REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*

2. REPORT TYPE
FINAL REPORT

3. DATES COVERED *(From - To)*
1 Jun 05 – 31 May 08

4. TITLE AND SUBTITLE
Testing Transitivity and Related Axioms of Preference for Individuals and Small Groups

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA9550-05-1-0356

5c. PROGRAM ELEMENT NUMBER
61102F

6. AUTHOR(S)
DR Michel Regenwetter

5d. PROJECT NUMBER
2313

5e. TASK NUMBER
BX

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF ILLINOIS
109 Coble Hall, 801 South Wright Street
Champaign, IL 61820-6200

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR.RSL
875 NORTH RANDOLPH ST, ROOM 3112
ARLINGTON, VA 22203-1678

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-OSR-VA-TR-2012-0482

12. DISTRIBUTION/AVAILABILITY STATEMENT
DISTRIBUTION A – Approve For Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project has been completed. The original aim was to revisit transitivity of preference in individuals and small groups. For individual preferences, the project has shown that the past 50 years of research has fallen victim to a variety of conceptual, mathematical and statistical errors. We have made tremendous progress in this area, and, in particular, set the stage for a new generation of research in individual decision making. The project has also made headway in models for the evolution of preferences, as well as in behavioral social choice.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*

AFOSR FA9550-05-1-0356 Final Project Report

PI: Michel Regenwetter

University of Illinois

September 2008

Status of Effort.

This project has been completed. The original aim was to revisit transitivity of preference in individuals and small groups. For individual preferences, the project has shown that the past 50 years of research has fallen victim to a variety of conceptual, mathematical and statistical errors. We have made tremendous progress in this area, and, in particular, set the stage for a new generation of research in individual decision making. The project has also made headway in models for the evolution of preferences, as well as in behavioral social choice.

Accomplishments / New Findings: Organized by Resulting Publications and Themes.

1) Early in the project, Clintin Davis-Stober made a breakthrough methodological discovery: He solved the problem of frequentist inference with linear order constraints on the parameters. This discovery opened up the entire project to a new level of methodological rigor. Davis-Stober's discovery will soon be published. A summary of the paper follows below, and a copy is attached at the end of the package.

Summary of Davis-Stober, C. (in press), "Multinomial models under linear inequality constraints: Applications to measurement theory," *Journal of Mathematical Psychology.*

> Multinomial random variables are used across many disciplines to model categorical outcomes. Under this framework, investigators often use a likelihood

20120918119

ratio test to determine goodness-of-fit. If the permissible parameter space of such models is defined by inequality constraints, then the maximum likelihood estimator may lie on the boundary of the parameter space. Under this condition, the asymptotic distribution of the likelihood ratio test is no longer a simple χ^2 distribution. This article summarizes recent developments in the constrained inference literature as they pertain to the testing of multinomial random variables, and extends existing results by considering the case of jointly independent multinomial random variables of varying categorical size. This article provides an application of this methodology to axiomatic measurement theory as a means of evaluating properly operationalized measurement axioms. This article generalizes Iverson and Falmagne's (1985) seminal work on the empirical evaluation of measurement axioms and provides a classical counterpart to Myung, Karabatsos, and Iverson's (2005) Bayesian methodology on the same topic.

2) At the core of the project lies the reconsideration of prior claims about intransitive preference in individual decision makers. Given that the literature has produced in excess of 100 papers reporting intransitive preference behavior, this was an enormous task. Our main result is that every paper has succumbed to one or more of a list of about a dozen conceptual, mathematical, or statistical errors. When modeled correctly and analyzed with suitable statistical methods, then the published data on "intransitive" preferences are not providing convincing evidence for axiom violations. Moreover, in our own experiments, we have not found compelling evidence even though we used the most prominent empirical paradigm in the field.

These results are now under revision for possible publication in the leading theory paper of psychology. We had five referees (!), none of who had any major reservations with the paper. The revision needs to primarily shorten and simplify the exposition, as well as adress additional questions raised by the referees and the action editor. A summary of the paper follows below, and a copy of the original submission to Psychological Review is attached at the end of the package.

Summary of Regenwetter, M., Dana, J. & Davis-Stober C. (under revision). "Transitivity of Preferences." Target journal: *Psychological Review*.

> Transitivity of preferences is a fundamental principle shared by most major contemporary rational, prescriptive and descriptive models of decision making. We discuss why unambiguous evidence for violations of transitive preferences in individual decision makers is currently lacking. In counterpoint to Tversky's seminal (*Psychological Review*, 1969) "Intransitivity of Preferences," we reconsider his data as well as those from more than 20 other papers on "intransitive" human or animal decision makers. We challenge the standard operationalizations of transitive preferences and discuss pervasive methodological problems in the collection, modeling and analysis of relevant empirical data. We argue that "stochastic transitivity" should be altogether abandoned as a model of preference transitivity. We use parsimonious mixture models, where the parameter space of permissible preference states is the family of (transitive) strict linear orders. We show that the data from many of the available studies designed to elicit intransitive choice are consistent with not just transitive preferences, but with variable strict linear order preferences. Using transitivity as an example, we discuss conceptual and methodological problems that permeate behavioral decision research.

3) Davis-Stober and I have also written a tutorial review about transitivity of preferences, that has appeared as a book chapter. I only provide the abstract here, as the full publication is available in print.

Summary of Regenwetter, M. & Davis-Stober, C. (2008). "There are many models of transitive preference: A tutorial review and current perspective." In *Decision Modeling and Behavior in Uncertain and Complex Environments,* (T. Kugler, J.C. Smith, T. Connolly, and Y.-J. Son, Eds.), Springer, NY, *Series on Optimization and its Applications,* 99-124.

