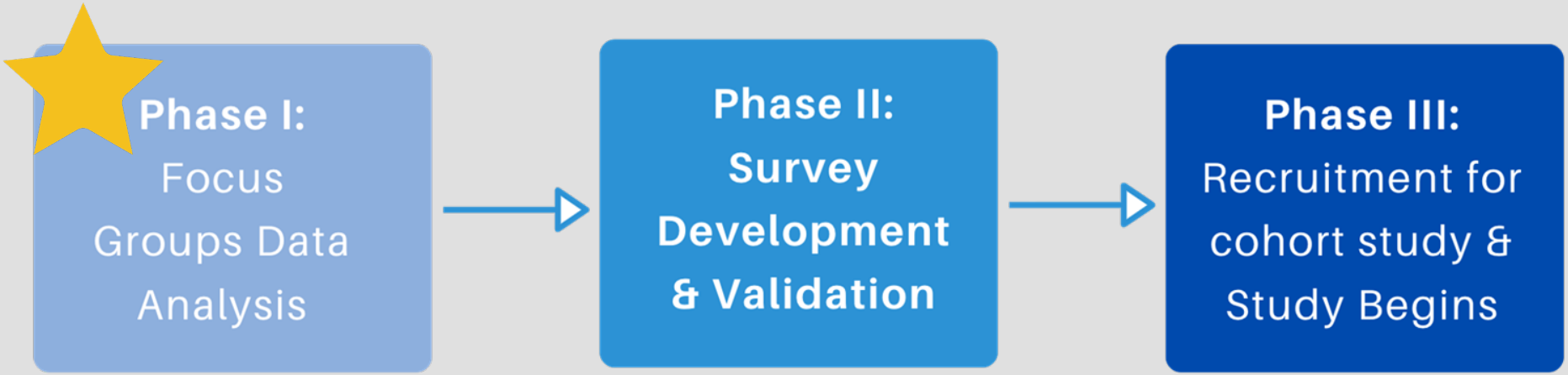
Qualitative Survey Methods: Interrater Reliability

Audrey Burnette[1]
Lauren Wisnieski[2]
Debusk College of Osteopathic Medicine, Lincoln Memorial University, Harrogate, TN 37752[1]
Richard A. Gillespie College of Veterinary Medicine, Lincoln Memorial University, Harrogate, TN 37752[2]

The Retention of Equine Veterinarians Longitudinal (REVL) study



Focus Groups

Vet Students Interested in Equine Practice

Current Equine Vets

Equine Vets That Have Left Equine Practice

Developing the codebook by integrating current knowledge on Equine Veterinarian burnout using Conservation of Resources theory to evaluate the Focus Groups:

