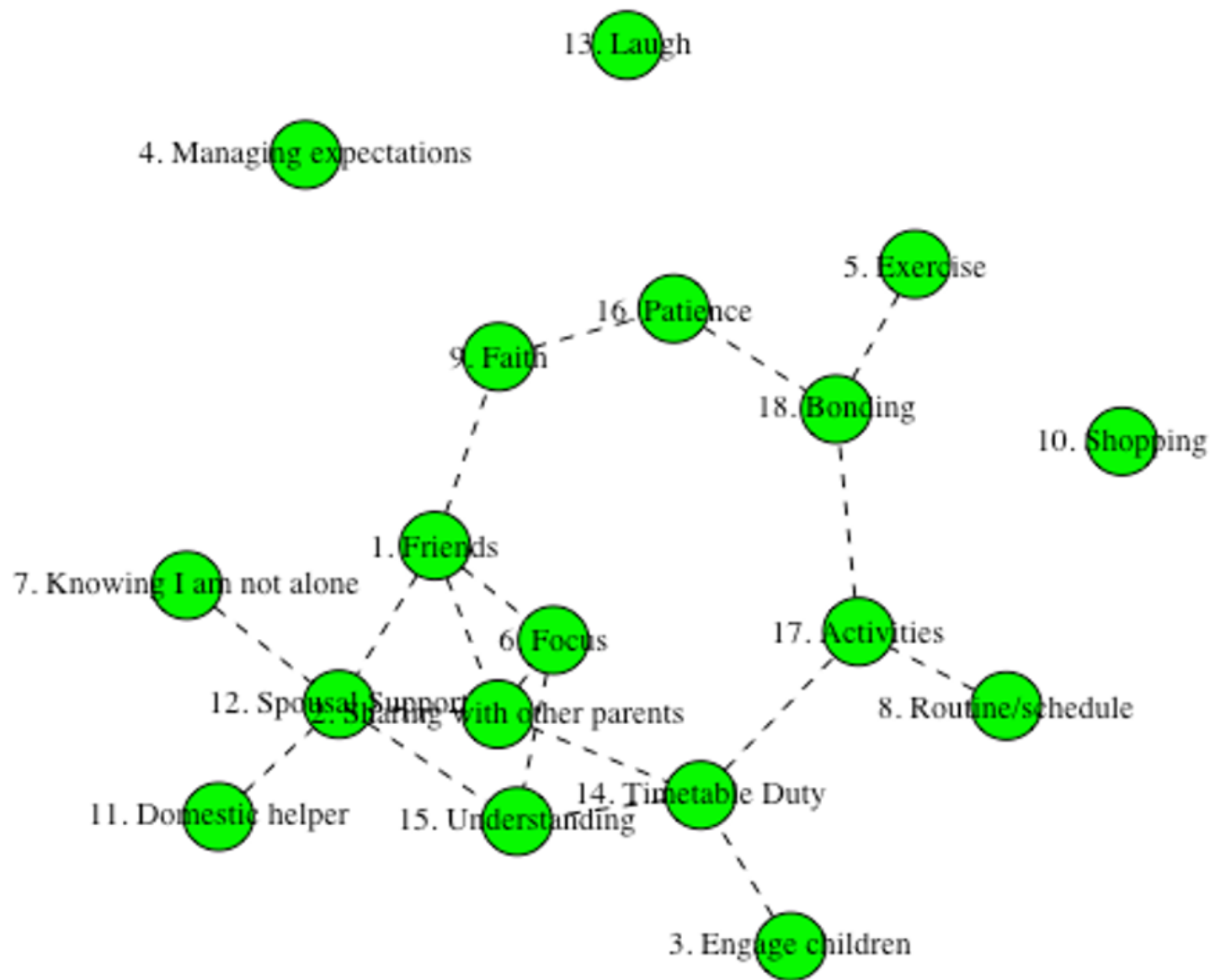Photograph: Allstar/Marvel/Sportsphoto Ltd