> Transitivity of preference is a fundamental rationality axiom shared by nearly all normative, prescriptive and descriptive models of preference or choice. There are many possible models of transitive preferences. We review a general class of such models and we summarize a recent critique of the empirical literature on (in)transitivity of preference. A key conceptual hurdle lies in the fact that transitivity is an algebraic/logical axiom, whereas experimental choice data are,

by design, the outcomes of sampling processes. We discuss probabilistic specifications of transitivity that can be cast as (unions of) convex polytopes within the unit cube. Adding to the challenge, probabilistic specifications with inequality constraints (including the standard "weak stochastic transitivity" constraint on binary choice probabilities) fall victim to a "boundary problem" where the log-likelihood test statistic fails to have an asymptotic χ^2-distribution. This invalidates many existing statistical analyses of empirical (in)transitive choice in the experimental literature. We summarize techniques to test models of transitive preference based on two key components: (1) we discuss probabilistic specifications in terms of convex polytopes, and (2) we provide the correct asymptotic distributions to test them. Furthermore, we demonstrate these techniques with examples on illustrative sample data.

4) One of the major problems with prior transitivity research resides in the fact that the standard empirical paradigm, namely the two-alternative forced choice task, forces the axioms of completeness and asymmetry to hold in every paired comparison. When we combine transitivity with those two axioms, then we obtain linear orders. In other words, testing transitivity in the presence of completeness and asymmetry is tantamount to testing linear order preferences. However, linear orders are only a microscopic fraction of all transitive relations. We have moved the empirical paradigm to two-alternative nonforced choice, in which decision makers are allowed to express indifference among choice options. We find an amazingly good fit of a weak order polytope to such data, even though the model is prohibitively restrictive. We are currently in the process of writing up these findings for publication in a major journal.

Preliminary summary of Regenwetter, M. & Davis-Stober, C. (in preparation). "Variability of Choice Behavior and Weak Order Preferences" (title tentative).

According to rational theory, a decision maker must have mutually consistent preferences among a sequence of pairs of options. More than a hundred papers report that animal and human choice violates a fundamental mathematical consistency axiom called "transitivity." Recent work has pointed out pervasive

> conceptual, mathematical and statistical errors in that literature. The present approach circumvents a number of possible sources of artifacts. Using new quantitative interdisciplinary methodologies we dissociate variability of choice from inconsistency of preference. We show that laboratory choice behavior in a classical "intransitivity" paradigm is, in fact, compatible with variable, but transitive, "weak order" preferences. We find that decision makers act in accordance with a restrictive mathematical model that, for the behavioral sciences, is extraordinarily parsimonious.

5) On a parallel research track, I also studied the evolution of preferences over time. One paper considers random walks on the graph of weak orders or the graph of semiorders. Since it has already appeared in print, I only include an abstract in this report.

Summary of Hsu, Y.-F. & Regenwetter, M. (2007). "Applications of stochastic media theory to 1992, 1996, and 2000 National Election Study panel data." *Chinese Journal of Psychology*, **49**, 225-244.

> We study a class of stochastic models of persuasion that form an application of media theory developed by Falmagne and others. These models describe the evolution of preferences over time. We consider the case where personal preferences are represented by (strict) weak orders and semiorders. Over time, these preferences may change under the influence of "tokens" of information arising in the environment. Successful applications of some weak order implementations of stochastic media theory to 1992 U.S. National Election Study (NES) panel data have been reported by Falmagne and various collaborators. However, past attempts to fit a semiorder model to the same data have failed. We successfully fit four media theoretic models, including two semiorder models based on the "neighboring" response mechanism, to 1992, 1996, and 2000 NES panel data. We compare the fit of these four models, discuss the psychological interpretation of key model parameters and illustrate applications to negative political campaigning.

6) The latest work in this line of research is just about to go out to a major decision

making journal. I include a summary below and attach a copy of a paper draft to the report.

Summary of Hsu, Y.-F. & Regenwetter, M. (in preparation). "Semiorder Preference, Persuasion, and Mover/Stayers."

> We capture three important properties of human judgment and decision making: 1) Individual preferences may take the form of semiorders, in which strict preference is transitive, but indifference is intransitive, 2) preferences may change over continuous time in reaction to persuasion processes in the environment, and yet, 3) some decision makers may be impervious to a given persuasion campaign.
>
> Semiorders have a long tradition as theoretical models of preference (Fishburn, 1979; Luce, 1956, 1959; Pirlot and Vincke, 1997) but have received little empirical attention (Böckenholt, 2001). We discuss stochastic processes on the graph of semiorders over a fixed finite set of choice alternatives, taking into account that decision makers can be "movers" or "stayers," depending on whether they react to the persuasion campaign or not. We apply these models to three illustrative sets of attitudinal panel data regarding United States presidential candidates. The data are from the American National Election Studies of the 1992, 1996 and 2000 campaigns.

7) In related work on subset choice, my collaborators and I considered a variety of decision heuristics that can play a role in approval voting. We developed a quantitative model and estimated the proportion of voters who act according to the different candidate heuristic models. This paper has now appeared in the top theory journal in psychology, and thus I only include an abstract below.

Summary of Regenwetter, M., Ho, M.-H. & Tsetlin, I. (2007). "Sophisticated approval voting, ignorance priors and plurality heuristics: A Behavioral Social Choice analysis in a Thurstonian framework." *Psychological Review*, **114**, 994-1014.

> This procjet reconciles historically distinct paradigms at the interface between individual and social choice, as well as rational and behavioral decision theory. We combine a utility-maximizing prescriptive rule for "sophisticated approval voting" with the "ignorance prior" heuristic from behavioral decision research, and two

> types of "plurality heuristics" to model approval voting behavior. When using a "sincere plurality heuristic," voters simplify their decision process by voting for their single favorite candidate. When using a "strategic plurality heuristic," voters strategically focus their attention on the two front runners and vote for their preferred candidate among these two. Using a hierarchy of Thurstonian random utility models, we implement these different decision rules and test them statistically on seven real world approval voting elections. We cross-validate our key findings via a psychological internet experiment. Although substantially many voters used the plurality heuristic in the real elections, they did so sincerely, not strategically. Moreover, even though Thurstonian models do not force such agreement, we find, in contrast to common wisdom about social choice rules, that the sincere social orders by Condorcet, Borda, Plurality, and Approval Voting are identical in all seven elections and in the internet experiment.

