Australian Mathematical Psychology 2020

Coogee Beach, Sydney, February 12-14

Welcome

UNSW is delighted to host the 2020 Australian Mathematical Psychology Conference in Sydney. The conference runs from the 12th-14th Feb with two and a half full days of talks.

The last time UNSW hosted the conference in 2013, there were 34 talks. We could safely host the conference in the small room of a nearby hotel. Attendees at that conference commented on how the conference seemed to have grown. There are 52 talks this year.

The committee would like to thank Alice Mason, Jake Embrey, and Amy Li for their assistance in preparing and running the meeting. We gratefully acknowledge the financial support of the School of Psychology at UNSW.

We would like to acknowledge the Gadigal and Bidjigal peoples who are the traditional custodians of the land on which the AMPC2020 meeting takes place.

Brett Hayes, Ben Newell, Chris Donkin, and Danielle Navarro
Organising Committee

Conference Code of Conduct

We are committed to making AMPC20 an inclusive, respectful and harassment-free experience. All attendees, speakers, sponsors and volunteers at our conference are required to agree with the following code of conduct. Our conference is dedicated to providing a harassment-free conference experience for everyone, regardless of gender, gender identity and expression, age, sexual orientation, disability, physical appearance, body size, race, ethnicity, religion (or lack thereof), or technology choices. We do not tolerate harassment of conference participants in any form. Sexual language and imagery is not appropriate for any conference venue, including talks, workshops, Twitter-posts and other online media connected with the conference.

If you have any concerns please contact one of the organisers.

Commitment to Sustainability

To promote more environment-friendly conference practices we will be using plastic-free namecards on recycled cardboard. We will not be supplying a conference bag which we suspect often wind up in landfill. So please bring your own backpack /tote. Pads and pens will be available at the venue.

Session Chairs

We would like the presenter of the **final** paper in each session to act as Session Chair. If you don’t want to chair a session please contact the conference organisers.

Conference Schedule (Session Chairs in Bold)

Wednesday 12-02-2020

Morning

Time	Authors	Title
09:10	Philip Smith, Simon Lilburn	Vision for the Blind: Visual Psychophysics and Blinded Inference for Decision Models
09:30	Andrew Heathcote, Angus Reynolds, Peter Kvam, Adam Osth	What do confidence and response time tell us about accuracy (and a model).
09:50	Thomas Narraway, Guy Hawkins, Scott Brown	A Novel Methodology For Manipulating The Speed-Accuracy Trade-Off
10:00	Sarah Moneer, Eleanore Williams, Daniel Little	The Effect of Distractors and Bilateral Presentation on Processing Architecture in Categorization Decisions
10:10	Caroline Kuhne, Scott Brown, Andrew Heathcote, Dora Matzke	Multivariate Normal Priors for Evidence Accumulation Models
10:20	**David K Sewell**, Rebecca Jaymes, Thomas J Palmeri	Modeling switch costs in a rule-plus-exception categorization task
10:40		**Coffee**
11:00	Roger Ratcliff, Philip L. Smith	Identifying Sources of Noise in Decision-Making
11:20	Gavin Cooper, Guy Hawkins, Ami Eidels	Modelling decision rules in consumer choice with the LBA
11:30	Steson Lo, Jeremy Ngo, Chris Donkin, Ben Newell	Nudging at drift rates: Using the diffusion model to study effects of probabilistic feedback on perceptual decisions
11:40	Jeremy Ngo, Chris Donkin	Using the Timed Racing Diffusion Model to account for responses to stimuli with limited information
11:50	Jörg Rieskamp, Ashley Luckman	The Use of Response Times as a Critical Test of Risky Decision-making Theories
12.10	**Guy Hawkins**, David Gunawan, Minh-Ngoc Tran, Robert Kohn, Mark Steyvers, Scott Brown	Latent class discovery for cognitive models
12:20		**Lunch**

