A Dynamic Systems Approach to Patterns of Affect and Cognitive Difficulty in Interviewer-respondent Interactions

Matt Jans[1]
[1]UCLA Center for Health Policy Research,
10960 Wilshire Blvd, Ste 1550, Los Angeles, CA 90024

Abstract
Interviewer-respondent (I-R) interactions are dynamic phenomena at their core, making dynamic systems (DS) theory an obvious framework for their study, yet DS theory and methods are rarely applied to I-R interactions. This study uses utterance-level coded data from audio recordings of phone interviews conducted by the Reuters/University of Michigan Surveys of Consumers (SCA). Recordings were transcribed into interviewer and respondent utterances, and rater judgments of affect were applied. Using GridWare v1.15a, a DS theory-based software (www.statespacegrids.com), trajectory plotting, attractor states identification, dynamic systems parameter estimation are demonstrated. The paper also motivates a graphical and intuitive understanding of cross-utterance I-R dynamics using the software. Initial results show patterns of affect dynamics varying by responses to income, interviewer gender, and across interviews conducted by the same interviewer. Potential for DS theory and methods to inform and guide the study of I-R interactions is briefly discussed. This research was sponsored by the Charles Cannell fund in Survey Methodology and the U.S. Census Bureau Dissertation Fellowship.

Key Words: interviewer-respondent interaction, dynamic systems, survey interviews

1. Introduction

Interactions between interviewers and respondents have long been conceptualized as dynamic (Kahn & Cannell, 1957), but very little research actually operationalizes and analyzes these interactions as dynamic systems (see Elzinga, Hoogendoorn, & Dijkstra, 2007 for an exception). Most quantitative research on interviewer-respondent (I-R) interaction summarizes behaviors of one or both participants, often at the question level (Cannell, Lawson, & Hausser, 1975), ignoring interaction between the I and R in the analysis. Much of the I-R research that has an explicit dynamic approach is qualitative (Maynard & Schaeffer, 2002; Schaeffer, 1991).

However, fields outside survey methodology have readily adopted dynamic systems (DS) methods for studying interpersonal interactions. One such approach is state space grids (Granic & Hollenstein, 2003; Granic, Hollenstein, Dishion, & Patterson, 2003; Hollenstein, 2007, 2013; Hollenstein, Tom, 2003; Lewis, Zimmerman, Hollenstein, & Lamey, 2004). The approach offers a visual representation of the interviewer-respondent dynamic in "state space", and the application of dynamic systems analysis techniques, bringing our analyses of I-R interactions more in-line with our conceptualizations of the system. This paper is an exploratory “proof-of-concept” application of state-space-grids to I-R interactions. It addresses several exploratory and empirical research questions:

1) Can dynamic systems concepts meaningfully describe I-R interactions?

- Do we get more from this approach, conceptually or empirically, than other approaches?

2) Are there differences in dynamic systems phenomena between different interviewer or respondent types?

1.1 Core Dynamic Systems (DS) Concepts

The primary dynamic systems concept motivating this paper is "state space", which is the range of all possible states of the system (Hollenstein, 2013). In this context, that is the affective states that an interviewer and respondent can occupy at measurement points over conversational turns on an income question. Figure 1 is an example of this state space, with the respondent's affect on the *y* axis and the interviewer's affect on the *x* axis. Measurements on which the respondent and interviewer both show positive affect would appear in the upper right corner. When both are negative, the observation is plotted in the lower left corner. Observations on which the respondent was positive, but the interviewer was negative would appear in the upper left, and those with negative respondent and positive interviewer affect would appear in the lower right.