8) My ongoing research on preference aggregation has led to paper invitations by two major journals. The first is an invited "Major Open Problem" paper to appear in *Perspectives on Psychological Science*. The paper is in press, pending minor editorial changes. I attach a copy to the report.

Summary of Regenwetter, M. (in press). "Perspectives on Preference Aggregation." *Perspectives on Psychological Science.*

> For centuries, the mathematical aggregation of preferences by groups, organizations or society has received keen interdisciplinary attention. Extensive 20th century theoretical work in Economics and Political Science highlighted that competing notions of "rational social choice" intrinsically contradict each other. This led some researchers to consider coherent "democratic decision making" a mathematical impossibility. Recent empirical work in Psychology qualifies that view. This nontechnical review sketches a quantitative research paradigm for the behavioral investigation of mathematical social choice rules on real ballot, experimental choice, or attitudinal survey data. The paper poses a series of open questions.

Some classical work sometimes makes assumptions about voter preferences that are descriptively invalid. Do such technical assumptions lead the theory astray? How can empirical work inform the formulation of meaningful theoretical primitives? Classical "impossibility results" leverage the fact that certain desirable mathematical properties logically cannot hold universally in all conceivable electorates. Do these properties nonetheless hold in empirical distributions of preferences? Will future behavioral analyses continue to contradict the expectations of established theory? Under what conditions and why do competing consensus methods yield identical outcomes?

9) A second invited paper, this time for a leading biology journal, provides a status report on behavioral social choice. This paper is now conditionally accepted for publication pending final approval of the entire journal issue by the Editor in Chief. I include an abstract and attach a copy of the manuscript.

Summary of Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. & Cavagnaro, D. (conditionally accepted). "Behavioral social choice: A status report," *Philosophical Transactions of the Royal Society B: Biological Sciences*).

Behavioral social choice has been proposed (Regenwetter et al., 2006) as a social choice parallel to seminal developments in other decision sciences, such as *behavioral decision theory*, *behavioral economics*, *behavioral finance* and *behavioral game theory*. Behavioral paradigms compare how rational actors *should* make certain types of decisions with how real decision makers behave *empirically*. We highlight that important theoretical predictions in social choice theory change dramatically under even minute violations of standard assumptions. Empirical data violate those critical assumptions. We argue that the nature of preference distributions in electorates is ultimately an empirical question, which social choice theory has often neglected.

We also emphasize important insights for research on decision making by individuals. When researchers aggregate individual choice behavior in laboratory experiments to report summary statistics, they are implicitly applying social choice rules. Thus, they should be aware of the potential for aggregation paradoxes. We hypothesize that such problems may substantially mar the conclusions of a number of (sometimes seminal) articles in behavioral decision research.

10) Returning to the core theme of my research project, namely the formulation and testing of models of transitive preferences, a group of collaboratos have helped to solve an important mathematical and methodological problem that arises in the formulation and testing of ordinal decision theories. Most of the research on this component project is done, and we are in the process of drafting a manuscript. The authorship and tentative title are as follows.

Davis-Stober, C., Doignon, J.-P., Fiorini, S., Glineur, F. & Regenwetter, M. (in preparation). "Quantitative Testing of Ordinal Decision Theories: Constrained Inference, Order Polytopes, and Network Flows" (title tentative).

This highly interdisciplinary paper combines network flows from discrete mathematics, order constrained inference from statistics and mathematical psychology, as well as advanced numerical analysis methods from operations research to formulate computationally tractable tests of certain ordinal decision theories.

11) Researchers face a fundamental conceptual hurdle in testing decision theories against empirical data. Many theories, including Cumulative Prospect Theory and the Priority Heuristic are formulated in terms of logical, algebraic statements that are deterministic, in the sense that they do not model a decision process via random variables. Empirical data, on the other hand are, by the design of the data collection method, outcomes of random variables. Indeed, most standard statistical analyses of empirical data require that the data form an independent and identically distributed (iid) random sample from a population of interest. In general, for algebraic theories and relevant empirical data resulting from a random sampling process, it is therefore necessary to bridge the conceptual (and mathematical) gap between theory and data.

This problem has long been known as one of the most profound (and unresolved) challenges to empirical testing, e.g., of decision theories (Luce and Narens, 1994; Luce, 1995, 1997). Luce's challenge is to 1) recast a deterministic theory as a probabilistic model (or hypothesis) with respect to a suitable empirical sample space, and 2) use the appropriate statistical methodology for testing the probabilistic model of the theory (or hypothesis) on available data.

Both problems are nontrivial. Luckily, the outlook on Luce's challenge is no longer grim. Important developments in the statistical domain have revolutionized the types of quantitative constraints we can now test (Davis-Stober, 2008; Iverson, 1990; Iverson and Falmagne, 1985; Myung et al., 2005). Luce's challenge has been recognized, sometimes independently, by other leading scholars (see, e.g., Carbone and Hey, 2000; Harless and Camerer, 1994; Hey, 1995, 2005; Hey and Orme, 1994; Iverson and Falmagne, 1985; Loomes and Sugden, 1995; Starmer, 2000; Tversky, 1969). Some of these researchers have emphasized that different probabilistic specifications of the same theory may lead to diametrically opposite quantitative predictions.

Jointly with Clintin Davis-Stober and William Messner, I have developed a panoply of probabilistic specification methods for algebraic decision theories. To our knowledge, most of these methods are new. We are currently in the process of developing a publication plan for these new developments.