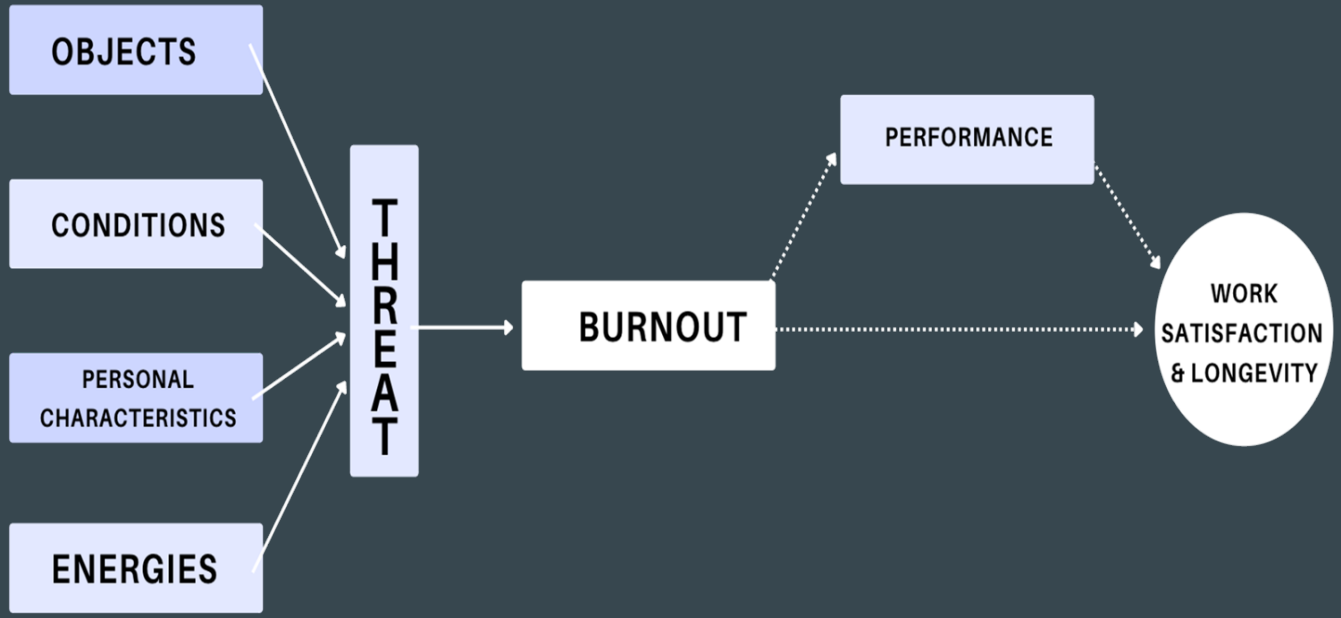
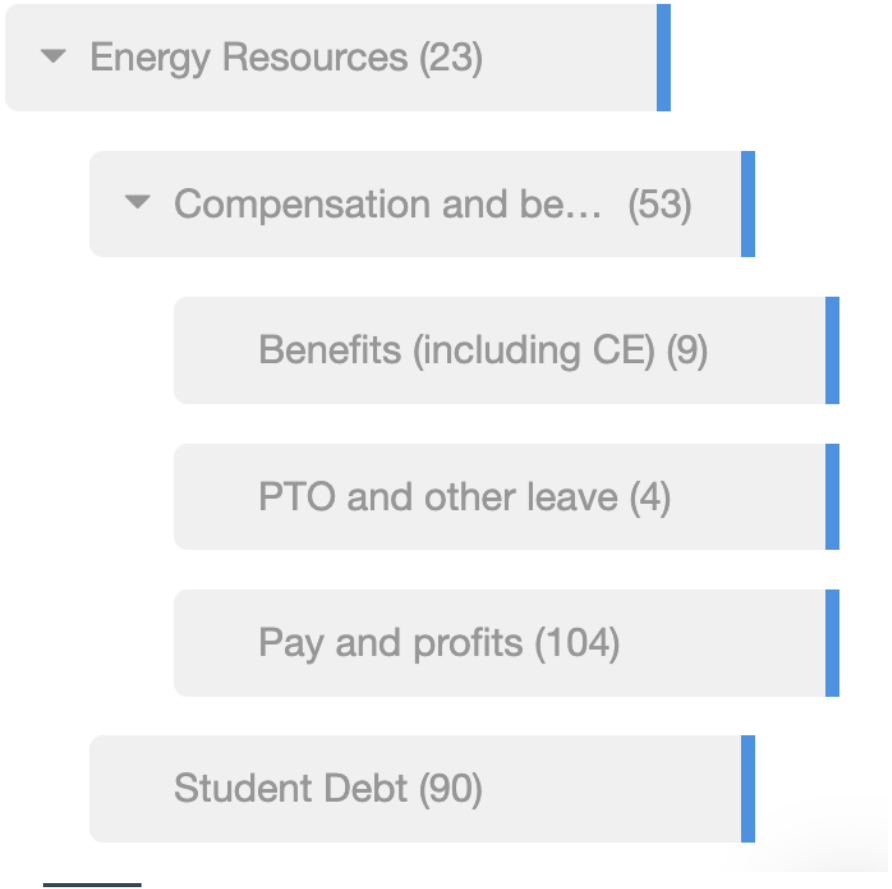