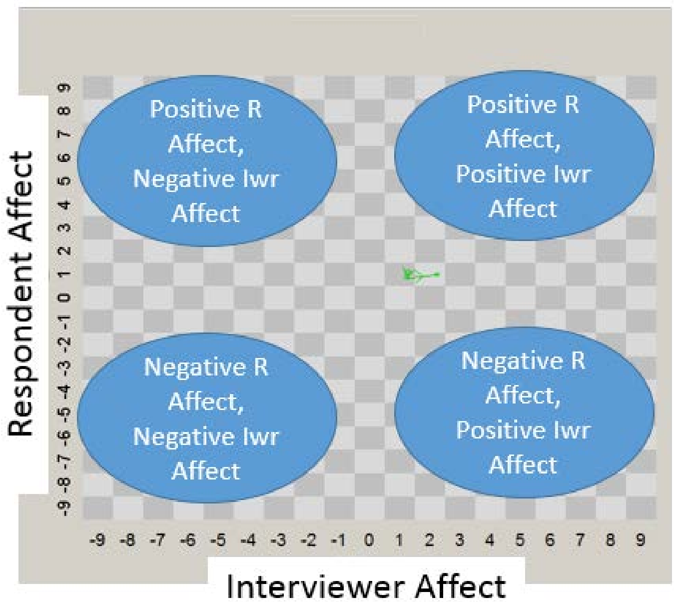

Figure 1: State space of interviewer and respondent affect across utterances

Each cell in the state space grid is referred to as a *state*, and the plotted points are referred to as *events*. Figure 1 shows three events from one I-R exchange in green (one at 1,1; a second at 1,1, and a third at 1,2). A series of events is a *trajectory*. If there is no change in the affective tone of the interaction, the trajectory would fall entirely in one cell. If there is completely random change over events, no discernable patterns will be seen. However, one of the goals of DS research is identifying *attractor states*, which are areas of the state space or cyclical patterns in trajectories toward which the system is repeatedly drawn over time. Other DS metrics can also describe movement of the system. For example, *dispersion* summarizes how much of the state space the system covers. Analysis of *visits to regions* of the state space is another way DS research searches for attractor states. For example, if the I-R trajectory is continually drawn into the region of state space where both interviewer and respondent are negative, that could be considered an attractor state. If that pattern correlates with how income questions are answered across trajectories, it would demonstrate something systematic about the relationship between I-R interactions and response. These DS concepts will be the focus of this paper.

2. Methods

2.1 Data

Interviewer-respondent exchanges on the income question from n = 170 interviews in the Reuters/University of Michigan Surveys of Consumers (SCA). This survey digitally records interviews for quality control purposes unless the respondent objects to being recorded. Recordings were selected to roughly balance the sample between income nonrespondents (i.e., those who provided no income information, n = 56), income bracketed respondents (i.e., those who provided income information in brackets only, n = 55), and those who provided an open-ended dollar amount income response (n = 59). The SCA income question read,

> "To get a picture of people's financial situation we need to know the general range of income of all people we interview. Now, thinking about [your/your family's] total income from all sources (including your job), how much did [you/your family] receive in the previous year?"

Respondents could answer with a dollar amount, refuse to answer, or say "I don't know." Those who refused or said "don't know" were read a series of unfolding brackets starting with "Was your income above or below $50,000?" If the respondent said "don't know" or refused again, that was their final income response (i.e., income nonrespondents in this study). If the respondent said "above," the interviewer continued to ask unfolding bracket questions in $10,000 steps until the appropriate category was reached or the respondent would answer no further. If the respondent said "below", the interviewer asked "Was it above $10,000" and worked up to $50,000 in $10,000 steps (see Yan, Curtin, & Jans, 2010 for more details about this question). Respondents who answered any level of bracketing were called "bracket respondents" in this study. This resulted in three final income response types: Dollar amount response, bracketed income response, and income nonresponse.

Undergraduate students coded interviewer and respondent affect (i.e., mood) valence and intensity at each utterance. Valence was coded as positive, negative, or neutral, and intensity was coded on a 9-point scale (1 = low intensity to 9 = high intensity). These measures were combined to create the state space dimensions for the current analysis. This created a 19x19 cell state space grid which was analysed in the GridWare v1.15a[1] software.

2.1.1 Dynamic systems trajectory plots and measures

GridWare allows for plotting trajectories and examining their patterns qualitatively (e.g., Figure 1). It also calculates common dynamic systems metrics, which can be output and analyzed in standard software. For this exploration, *whole-dispersion*, and *visits to regions* of the state space were calculated and analyzed (see Hollenstein, 2013 for deeper discussion of these concepts and their measurement).

[1] http://130.15.96.140/SSG/

Whole-grid dispersion is essentially a measure of how broadly each trajectory spans the available state space. It is calculated as,

$$\mathbf{1} - \frac{\left(\boldsymbol{n}\sum(\boldsymbol{d_i}/\boldsymbol{D})^{\mathbf{2}}\right) - \mathbf{1}}{\boldsymbol{n} - \mathbf{1}}$$

where D is the total duration, d_i is the duration in cell (across events), n is the total number of cells occupied. The measure ranges from 0 (no dispersion) to 1 (maximum dispersion). In this analysis, there is no duration measure for each event (i.e., we only know whether and how many times a cell was visited, but not for how long), so dispersion is based on the count of events instead of duration.

Calculating the number and proportion of visits to regions of the state space involved highlighting regions of interest on the grid. In this analysis, regions of interest were defined as those where there was mutually positive affect (upper right), mutually negative affect (lower right), positive respondent affect, but negative interviewer affect (upper left), and positive interviewer affect, but negative respondent affect (lower right). Figure 2. highlights those areas in yellow. Regions where either one of the speakers was neutral were excluded, so that the regions defined had clear affective tone.