12) The dominant paradigm in the literature on transitive preference in individual decision makers has been the notion of "weak stochastic transitivity." Basically, a decision maker's preference is modeled through their modal pairwise choices among gambles. While we show in Regenwetter, Dana & Davis-Stober (under revision) that weak stochastic transitivity is not an appropriate model for transitive individual preference, it is nonetheless interesting to check to what extent this property holds in empirical data. Weak stochastic transitivity holds if and only if majority aggregated choices are transitive. In other words, this property is interesting from a social choice perspective.

Nearly all previous statistical analyses of weak stochastic transitivity are incorrect because they have failed to treat the test as an order constrained statistical inference task. We have started work on checking weak stochastic transitivity on more than 100 published data sets, as well as our own data. This work is ongoing beyond the end of the project.

13) While work on individual decision making dominated most of the project, we have also started work on small group decision making. Towards the end of the project, our lab ran a variation of Tversky's (Tversky, 1969) study, but with the main additional feature

that choices were recorded both by individual decision makers, as well as by diads of decision makers. We considered both established diads (room mates) and ad hoc diads (room mates were split up and matched with others). We have not yet started to analyze the data from this experiment.

Personnel Supported.

From August 2005 to May 2006, quantitative psychology PhD student Aeri Kim contributed to laboratory research and to the analysis of empirical data. Mrs. Kim later took a leave of absence to accompany her husband on a job assignment on the West Coast.

From August 2006 to May 2008, quantitative psychology PhD student William Messner contributed to laboratory research, the analysis of empirical data, and to publication of findings. His main task was the design, programming and execution of laboratory experiments. Mr. Messner is planning to write his Masters thesis on research that was partially funded by this project, and will most likely write his PhD thesis on follow-up work.

For one summer month in 2005, and then from August 2005 to August 2006, quantitative psychology PhD student Clintin Davis-Stober has made very extensive contributions to the theoretical developments, as well as to experimental design, and manuscript write-up. Mr. Davis-Stober is currently writing his PhD thesis on research that was supported by this project.

From summer 2007 to summer 2008, quantitative psychology PhD student Anna Popova contributed primarily to literature review and data analysis efforts for the project.

In the summer of 2007, quantitative psychology PhD student Jay Verkuilen contributed to the programming of some data analysis routines. Dr. Verkuilen has since graduated and taken a faculty position in academia.

Publications in the last 12 Months

Davis-Stober, C. (in press), "Multinomial models under linear inequality constraints: Applications to measurement theory," *Journal of Mathematical Psychology.*

Hsu, Y.-F. & Regenwetter, M. (rejected by *Management Science* on the basis of fit). "Semiorder Preference, Persuasion, and Mover/Stayers." Under revision for submission to a major decision making journal.

Regenwetter, M. (in press). "Perspectives on Preference Aggregation." *Perspectives on Psychological Science.*

Regenwetter, M., Dana, J. & Davis-Stober C. (revise and resubmit). "Transitivity of Preferences." *Psychological Review.*

Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. & Cavagnaro, D. (conditionally accepted). "Behavioral social choice: A status report," *Philosophical Transactions of the Royal Society B: Biological Sciences.*

Interactions / Transitions

A) PARTICIPATION/PRESENTATION AT MEETINGS, CONFERENCES, SEMINARS

1) Conference Organizer, Session Organizater, Session Chair (M. Regenwetter)

Conference Organizer Annual *European Mathematical Psychology Group* meeting, September 2007, Luxembourg (co-organized with Raymond Bisdorff and others). Invited Guest Speakers: Birnbaum (Fullerton), Böckenholt (McGill), Doignon (Brussels), Erdfelder (Mannheim), Iverson (UC Irvine), Pekeč (Duke), Rieskamp (Max Planck Berlin), Roberts (Rutgers), Schweickert (Purdue), Wakker (Rotterdam).

Session Chair Society for Mathematical Psychology (annual meeting 2007),

European Mathematical Psychology Group (annual meetings 2006, 2007, 2008).

Session Organizer and Chair European Mathematical Psychology Group (annual meeting, summer 2006). Symposium on *Transitive or Intransitive Preference.* Invited Speakers: Doignon (Brussels), Davis-Stober (U.I. Urbana-Champaign).

Institute for Operations Research and the Management Sciences (fall 2007). Session on *Testing Decision Theories and Decision Axioms.* Invited speakers: Clintin Davis-Stober (U.I. Urbana-Champaign), William Messner (U.I. Urbana-Champaign).

Program Committee Society for Judgment and Decision Making (annual conference, fall 2008).

2) Invited Speaker (M. Regenwetter)

2005 Decision Sciences Seminar, INSEAD.

2006 Cognition and Decision Program Review, Air Force Office of Scientific Research (with C. Davis-Stober, presented by CDS).

Coombs Memorial Lecture, Department of Psychology, University of Michigan.

Workshop on Decision Modeling and Behavior in Uncertain and Complex Environments, University of Arizona.

Workshop on Polyhedral Combinatorics of Random Utility, Center for Discrete Mathematics and Theoretical Computer Science, Rutgers University.

Workshop on Voting Theory and Preference Modelling, Laboratoire d'Analyse et Modélisation de Systèmes pour d'Aide à la Décision, Université Paris Dauphine and Center for Discrete Mathematics and Theoretical Computer Science, Rutgers University.

Institute for Mathematical Behavioral Sciences, U.C. Irvine.

Psychology Colloquium, Department of Psychology, Universität Heidelberg.

Psychology Colloquium, Department of Psychology, Universität Mannheim.

Colloquium, Applied Mathematics Unit, University of Luxembourg.

Decision Processes Seminar, Wharton School of Business, University of Pennsylvania.