Figure 2: Resources of the conservation of resources theory

Using COR theory to build coding framework



Example of Delve coding:

FG 1-A Coder 2

Transcript and Codes | Participant Information

Vet 1-3: So, I do love working for myself. I do like the flexibility it allows me, but there was actually a place when I was in vet school in Temecula where five doctors, I think it was, came together bought a house so that they each had like a room for their office. And they shared paying for one receptionist and one person that did billing, and I thought that was the coolest thing ever. If the people around here had gone for that, I would definitely have but coming in, as the only young vet, I was pretty much shunned. So my perfect practice is independent but have camaraderie with the other independents around.

Coworkers | Respect, teamwork, and feeling valued | Teamwork | Social Support | Supervisory support and mentorship | Coworkers and/or supervisor/boss | Work Environment

Why do we need Interrater Reliability?

- To ensure the accuracy and quality of the coders.
- To further refine the codebook.

Interrater Reliability Best Practices:

- Appropriate # of Coders
- Independence of each coder and Interaction between Coders
- Data segmentation
- How results should be acted on
- Re-iterative Process

Cronbach's Alpha: statistical measure of **inter-rater reliability** for categorical variables

0 — 1

Closer to **zero** indicates that there is no correlation between the items at all. They are entirely independent and raters are not in agreement.

Cronbach's Alpha Cutoff : 0.6 or higher

Closer to **one** indicates that there is perfect correlation or perfect agreement between raters.

Checking For Interrater Reliability

Category	Subcategories	alpha	times used
Work-life balance	Time for self-care	0.63	12
	Time for family/friends	0.66	25
	Time for their own horses	0.95	8

Using **STATA** we were able to calculate Cronbach's Alpha among 3 individual coders. 10% of data from each focus group is being used for reliability checks.

The Evolving Codebook

From ~100 codes to 35 codes

- Achieving satisfactory reliability should work with an upper limit of 30–40 codes (MacQueen et al. 1998)
- In order to narrow down codes that were most used, codes that proved to be unused or repetitive were combined under a higher-order code category.

Most Used Codes

- Pay
- On-call
- Work Atmosphere
- Sharing workload with other Vets
- Mentorship
- Time for Family/Friends

Unused Codes

- Electronic Medical Records (vs paper)
- Dictation Devices
- Administrative Tasks
- Collecting money from clients
- Travel to farms
- Business Management Skills

Most Important Codes: Besides from "Pay" it seems that the most used codes with the highest reliability cronbach's alpha score fall in the **condition resources** category of stressors.