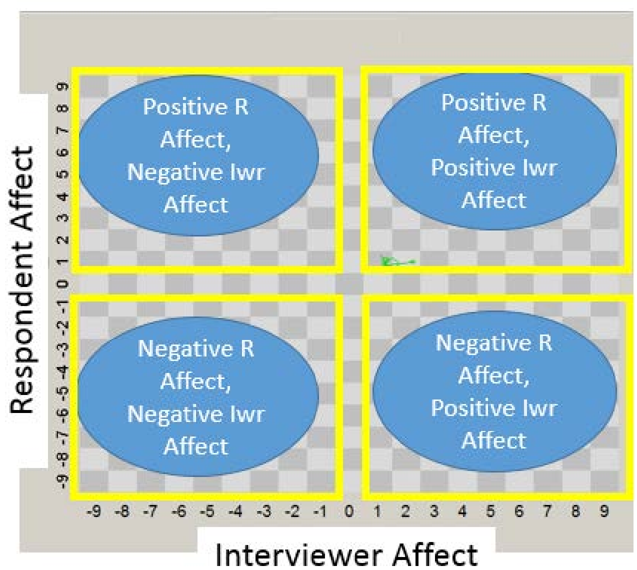

Figure 2: Definition of state space regions of interest for “region visit” analysis

3. Results

3.1 Assessing I-R Interactions Qualitatively

One benefit of GridWare is the visual representation of trajectories for each I-R pair. Figures 3 and 4 each show the trajectory of an I-R pair across utterances on the income question. Because the state space is large, zoomed views of the trajectories are provided. The pattern of the trajectories revising states across events may be signs of attractor states. In all plots, red denotes income nonrespondents, blue denotes bracketed respondents, and green denotes dollar amount respondents.

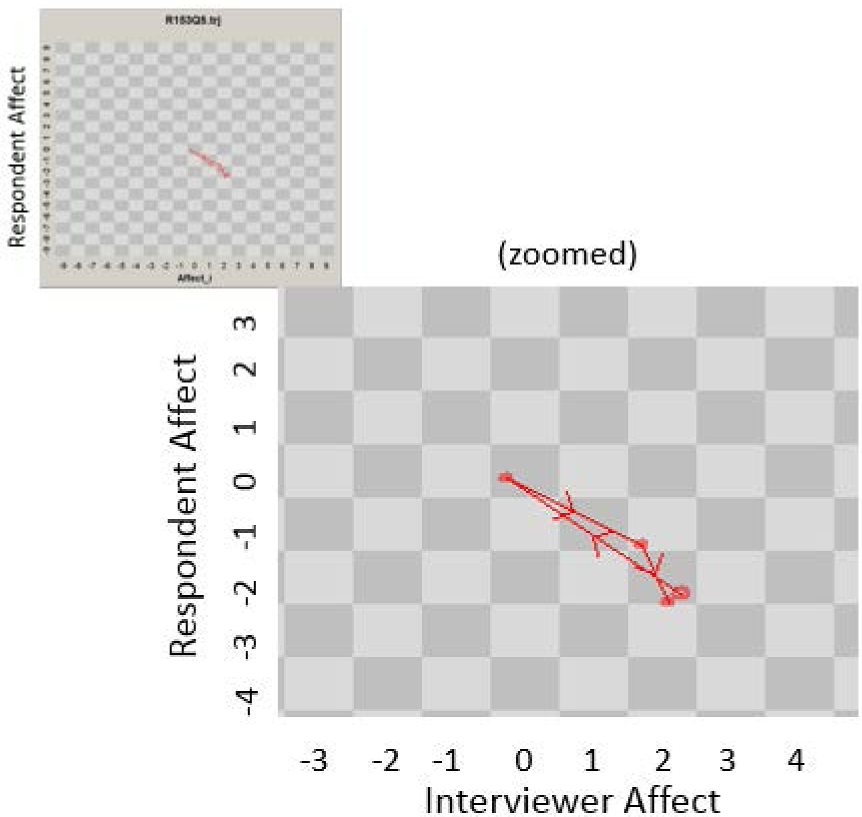

Figure 3: One I-R trajectory that revisited states (possible attractor at -2,2)