2007 Cognition and Decision Program Review, Air Force Office of Scientific Research.

Invited Inaugural Address for the *NSM Decision Lab*, Faculty of Management Sciences, University of Nijmegen,

Max Planck Institute for Human Development, Berlin.

2008 Cognition and Decision Program Review, Air Force Office of Scientific Research.

Formal Models of Memory, Judgment, & Decision Making, University of Mannheim.

Psychology Colloquium, Department of Psychology, University of Oklahoma at Norman.

Psychology Colloquium, Department of Psychology, University of Maryland at College Park.

3) Regular Conference Presentations (M. Regenwetter)

2005 Annual meeting, Society for Mathematical Psychology, University of Memphis (with C. Davis-Stober, presented by CDS).

2006 Annual meeting, (Ward Edwards) Bayesian Research Conference, Fullerton.

Workshop on Decision Research at UIUC, University of Illinois at Urbana-Champaign.

Annual meeting, Society for Mathematical Psychology, Vancouver.

Annual meeting, European Mathematical Psychology Group, École Nationale Supérieure des Télécommunications de Bretagne.

2007 Annual meeting, Society for Mathematical Psychology, Costa Mesa (co-author on four talks: with C. Davis-Stober and J. Dana; with C. Davis-Stober, presented by CDS, with W. Messner, presented by WM; with Y.-F. Hsu, presented by YFH).

Annual meeting, European Mathematical Psychology Group, University of Luxembourg. (co-author on three talks: with C. Davis-Stober and J. Dana; with C. Davis-Stober, presented by CDS, with W. Messner, presented by WM).

Annual Meeting, Institute for Operations Research and the Management Sciences, Seattle. (co-author on three talks: with C. Davis-Stober; with C. Davis-Stober, presented by CDS, with W. Messner, presented by WM).

2008 Annual meeting, European Mathematical Psychology Group, Universität Graz (with C. Davis-Stober).

4) Posters at Conferences

2005 Annual meeting, Society for Judgment and Decision Making, Toronto (co-author on two posters: with A. Kim; with C. Davis-Stober).

5) Colloquium Talks at Home Institution

2005 Program in Systems and Entrepreneurial Engineering Seminar, Department of General Engineering, U.I. Urbana-Champaign.

2006 Quantitative Seminar, Department of Psychology, U.I. Urbana-Champaign.

6) Presentations by Research Assistants Davis-Stober, Messner, and Popova

2005 Annual Meeting, Society for Mathematical Psychology, Memphis (paper presentation: Davis-Stober, C. P. with Regenwetter, M.).

2006 Annual Meeting, European Mathematical Psychology Group, Brest, France (paper presentation: Davis-Stober, C. P. with Regenwetter, M.).

Workshop on Polyhedral Combinatorics of Random Utility, Discrete Mathematics and Computer Science (DIMACS), Piscataway, NJ (paper presentation: Davis-Stober, C. P.).

Cognitive and Decision Making Program Review, Air Force Office of Scientific Research (AFOSR), Fairborn, OH (paper presentation: Davis-Stober, C. P. with Regenwetter, M.).

Annual Ward Edwards Bayesian Conference, Fullerton, CA (paper presentation: Davis-Stober, C. P.).

2007 Quantitative Seminar, Department of Psychology, U.I. Urbana-Champaign (paper presentation: Messner, W. with Regenwetter, M. and Davis-Stober, C. P.).

Annual Meeting, Society for Mathematical Psychology, Costa Mesa (paper presentation: Davis-Stober, C. P. with Regenwetter, M.).

Annual Meeting, Society for Mathematical Psychology, Costa Mesa (poster: Messner, W. with Regenwetter, M. and Davis-Stober, C. P.).

Annual Meeting, European Mathematical Psychology Group, University of Luxembourg (paper presentation: Davis-Stober, C. P. with Regenwetter, M.).

Annual Meeting, European Mathematical Psychology Group, University of Luxembourg (paper presentation: Messner, W. with Regenwetter, M. and Davis-Stober, C. P.).

Annual Meeting, Institute for Operations Research and the Management Sciences, Seattle (paper presentation: Messner, W. with Regenwetter, M. and Davis-Stober, C. P.).

2008 Annual Meeting, Public Choice Society, San Antonio (paper presentation: Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. P. and Cavagnaro, D., presented by AP).

Quantitative Seminar, Department of Psychology, U.I. Urbana-Champaign (paper presentation: Regenwetter, M., Grofman, B., Popova, A., Messner, W., Davis-Stober, C. P. and Cavagnaro, D., presented by AP).

B) CONSULTATIVE AND ADVISORY FUNCTIONS.

None.

C) TRANSITIONS.

None known.

New Discoveries, Inventions, Patent Disclosures.

None.

Honors/Awards

2007-2008 Associate, *Center for Advanced Study*, University of Illinois.

1999 Young Investigator Award, *Society for Mathematical Psychology.*

2006 Fellow, *Association for Psychological Science* (formerly *American Psychological Society*).

References

Böckenholt, U. (2001). Thresholds and intransitivities in pairwise judgments: A multilevel analysis. *Journal of Educational and Behavioral Statistics*, 26:269–282.

Carbone, E. and Hey, J. D. (2000). Which error story is best? *Journal of Risk and Uncertainty*, 20:161–176.

Davis-Stober, C. P. (2008). A frequentist framework for axiom testing. Under revision for Journal of Mathematical Psychology.

Fishburn, P. C. (1979). *Utility Theory for Decision Making.* R. E. Krieger, Huntington.

Harless, D. W. and Camerer, C. F. (1994). The predictive value of generalized expected utility theories. *Econometrica*, 62:1251–1289.

Hey, J. D. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review*, 39:633–640.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental Economics*, 8:325–345.

Hey, J. D. and Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 62:1291–1326.

Iverson, G. J. (1990). Probabilistic measurement theory. Available as MBS 90-23 Technical Report at the IMBS, University of California, Irvine.