Afternoon

Time	Authors	Title
13:45	Timothy Ballard, Andrew Neal, Simon Farrell, Andrew Heathcote	Where do we go wrong? Dynamic strategy regulation during dynamic, goal-directed decision making
14:05	Luke Strickland, Vanessa Bowden, Russell Boag, Andrew Heathcote, Shayne Loft	Modelling how humans use decision aids in simulated air traffic control
14:25	Erin Lloyd, Andrew Neal, Timothy Ballard	Team Coordination: A Multiple Goal Pursuit Problem
14:35	Andrew Morgan, Timothy Ballard, Andrew Neal	The Effects of Competition Type on Goal-Directed Decision Making
14:45	Alexander Thorpe, Sylvia Evans, Keith Nesbitt, Ami Eidels	Assessing Workload Capacity in a Continuous Task
14:55	**Zach Howard**, Paul Garrett, Daniel Little, Ami Eidels, James Townsend	Questioning the assumptions of the Capacity Coefficient
15:15		**Coffee**
15:40	Michael D. Lee, Holly A. Westfall	A model of triadic comparison applied to the problem of understanding semantic memory impairment
16:00	Quentin F. Gronau, Michael D. Lee	Bayesian Inference for Multidimensional Scaling Representations with Psychologically-Interpretable Metrics
16:20	Guangyu Zhu, Yiyun Shou, Michael Smithson	Extreme Response Style: Asymmetry and Domain Specificity
16:40	**Alexander Ly**, Udo Boehm, Simon Kucharsky, Botond Szabo, Dora Matzke, Andrew Heathcote	Using latent network structures to link neurophysiology and behaviour
17:00		**Soccer**

Thursday 13-02-2020

Morning

Time	Authors	Title
09:00	Amy Perfors, Keith Ransom	Polarisation, learning, and trust in populations of Bayesian agents
09:20	Keith Ransom, Charles Kemp, Amy Perfors	The role of utility in semantic categorisation
09:40	Brett Hayes, Yuki Wen, Danielle Navarro	When is absence of evidence evidence of absence? And who can tell the difference?
09:50	Belinda Xie, Danielle Navarro, Brett Hayes	Bayes for New and Old: Modelling type and token effects in induction
10:10	Shi Xian Liew, Danielle Navarro, Benjamin Newell	Wanting to know, or not wanting to not know? A comparison of information seeking models
10:20	**Charles Kemp**, Brett Hayes, Danielle Navarro	A causal Bayes net framework for reasoning from non-random samples
10:40		**Coffee**
11:00	Jessica C. Lee, Peter F. Lovibond, Brett K. Hayes, Stephan Lewandowsky	A mixture of experts in associative generalization
11:20	Aba Szollosi, Ben Newell	People as intuitive scientists: Reconsidering statistical explanations of human cognition
11:40	Jon-Paul Cavallaro, Guy Hawkins	Investigating the time course of feature processing in consumer-like choices
11:50	Tehilla Mechera-Ostrovsky, Steven Heinke, Sandra Andraszewicz, Jörg Rieskamp	On the Link Between Risk Preferences and Cognitive Abilities: A Meta-Analysis
12:00	**Christin Schulze**, Juliane E. Kammer, Mehdi Moussaid, Wolfgang Gaissmaier, Thorsten Pachur	Shall we risk it? Modeling risky choice in groups and individuals
12:20		**Lunch**

Afternoon

Time	Authors	Title
13:45	Michael Smithson, Yiyun Shou, Guangyu Zhu	Understanding Middle-Response Bias
14:05	Yonatan Vanunu, Jared M. Hotaling, Ben R. Newell	Top-down and Bottom-up Processes in Risky Choice
14:25	Murray Bennett, Rachel Mullard, Marc Adam, Mark Steyvers, Ami Eidels, Scott Brown	Competitive Decision Making in Dutch Auction
14:35	Scott Brown, Murray Bennett, Rachel Mullard, Ami Eidels	An Iterated Prospect Theory Model for the Dutch Auction
14:45	Paul Garrett, Zachary Howard, Yu-Tzu Hsieh, Murray Bennett, Cheng-Ta Yeng, Daniel Little, Ami Eidels	Symbolic wheel of fortune: confusion analysis and the mental representations of digits across cultures
14:55	**Alice Mason**, Christopher Maddan, Marcia Spetch, Elliot Ludvig	How recalling values from the edges of the distribution supports evaluation and choice
15:15		**Coffee**
15:40	John Dunn, Kym McCormick, David Kellen, Carolyn Semmler	The Geology of Eyewitness Memory
16:00	Yiyun Shou, Guangyu Zhu, Michael Smithson	Trait Risk Attitudes as a Result of Response Bias?
16:10	Kevin D. Shabahang, Hyungwook Yim, Simon Dennis	Propositions versus Associations in Memory for Sentences
16:20	Kym M McCormick, John C. Dunn, Carolyn Semmler, David Kellen	Forced choice eyewitness memory: Testing the assumption of random utility
16:30	Nicole Cruz, Brett K. Hayes, Anagha Kaluve, Jasmine Choi-Christou, John Dunn, Rachel Stephens	Using signed difference analysis to differentiate between reasoning processes
16:50	**Garston Liang**, Jennifer Sloane, Chris Donkin, Ben Newell	To cue or not to cue? Seeking and integrating external information into choice
17:00	Session Close	