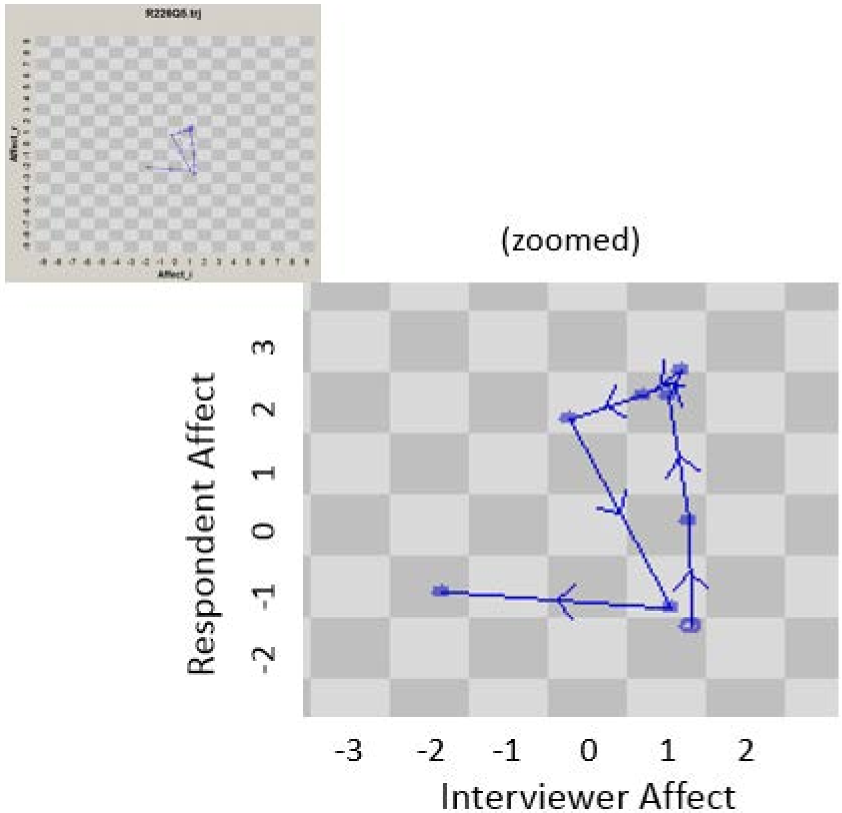

Figure 4: Another I-R trajectory that revisited states (possible attractor at -1,1 or -2,1)

Figure 5 shows variability across interactions within one interviewer. It is interesting to see how different interactions can be across respondents interviewed by the same interviewer. In trajectory #2, there maybe be signs of an attractor state around -3,-2 to -3,-1 and at 1,-2. Attractors that revisit states or regions cyclically are called *cyclical attractors* (Hollenstein, 2013).

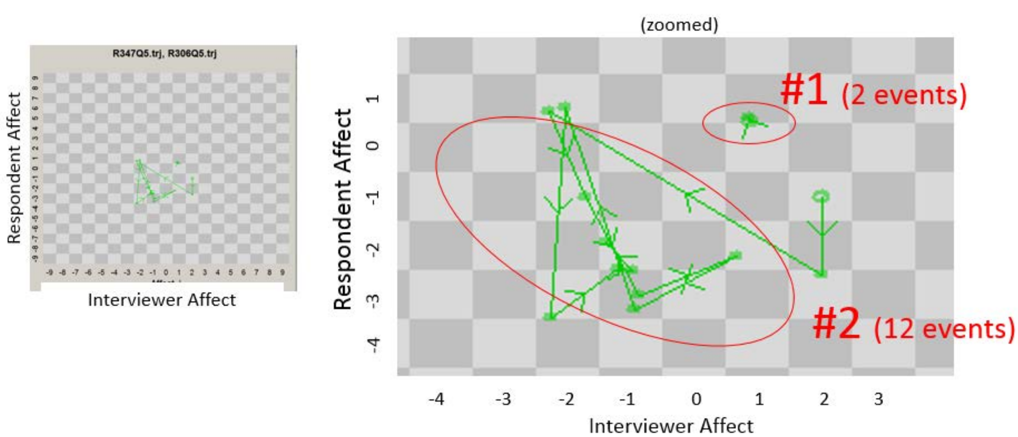

Figure 5: Two dollar amount respondents from one interviewer

Figure 6 shows another interviewer and the variability of their interactions across 7 respondents.. As with above, green trajectories denote dollar amount respondents, blue denote bracketed respondents, and red denote nonrespondents.