Iverson, G. J. and Falmagne, J.-C. (1985). Statistical issues in measurement. *Mathematical Social Sciences*, 10:131–153.

Loomes, G. and Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, 39:641–648.

Luce, R. D. (1956). Semiorders and a theory of utility discrimination. *Econometrica*, 26:178–191.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis.* John Wiley, New York.

Luce, R. D. (1995). Four tensions concerning mathematical modeling in psychology. *Annual Review of Psychology*, 46:1 – 26.

Luce, R. D. (1997). Several unresolved conceptual problems of mathematical psychology. *Journal of Mathematical Psychology*, 41:79–87.

Luce, R. D. and Narens, L. (1994). Fifteen problems in the representational theory of measurement. In Humphreys, P., editor, *Patrick Suppes: Scientific Philosopher, volume 2: Philosophy of Physics, Theory Structure, Measurement Theory, Philosophy of Language, and Logic*, pages 219–245. Kluwer, Dordrecht.

Myung, J., Karabatsos, G., and Iverson, G. (2005). A Bayesian approach to testing decision making axioms. *Journal of Mathematical Psychology*, 49:205–225.

Pirlot, M. and Vincke, P. (1997). *Semiorders: Properties, Representations, Applications.* Kluwer Academic Publishers, Dordrecht.

Regenwetter, M., Grofman, B., Marley, A., and Tsetlin, I. (2006). *Behavioral Social Choice.* Cambridge University Press, Cambridge, UK.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38:332–382.

Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76:31–48.

Appendix 1: Experiments.

I) 2AnFC EXPERIMENT

Background

This study was an expansion of Tversky's (1969) main experiment. Tversky had eight subjects do altogether 20 separate two-alternatives forced choice (2AFC) decisions for each pair among five gambles. We used similar gambles as stimuli, but decisions were two-alternatives non-forced choice (2AnFC), and the experiment involved a greater number of subjects and trials.

Details of the Experiment

Each subject attended three sessions, and each session took approximately one hour to complete. Each session consisted of the subject working at a computer, making choices between comparable gambles (described below). The computer program iteratively displayed pairs of gambles, and the subject used the computer mouse to click on the gamble in the pair that he/she preferred. If the subject could not decide between the two gambles, he/she clicked the third option, labeled "indifferent." After a few warmup choices at the beginning of each session, the subject made 600 choices between gambles. The subjects were allowed to take short breaks whenever necessary to reduce fatigue during a session.

The gambles were organized into four sets: *Cash1*, *Cash2*, *Noncash*, and *Distractors*. The gambles in these sets are described below. There are five gambles in each set, so each set's gambles can be paired ten distinct ways. The computer cycled through the four sets as it chose gamble pairs for the subject, i.e., one pair of gambles from each set was shown to the subject over a four-trial cycle, and then the cycle repeated in the same order with new gamble pairs from each set. The gamble pairs were drawn from the sets in a quasi-random fashion with the conditions that the same gamble pair did not appear in consecutive cycles, and each gamble pair (except from the *Distractor* set) was shown to the subject 15 times per session. The gambles are listed below (in 2006 dollars):

Cash1

Gamble	A	B	C	D	E
Probability of Winning	7/24	8/24	9/24	10/24	11/24
Payoff	$26.32	$25.00	$23.68	$22.36	$21.04

Cash2

Gamble	A	B	C	D	E
Probability of Winning	0.28	0.32	0.36	0.40	0.44
Payoff	$31.43	$27.50	$24.44	$22.00	$20.00

Noncash

Gamble	A	B	C	D	E
Probability of Winning	0.26	0.24	0.22	0.20	0.18

Payoffs included gift cards redeemable for the following prizes from local shops: 15 sandwiches, 40 movie rentals, 4 CDs, 40 coffees, or 7 books. The gift cards were worth about $75. At the beginning of the experiment, subjects ranked these prizes 1-5 in order of preference. The first-ranked prize was associated with the lowest probability of winning, the second-ranked prize associated with the second lowest probability, and so on.

Distractors

These gambles were randomly generated (using uniform distributions). The payoffs in this set were randomly chosen from the 15 payoffs in the three sets above. The probabilities of winning were randomly generated from the interval [0.02, 0.42] .

Subject Information and Recruiting

30 subjects completed all three sessions. One additional subject completed only two sessions. No information was collected in regard to age, gender, or other demographic factors, but the subjects were roughly balanced in terms of gender, and most, if not all, were college students.

Subjects were recruited through posted advertisements in academic buildings and dormitories around the University of Illinois campus. Subjects were informed that the study involved decision making, and that they would be paid for their participation in the study. As the research applies to members of the general adult population who are able to read and perform a simple choice or ranking task on the computer, no particular subject groups were targeted or excluded.

After a short instructional presentation in the first session, participants signed a consent form and were given the opportunity to try the task and decide if they were able to perform it. All subjects chose to proceed.

Payment

At the end of each session a subject completed, he/she received a \$5 minimum payment.

In addition to this baseline payment, a gamble was randomly selected (using uniform distributions) from a collection of cash gambles that the subject chose during the session, and the subject played this gamble for real money. For gamble pairs where the subject clicked 'indifferent,' a gamble from the pair was chosen at random and added to the collection of chosen gambles (only for the purpose of payment). This gave subjects incentive to choose gambles carefully, as they might eventually play any gamble they chose during the session, and to use the indifference option sparingly, since it could result in either gamble in a pair entering the collection of potential gambles subjects could ultimately face. Since cash gambles ranged in value from \$20-\$31.43, subjects earned \$5-\$36.43 per session.

To add further incentive for the subjects to complete all three sessions, at the end of the third session, a noncash gamble (such as a chance for free movie rentals) was also played in addition to the cash gamble.

Record-keeping

All data were stored on a password-protected computer in a locked laboratory space.