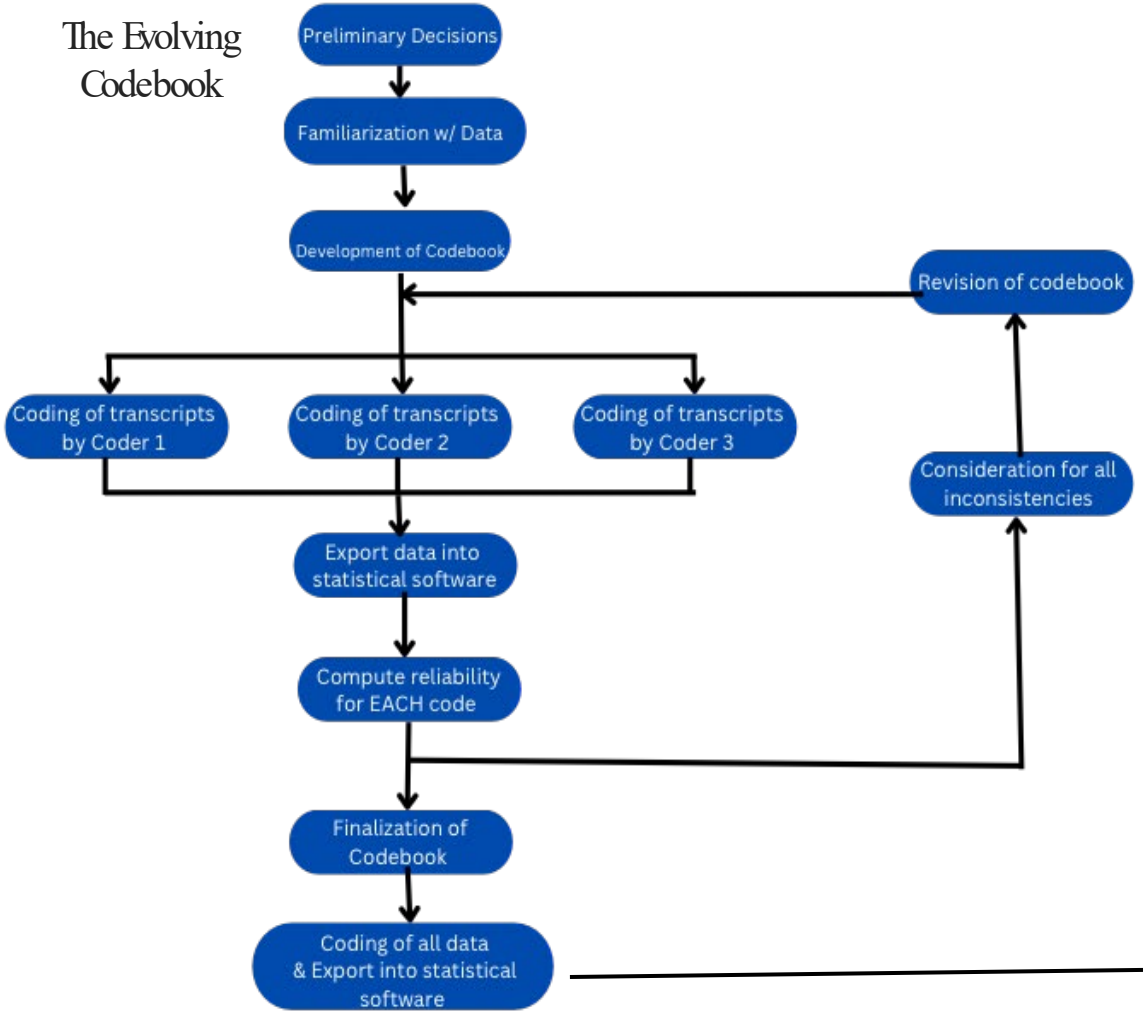
Early Codebook

Category	Subcategories	alpha	times used
Work-life balance	Time for self-care	0.63	12
	Time for family/friends	0.66	25
	Time for their own horses	0.95	8

Newest Codebook

Category
Work-life balance

The Evolving Codebook



Phase 1:

As we continue data analysis, we will continue **rechecking our interrater reliability** and iteratively **editing the codebook** until we have it finalized for the next step of the study.

Phase 2: Survey Development

Conclusions

- Codebook reliability is important for qualitative research because it allows us to accurately derive meaning behind focus group interviews and allows for quality interpretation of data.
- Reliability was checked with 3 independent coders. Cronbach's alpha and how many times each code was used was calculated..
- Most codes were related to conditions resources, such as workload, work atmosphere, time for family/friends, etc.
- Reduced codebook from ~100 codes to ~35 codes so far.
- Iterative Process! We will be going through more rounds of reliability checks before coding the rest of our focus group transcripts.

Resources

- O'Connor, C., & Joffe, H. (2013). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, Volume 19: 1-13, https://journals.sagepub.com/home/ijq
- MacQueen, K. M., McLellan, E., Kay, K., & Milstein, B. (1998). Codebook development for team-based qualitative analysis. *CAM Journal,* 10, 31-36. https://doi.org/10.1177/1525822X980100020301

Acknowledgment

- We would like to thank the individuals that participated in the focus groups and our funding sponsor, Boehringer Ingelheim.