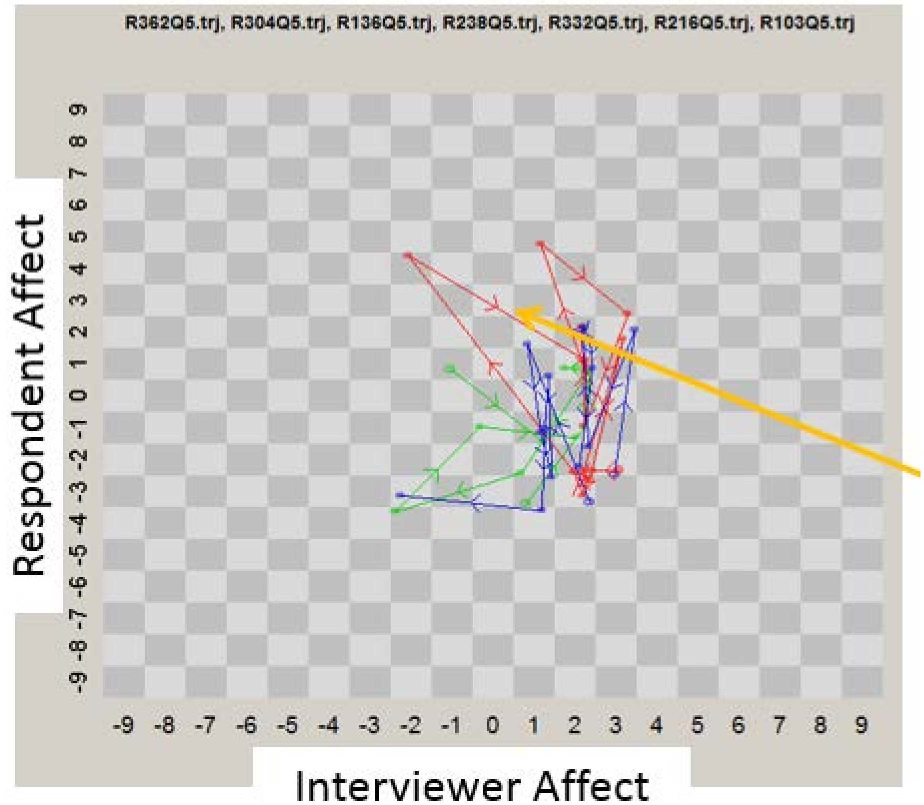

Figure 6: Variability in trajectories and income response outcomes within one interviewer (orange arrow points out that the trajectory with the most positive respondent affect, but negative interview affect, ended up as income nonresponse)

3.2 Assessing I-R Interactions Quantitatively

In this section, state space plots are paired with traditional statistical analyses to provide a more complete picture of the I-R dynamics.

3.1.1. Whole-grid dispersion

Figures 7 and 8 display the same trajectories. Figure 8 plots the average whole-grid dispersion of the trajectories displayed in Figure 7. Bracketed respondents had a significantly higher average dispersion than income nonrespondents and dollar amount respondents. This may be due, in part, to the additional events involved in asking and answering the unfolding brackets. Referring to Figure 7 shows that the variability in bracketed respondent dispersion is more along the respondent affect dimension than the interviewer affect dimension.

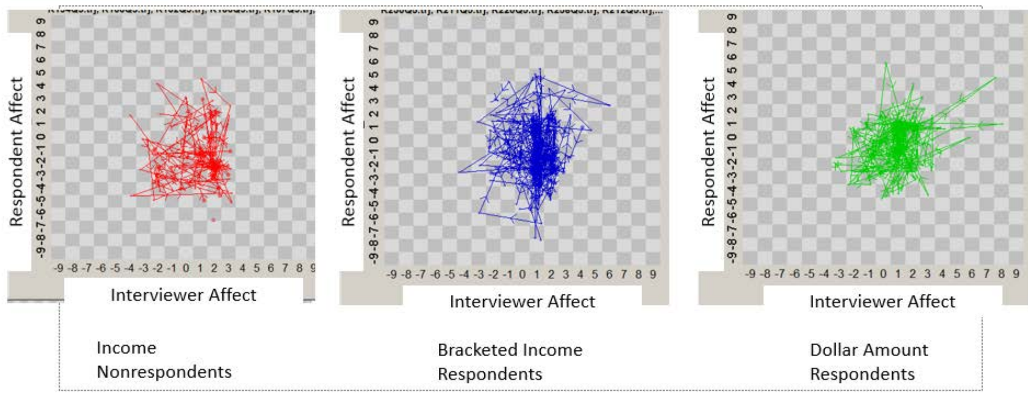

Figure 7: Relative variability of trajectories by income response.