To ensure that subject data could be linked between sessions, subjects' names and arbitrarily chosen experimental ID numbers were recorded together on notecards. These records were kept in a secure location until the experiment ended and were then destroyed. The participants' names were also recorded on a signed consent form.

In the data files, the participants are only identified by their experimental IDs. Since there is no remaining link between their names and these ID numbers once the notecards are destroyed, the confidentiality of the data is ensured.

II) 2AFC Experiment

This study was an expansion of Tversky's (1969) main experiment. Tversky had eight subjects do altogether 20 separate two-alternatives forced choice (2AFC) decisions for each pair among five gambles. We used similar gambles as stimuli and the 2AFC paradigm, but the experiment involved a greater number of subjects and trials.

Details of the Experiment

Each subject attended one experimental session which took approximately one-and-a-half hours to complete. This session consisted of the subject working at a computer, making choices between comparable gambles (described below). The computer program iteratively displayed pairs of gambles, and the subject used the computer mouse to click on the gamble in the pair that he/she preferred. After a few warmup choices, the subject made 800 choices between gambles. The subjects were allowed to take short breaks whenever necessary to reduce fatigue during a session.

The gambles were organized into four sets: *Cash1*, *Cash2*, *Noncash*, and *Distractors*. The gambles in these sets are described below. There are five gambles in each set, so each set's gambles can be paired ten distinct ways. The computer cycled through the four sets as it chose gamble pairs for the subject, i.e., one pair of gambles from each set was shown to the subject over a four-trial cycle, and then the cycle repeated in the same order with new gamble pairs from each set. The gamble pairs were drawn from the sets in a quasi-random fashion with the conditions that the same gamble pair did not appear in consecutive cycles, and each gamble pair (except from the *Distractor* set) was shown to the subject 20 times. The gambles are listed below (in 2007 dollars):

Cash1

Gamble	A	B	C	D	E
Probability of Winning	7/24	8/24	9/24	10/24	11/24
Payoff	$28.00	$26.60	$25.20	$23.80	$22.40

Cash2

Gamble	A	B	C	D	E
Probability of Winning	0.28	0.32	0.36	0.40	0.44
Payoff	$31.43	$27.50	$24.44	$22.00	$20.00

Noncash

Gamble	A	B	C	D	E
Probability of Winning	0.26	0.24	0.22	0.20	0.18

Payoffs included gift cards redeemable for the following prizes from local shops: 15 sandwiches, 40 movie rentals, 4 CDs, 40 coffees, or 7 books. The gift cards were worth about $75. At the beginning of the experiment, subjects ranked these prizes 1-5 in order of preference. The first-ranked prize was associated with the lowest probability of winning, the second-ranked prize associated with the second lowest probability, and so on.

Distractors

These gambles were randomly generated (using uniform distributions). The payoffs in this set were randomly chosen from the 15 payoffs in the three sets above. The probabilities of winning were randomly generated from the interval [0.02, 0.42] .

Subject Information and Recruiting

18 subjects completed the experiment. No information was collected in regard to age, gender, or other demographic factors, but the subjects were roughly balanced in terms of gender, and most, if not all, were college students.

Subjects were recruited through posted advertisements in academic buildings and dormitories around the University of Illinois campus. Subjects were informed that the study involved decision making, and that they would be paid for their participation in the study. As the research applies to members of the general adult population who are able to read and perform a simple choice or ranking task on the computer, no particular subject groups were targeted or excluded.

After a short instructional presentation, participants signed a consent form and were given the opportunity to try the task and decide if they were able to perform it. All subjects chose to proceed.

Payment

At the end of the experiment, each subject received a $10 minimum payment.

In addition to this baseline payment, a gamble was randomly selected (using uniform distributions) from a collection of all gambles (both cash and noncash) that the subject had chosen during the session, and the subject played this gamble for a real prize. This gave subjects incentive to choose gambles carefully, as they might eventually play any gamble they chose during the session. Since gambles ranged in value from $20 (lowest cash amount) to $75 (approximate value of the noncash prizes), subjects earned $10-$85 per session.

Record-keeping

All data were stored on a password-protected computer in a locked laboratory space.

The participants' names were recorded on a signed consent form.

Subjects were assigned arbitrarily chosen experimental ID numbers at the beginning of the experiment. In the data files, participants are only identified by their experimental IDs. Since there is no recorded link between their names and these ID numbers, the confidentiality of the data is ensured.

III) DYAD EXPERIMENT

This study was an expansion of Tversky's (1969) main experiment. Tversky had eight subjects do altogether 20 separate two-alternatives forced choice (2AFC) decisions for each pair among five gambles. We used similar gambles as stimuli, but decisions were two-alternatives non-forced choice (2AnFC), and the experiment involved a greater number of subjects and trials. Most importantly, this experiment compared the degree of intransitivity present in decisions made by individuals, established dyads (here, roommates), and dyads formed in an ad hoc fashion (here, reshuffled roommates).

Details of the Experiment

The experiment was conducted using laptop computers in a University of Illinois Department of Psychology lab and occurred over three sessions. Subjects attended each of three sessions with their roommate (i.e., two people who lived together at least six months). Each session took approximately one and a half hours to complete.

Subjects were divided into two counterbalanced groups. In one group, subjects were partnered with their roommates. In the other group, subjects were partnered with another participant who was not their roommate. In two of the three sessions, subjects made a series of choices between gambles with their partner as a two-person team (dyad). In the other session, each subject made a series of choices alone. Roughly one-third of subjects chose by themselves in their first session, one-third in their second session, and one-third in their final session.