CONFERENCE DINNER

Friday 14-02-2020

Morning

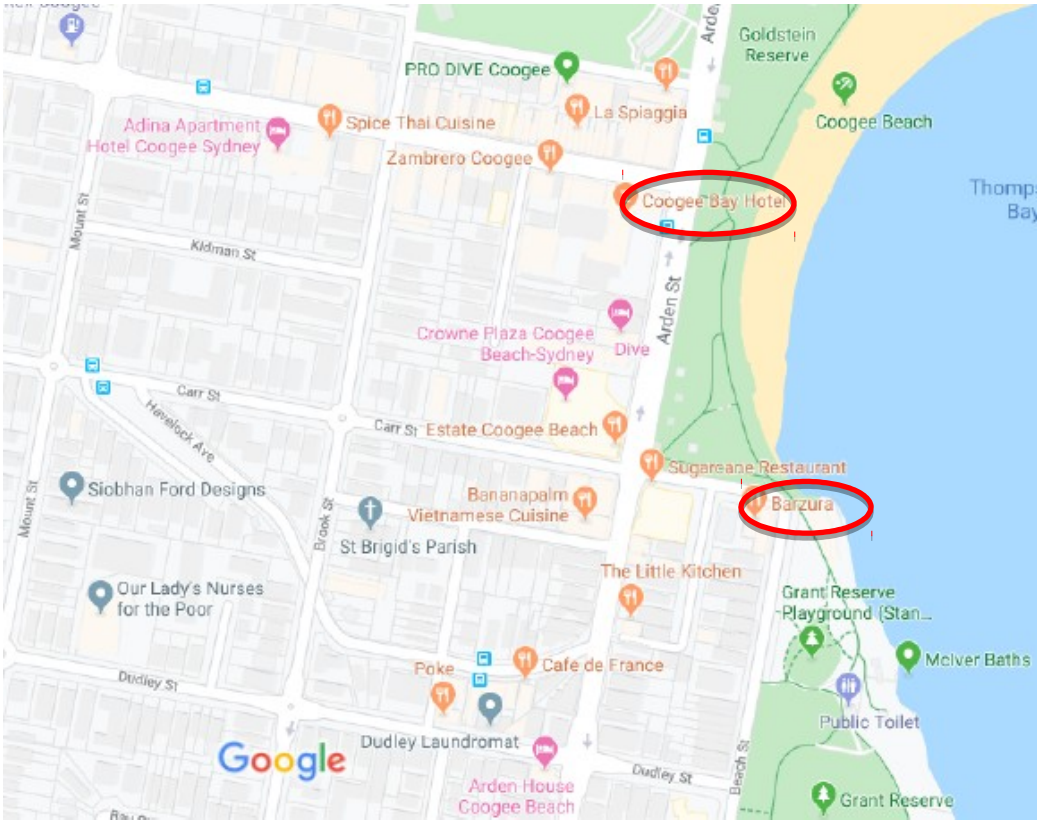
Time	Authors	Title
09:00	Simon Dennis, Elizabeth Laliberte, Hyungwook Yim, Benjamin Stone	Understanding Memory for WHERE using Smartphone Data
09:20	Adam F. Osth, Daniel Wraith-Franck	Integration of semantic and orthographic information in episodic memory
09:40	Jennifer Sloane, Corey White, Chris Donkin, Ryan Curl	Effects of Item Relatedness on Output Interference
09:50	Ed Stewart, Chris Donkin, Mike Le Pelley	The role of eye gaze in gaze contingent visual working memory tasks
10:00	Kevin D. Shabahang, Hyungwook Yim, Simon Dennis	Beyond Pattern Completion with Short-term Plasticity
10:10	**Hyungwook Yim**, Adam F. Osth, Simon J. Dennis	A systematic re-examination of the list-length effect in recognition memory
10:30		**Coffee**
10:50	Bruce D Burns, Duyi Chi	How well do people fit to Benford's law?
11:10	Duyi Chi, Bruce Burns	Testing the explanations of the first digit phenomenon in number generation
11:30	Erin Walsh, Ginny Sargent	The state of statistics in Knowledge Exchange Transfer questionnaire construction
11:40	Pablo Franco, Karlo Doroc, Nitin Yadav, Peter Bossaerts, Carsten Murawski	Computational cognitive requirements of random decision problems
12:00	**Chris Donkin**, Aba Szollosi	Fit is not a salvageable concept
12:10		**END OF FORMAL SESSIONS**
12.20		**BUSINESS MEETING**