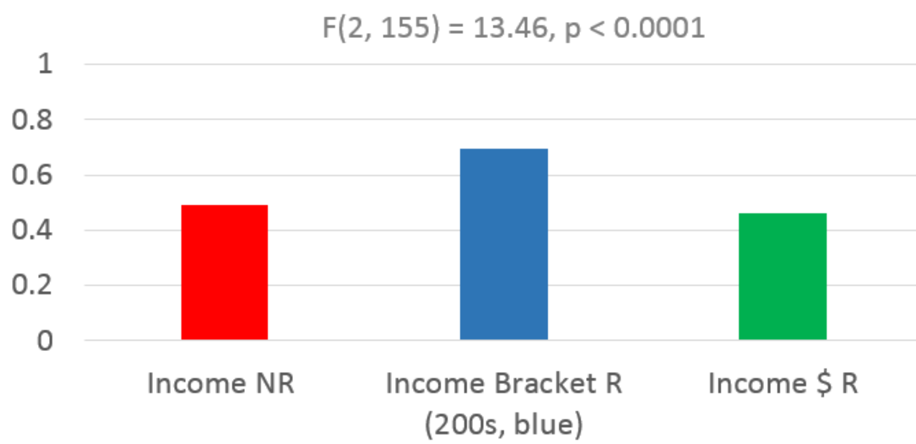

Figure 8: Mean differences in whole-grid dispersion of trajectory groupings in Figure 6.

Figures 9 and 10 show that interviews conducted by female interviewers had marginally-higher dispersion than those conducted by men.

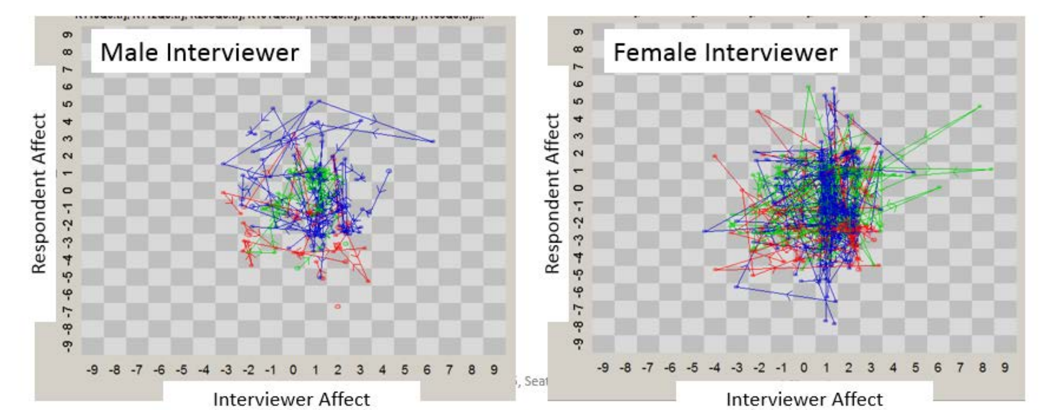

Figure 9: Trajectory plots of interviews conducted by male interviewers and female interviewers.

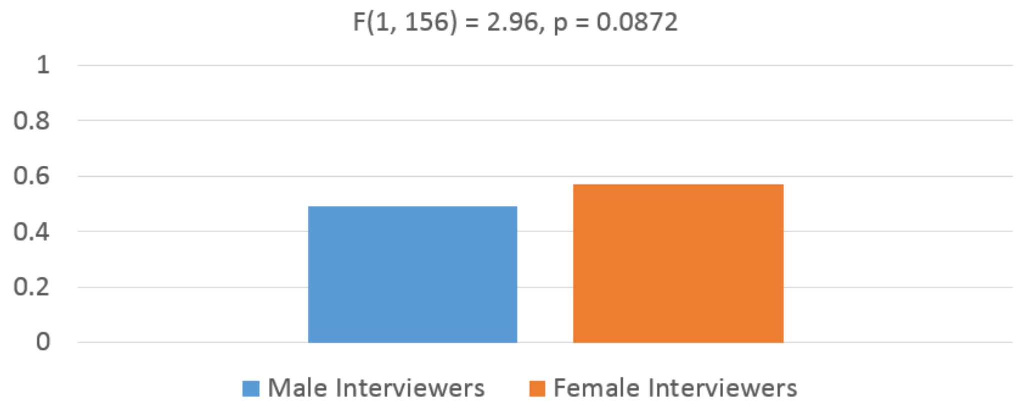

Figure 10: Mean differences in whole-grid dispersion for interviews conducted by male interviewers and female interviewers.