Each session consisted of the subject working at a computer (with or without a partner) and making choices between comparable gambles (described below). To begin a session, the computer introduced the subjects to the experiment through an instruction set and several practice choices. Then the computer iteratively displayed pairs of gambles, and participants chose which gamble they preferred to play or stated they were indifferent by clicking on the appropriate box on the computer screen. Each choice was be recorded by the computer. In all, the computer presented 600 gamble pairs per dyad session and 1200 gamble pairs in the individual session. Subjects were instructed that they could take short breaks whenever necessary to reduce fatigue while completing the experiment.

The gambles were composed solely of nonnegative monetary outcomes. Every choice was between two gambles belonging to one of the following five sets: 1) *Tversky Gambles*, 2) *Equal Expected Value Gambles*, 3) *Priority Heuristic Gambles*, 4) *Zero Outcome Distractor Gambles*, and 5) *Nonzero Outcome Distractor Gambles.*

Subjects were presented with all possible pairings of gambles within each set and were never presented with a pair composed of two gambles from different sets. The first three sets above each contain five gambles, so each of these sets' gambles can be paired ten distinct ways. Each and every such pair was shown to individuals 30 times per individual session and 15 times per dyad session. The distractors were randomly (using uniform distributions) generated during the experiment.

During a session, the subjects were presented with a pair of gambles from each set in the following order:

1) *Tversky Gambles*
2) 50-50 chance of seeing *Zero Outcome Distractor* or *Nonzero Outcome Distractor*
3) *Priority Heuristic*
4) *Equal Expected Value*

In the individual session, the above pattern repeated 300 times, so that the ten pairs in each set were repeated 30 times per session. In the dyad sessions, the pattern was repeated 150 times, so that the ten pairs in each set were repeated 15 times per session. The gamble pairs were drawn from the sets in a quasi-random fashion with the condition that the same gamble pair did not appear in consecutive cycles.

A description of each gamble set follows (in 2008 dollars):

Tversky Gambles - These gambles were all of the form "Win X with probability p, otherwise nothing with probability 1-p."

Gamble	A	B	C	D	E
Probability of Winning	7/24	8/24	9/24	10/24	11/24
Payoff	\$28.00	\$26.60	\$25.20	\$23.80	\$22.40

Equal Expected Value Gambles - These gambles were all of the form "Win X with probability p, otherwise nothing with probability 1-p."

Gamble	A	B	C	D	E
Probability of Winning	0.28	0.32	0.36	0.40	0.44
Payoff	\$31.43	\$27.50	\$24.44	\$22.00	\$20.00

Priority Heuristic Gambles - These gambles were all of the form "Win X (big payoff) with probability p, otherwise Y (small payoff) with probability 1-p."

Gamble	A	B	C	D	E
Probability of Winning Big	0.50	0.54	0.58	0.62	0.66
Big Payoff	$26.90	$25.28	$24.10	$22.98	$22.25
Small Payoff	$15.50	$16.40	$17.20	$18.30	$19.15

Zero Outcome Distractors - These gambles are generated by the computer during the experiment and are subject to the following constraints:

*They are randomly generated by the computer from a uniform distribution.

*One outcome is 0.

*The other outcome is greater than 0, but less than 30.

*Neither outcome has 0 or 1 probability, but their total probability is 1.

*The expected value of each gamble is less than $8.80.

Nonzero Outcome Distractors - These gambles are generated by the computer during the experiment and are subject to the following constraints:

*They are randomly generated by the computer from a uniform distribution.

*Both outcomes are greater than 0, but less than 30.

*Neither outcome has 0 or 1 probability, but their total probability is 1.

*The expected value of each gamble is less than $30.

Subject Information and Recruiting

Altogether 20 subjects completed all three sessions. Twelve subjects worked in dyads with their roomate; the other eight worked in dyads with another randomly chosen participant. No information was collected in regard to age, gender, or other demographic factors, but the subjects were roughly balanced in terms of gender, and most, if not all, were undergraduate or graduate students at the University of Illinois.

Subjects were recruited through the University of Illinois Psychology Department's website and posted advertisements in academic buildings and dormitories around the University of Illinois. Subjects were informed that the study involved decision making, and participation was paid, voluntary, and open to anyone who had a willing roommate. As the research applies to members of the general adult population who are able to read and perform a simple choice task on the computer, no particular subject groups were targeted or excluded.

After a short instructional, computer-based presentation in the first session, participants signed a consent form and were given the opportunity to try the task and decide if they were able to perform it. All subjects chose to proceed.

Payment

At the end of each session a subject completed, he/she received a \$16 baseline payment. In addition to this baseline payment, two gambles were chosen from the collection of gambles that the subject (dyad) had chosen during the session, and the subject (dyad) played these gambles for real money. One of these two gambles came from the *Tversky*, *Equal Expected Value*, or *Zero Outcome Distractor* sets. The other gamble came from the *Priority Heuristic* or *Nonzero Outcome Distractor* sets. For gamble pairs where the subject (dyad) expressed indifference, a gamble from the pair was chosen at random and added to the collection of chosen gambles (solely for the purposes of payment). This gave subjects incentive to choose gambles carefully, as they might eventually play any gamble they chose during the session, and to use the indifference option sparingly, since it could result in either gamble in a pair entering the collection of potential gambles subjects would ultimately face. Since gambles ranged in value from \$0 to approximately \$30, subjects earned \$16 to roughly \$76 per session. For the dyad sessions, each member of the team received the full amount from winning a game; prizes were not split.

Record-keeping

All data were stored on a password-protected computer in a locked laboratory space. To ensure that subject data could be linked between sessions, subjects' names and arbitrarily chosen experimental ID numbers were recorded together on notecards. These records were kept in a secure location until the experiment ended and were then destroyed. The participants' names were also recorded on a signed consent form.

In the data files, the participants were only identified by their experimental IDs. Since there is no remaining link between their names and these ID numbers once the notecards are destroyed, the confidentiality of the data is ensured. Since participants were paired with other participants in some sessions, we could not guarantee confidentiality from other participants, but we asked that all participants respect the privacy of the experiment.