Topic modeling

- A computational method to extract knowledge from large amounts of textual data
    - inductively identifies latent topics among a large collection of documents and ...
    - assigns individual documents to one or more topical categories.
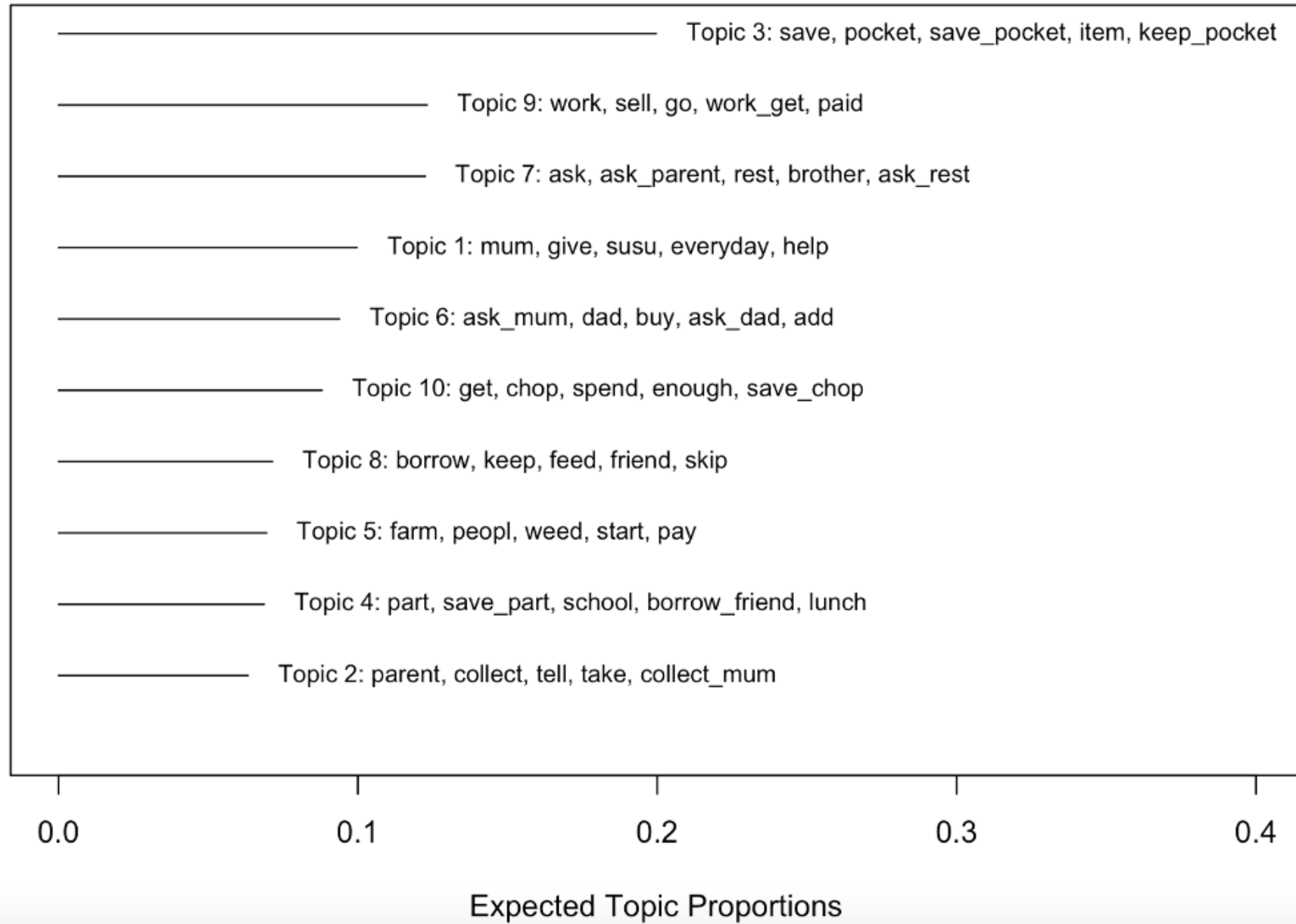
"What has help you cope with parenting?"

Figure S2. This figure offers a network analysis of topic correlations from the 18-topic model

".. what would you do to set aside enough money to buy it later?"

Top Topics

Topic 3: save, pocket, save_pocket, item, keep_pocket

Topic 9: work, sell, go, work_get, paid

Topic 7: ask, ask_parent, rest, brother, ask_rest

Topic 1: mum, give, susu, everyday, help

Topic 6: ask_mum, dad, buy, ask_dad, add

Topic 10: get, chop, spend, enough, save_chop

Topic 8: borrow, keep, feed, friend, skip

Topic 5: farm, peopl, weed, start, pay

Topic 4: part, save_part, school, borrow_friend, lunch

Topic 2: parent, collect, tell, take, collect_mum

0.0   0.1   0.2   0.3   0.4

Expected Topic Proportions

Challenges

- Potentially useful text data but hard to access

  - case notes, risk assessments, social reports etc.

- Short text challenge

  - severe data sparsity

  - biterm topic model

- Working with qualitative approaches

  - computational grounded theory framework (Nelson, 2020)

  - 3-step: pattern detection -> refinement -> confirmation

- Corpus linguistics methods

Understanding the experiences of social work clients remains a hallmark goal of social work research

Find me at:

🐦 **@GerardChung**
🐙 **@gerardchung**
🔗 **gerardchung.com**

References

- Geer, J. G. (1991). Do open-ended questions measure "salient" issues? Public Opinion Quarterly, 55(3), 360–370. https://doi.org/10.1086/269268

- Thyer, B. A. (2012). The scientific value of qualitative research for social work. Qualitative Social Work, 11(2), 115–125. https://doi.org/10.1177/1473325011433928

Understanding the EXPERIENCES OF SOCIAL WORK CLIENTS REMAINS A HALLMARK GOAL OF SOCIAL WORK RESEARCH

Find me at:

🐦 **@GerardChung**

💻 **@gerardchung**

🔗 **gerardchung.com**

Error   ×