3.1.2 Region assessment

Figure 11 plots the mean visits to regions of interest by income response type. It shows that there is variability in how frequently trajectories of each income respondent type visit the respective regions of the state space. This plot is particularly exploratory, so significance tests are not shown. Comparing the average number of visits of each type of respondent (within each defined region), dollar amount respondents have the highest average visits to the region where both I and R are positive (I Pos, R Pos). Bracketed respondents follow, then income nonrespondents. Referring to Figure 7 helps confirm this difference. While the grouped trajectory plots in Figure 7 all seem roughly centered in the middle of the state space, the dollar amount income respondents are shifted slightly right and up of center, and have a few trajectories that enter far into the positive-positive region on one or more events, a pattern not seen in the other to income respondent types. This suggests that income dollar amount interactions are characterized by positive affect by both speakers, on average, despite wide variability across respondent trajectories. There does not seem to be a difference in the average number of visits to the “negative, negative” region, suggesting that income nonresponse and bracketed response are not due to clearly negatively-toned interactions. Moving to the two off-diagonal regions, the most notable difference is that bracketed respondents, relative to dollar amount and nonrespondents, have higher average number of visits to regions where the interviewer was positive, but the respondent was negative, or where the interviewer was negative and respondent was positive. This suggests that bracketed respondent trajectories may be spread diagonally (NW to SE) more than dollar amount of nonrespondent trajectories. This is not evident from a review of Figure 7. The finding may also be due, at least in part, to the longer length of bracketed respondent trajectories on average. However, even when averages are out of events in individual trajectories (i.e., controlled for trajectory length), longer trajectories have more opportunities to cover more state space. These differences deserve more exploration.

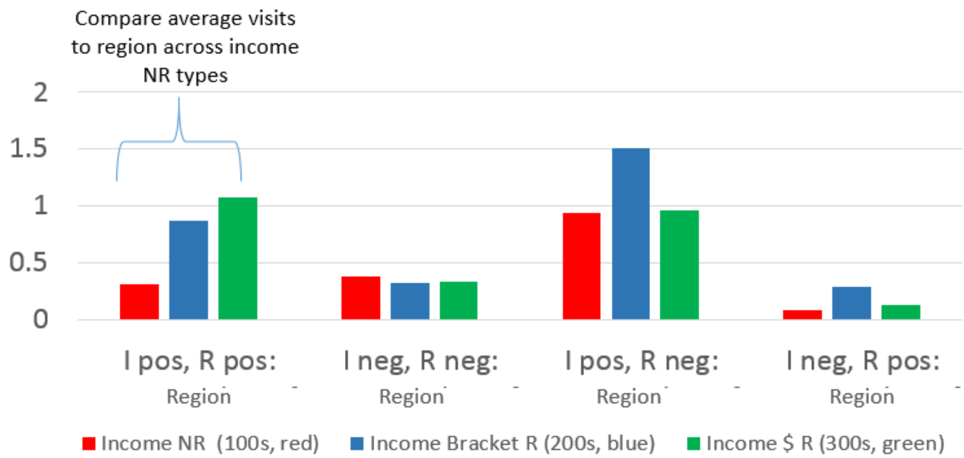

Figure 11: Mean differences in visits by region visited and income response type.

4. Discussion

Dynamic systems concepts, particularly state space, trajectories, events, dispersion, region visits, and attractor states seem to have a promising role in studying interviewer respondent interactions. Although the interactions explored here were not highly dynamic (i.e., they did not traverse the entire state space), that may be due to the definition of the state space, and the length of the trajectories. We did not need 9 level of positive and negative affect intensity to capture the mood of most interactions (although a few I-R pairs reached near the upper ends of the intensity scale, particularly in the positive direction for both speakers). Concerning length, trajectories plotted over several questions or an entire interview might show even wider dispersion, more attractor states, and distinguish between types of income respondents more clearly. Indeed, the model research on parent-child interactions (Granic & Hollenstein, 2003; Granic, I., Hollenstein, T., Dishion, T. J., & Patterson, G. R., 2003) use exchanges of 30 minutes or more, compared to the tens of seconds of interaction on a single survey question.

Substantively, the results appear to show that bracketed income respondents are have trajectories that are more dispersed on average than nonrespondents or dollar amount respondents, but this needs to be explored further to determine if the length of these trajectories contributes to that dispersion. Female interviewers had trajectories that were only marginally more dispersed than male interviewers, but gender may play a stronger role in other DS measures or on other questions (e.g., those with content related to sex and gender). Visual examination of grouped trajectory plots by interviewer gender appear to show more extreme positive interactions with female interviewers than males, but individual trajectories with female interviewers also reaches deeper into negative respondent and interviewer affect regions than those conducted by male interviewers.

Region assessment results are particularly difficult to interpret without clear statistical testing and visual display of grouped trajectory plots with target regions highlighted. These are clear next steps for this research. There is suggestive evidence that dollar amount respondent interactions may be more positive (in both interviewer and respondent affect) than bracketed amount or refusals, but those results may be due to a few extremely positive trajectories. This begs the question of how to treat outliers in trajectory analysis. Is a single trajectory that dramatically departs from the mass of the other trajectories removed, just like one might remove an outlying point in a scatter plot? Should the whole trajectory be removed, or just the outlying event? What value should be replaced to maintain the trajectory (e.g., the average of all other events for that trajectory)?