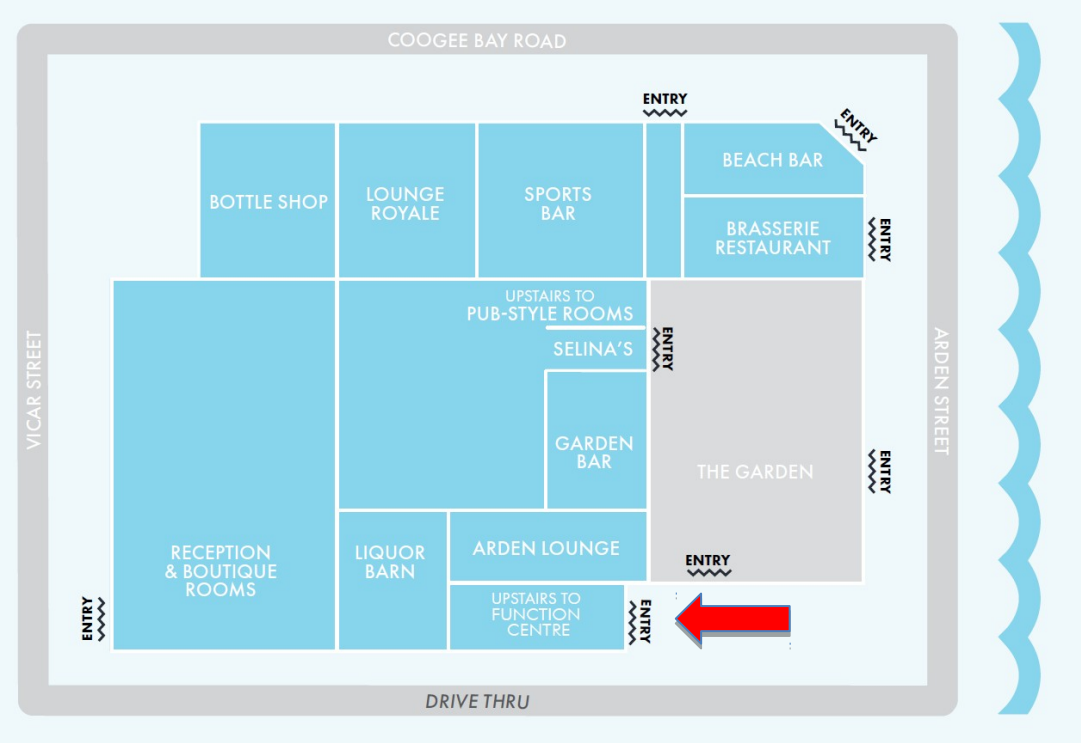
Venue Information

All sessions will be held in the Function Centre of Coogee Bay Hotel, 253 Coogee Bay Rd, Coogee, NSW. (see Map 1 for general location). WiFi will be available in the venue.

Please note that the meeting rooms are on Level 1 – upstairs from the pub and bistro. The main entrance via stairs or elevator is from Arden Street on the south side of the pub. See Red Arrow in Map 2. ***You don't need to go into the pub to enter the conference centre***!

Morning and afternoon tea /coffee for Days 1 and 2, and morning tea/coffee for Day 2, is included in your registration. There are plenty of options for lunch and coffee a short walk from the venue. *Please note however that you CANNOT bring food purchased elsewhere back to the hotel/conference rooms.*





Eating and Drinking

There are many excellent options in Coogee within easy walk of the venue, as well as in the nearby suburbs of Randwick and "The Spot" (Randwick South). Here are some committee members' Coogee favourites.