At minimum, state space grids and their related concepts help us define the dynamic phenomena we truly seek to study more clearly than other methods, and allow visual presentations of those systems (i.e., trajectory plots and grouped trajectory plots) that can produce insights about I-R interactions that would not be gained through standard statistical analyses or in-depth qualitative analyses. It is important to remember that there are other DS measures that can be calculated in GridWare, such as entropy and other ways to search for attractor states than those presented here.

It is also important to remember that dynamic systems theory, complexity theory, chaos theory, and the many names that categorize similar approaches represent a very wide range of statistical theory and technique, applied to a wide range of phenomena. While state space grids are an intuitive application for I-R interactions, other models may apply as well (e.g., Elzinga, Hoogendoorn, & Dijkstra, 2007). This exploration is offered as one attempt to match our analytic methods with our concepts and theory in the study of I-R interactions. It is not the only way to do it. There is certainly much more for survey methodology to learn and explore about dynamic systems modeling and visualization of interviewer-respondent interaction, but this family of approaches seem promising.

Acknowledgements

The author thanks the Charles F. Cannell Fund in Survey Methodology and U.S. Census Bureau Dissertation Fellowship for financial support for the coding. Maya Burns, Alex Dopp, Zach Hartley, Clare Levioki, Lian Liu, Miyuki Nishimura, Melinda Mosher, Travis Pashak, Katie Singer, and Jenna Stein were essential to producing the coded data for this study.

References

Cannell, C. F., Lawson, S. A., & Hausser, D. L. (1975). A technique for evaluating interviewer performance: a manual for coding and analyzing interviewer behavior from tape recordings of household interviews. Survey Research Center, Institute for Social Research, University of Michigan.

Elzinga, C. H., Hoogendoorn, A. W., & Dijkstra, W. (2007). Linked Markov sources: Modeling outcome-dependent social processes. Sociological Methods & Research, 36(1), 26–47. http://doi.org/10.1177/0049124106295975

Granic, I., & Hollenstein, T. (2003). Dynamic systems methods for models of developmental psychopathology. Development and Psychopathology, 15(3), 641–669.

Granic, I., Hollenstein, T., Dishion, T. J., & Patterson, G. R. (2003). Longitudinal analysis of flexibility and reorganization in early adolescence: A dynamic systems study of family interactions. Developmental Psychology, 39(3), 606–617. http://doi.org/10.1037/0012-1649.39.3.606

Granic, I., Hollenstein, T., Dishion, T. J., & Patterson, G. R. (2003). Longitudinal analysis of flexibility and reorganization in early adolescence: A dynamic

systems study of family interactions. Developmental Psychology, 39(3), 606–617. doi:10.1037/0012-1649.39.3.606.
Hollenstein, T. (2007). State space grids: Analyzing dynamics across development. International Journal of Behavioral Development, 31(4), 384–396. http://doi.org/10.1177/0165025407077765
Hollenstein, T. (2013). State Space Grids. Boston, MA: Springer US. Retrieved from http://link.springer.com/10.1007/978-1-4614-5007-8
Hollenstein, Tom. (2003). State space grids: Analyzing dynamics across development. International Journal of Behavioral Development, 31(4), 384–396. doi:10.1177/0165025407077765.
Kahn, R. L., & Cannell, C. F. (1957). The dynamics of interviewing; theory, technique, and cases. New York: Wiley.
Lewis, M. D., Zimmerman, S., Hollenstein, T., & Lamey, A. V. (2004). Reorganization in coping behavior at 1½ years: dynamic systems and normative change. Developmental Science, 7(1), 56–73. http://doi.org/10.1111/j.1467-7687.2004.00323.x
Maynard, D. W., & Schaeffer, N. C. (2002). Standardization and its discontents. In Douglas W. Maynard, Hanneke Houtkoop-Steenstra, Nora Cate Schaeffer, & Johannes van der Zouwen (Eds.), Standardization and tacit knowledge: Interaction and practice in the survey interview (pp. 3–47). New York: John Wiley & Sons.
Schaeffer, N. C. (1991). Conversation with a purpose—or conversation? Interaction in the standardized interview. Measurement Errors in Surveys, 365–391.
Yan, T., Curtin, R., & Jans, M. (2010). Trends in Income Nonresponse Over Two Decades. Journal of Official Statistics, 26(1), 145–164.