CLOSE TO THE VENUE

Zambrero Coogee, 231 Coogee Bay Road, Lunch, Dinner all week
Casual Aus-Mex eat in or take-away with several vegetarian options

Sugarcane, 1/56 Carr St. Coogee; Lunch Thursday-Friday, Dinner all week
South-east Asian Cuisine with a very useful cocktail list

A Fish Called Coogee, 229 Coogee Bay Rd, Coogee NSW 2034 Lunch, Dinner all week
Casual eat in or take-away fish and seafood

Osteria Coogee, 31 Alfreda St. Coogee Dinner all week, Lunch Friday
Modern spin on Italian classics. You may need to book…

Bananapalm Vietnamese, 260 Arden St. Coogee, Dinner only
Authentic Southern-Vietnamese

AT THE NORTH END OF THE BEACH

Coogee Pavilion, 169 Dolphin St, Coogee NSW 2034, Lunch, Dinner all week
Large pub complex with a range of dining options. Having drinks on the upstairs balcony overlooking the beach is a great way to relax after a hard day's Maths Psyching

The Lure Seafood, 139 Dolphin St, Coogee NSW 2034, Lunch, Dinner all week
Quality seafood take-away and chippie. Great option if you want to eat in the park.

Conference Dinner (see map 1)

The conference dinner is after the second day of talks (Thursday, 13th February) at Barzura Restaurant. Set menu with vegetarian options. The venue is a short walk from the conference.

Soccer

We will have the traditional cricket game (now soccer) at the end of the first day of the conference (Wednesday, February 12th). At the end of the schedule (~5.00 pm) we will walk to Grant Reserve, which is approximately 10-minute (uphill) walk from the conference venue. For anyone who would prefer not to play, there lots of places nearby to sit and talk.

Abstracts

Where do we go wrong? Dynamic strategy regulation during dynamic, goal-directed decision making

Timothy Ballard, Andrew Neal, Simon Farrell, Andrew Heathcote
The University of Queensland

Goal pursuit involves making a series of interdependent decisions in an attempt to progress towards a performance target. Although much is known about the intra-decision dynamics of one-shot decisions, far less is known about how this process changes over time as people get closer to achieving their goal and/or as a deadline looms. We have developed an extended version of the linear ballistic accumulator model that accounts for the effects that the dynamics of goal pursuit exert on the decision process and provides a normative benchmark for comparing the effectiveness of people's strategies. We test the model using a paradigm in which participants perform a random dot motion discrimination task, gaining one point for correct responses and losing one point for incorrect responses. Their objective was to achieve a certain number of points within a certain timeframe (e.g., at least 24 points in 40 seconds). The results from two experiments suggest that people generally set lower thresholds than would be needed to maximize the likelihood of goal achievement. This is particularly true when stimuli are easy to classify or the person has a lot of time available relative to the number of points they need to accumulate. This over-prioritization of speed relative to accuracy decreases the probability of goal achievement by as much as 30%. These findings illustrate the sensitivity of decision making to the higher order goals of the individual, and provide an important step in the development of a formal theory that describes how these higher level dynamics playout

Competitive Decision Making in Dutch Auction

Murray Bennett, Rachel Mullard, Marc Adam, Mark Steyvers, Ami Eidels, Scott Brown
University of Newcastle

In a Dutch Auction an item is offered for sale at a set maximum price. The price is then gradually lowered over a fixed amount of time until a bidder makes a bid, at which point that bidder is guaranteed the purchase, at the current price. Bidders must tradeoff certainty and price: bid early and you secure the item but pay a premium; bid later and the price is lower but you risk losing the bid to another. These properties make Dutch Auction a

fascinating context for the study of competitive decision making. We conducted a set of experiments in which triplets of participants took part in a computerised Dutch Auction, playing against each other in various conditions. In another experiment, the same participants played individually against a computer opponent. I present our computerised platform and results, including the effects of price-step and price volatility on bidding behaviour.

An Iterated Prospect Theory Model for the Dutch Auction

Scott Brown, Murray Bennett, Rachel Mullard , Ami Eidels
University of Newcastle

This talk follows directly from an earlier talk by Murray Bennett. We present a model for how people make the decision to buy during a Dutch auction. The model uses a very standard version of Prospect Theory, and iterates it. At each moment of the auction, the buyer is faced with a decision that can be framed in the classical Prospect Theory manner: a certain option (buy now!) or a risky option (wait a little longer for the price to fall, and hope that no-one else buys before then). We show that this model reproduces the basic phenomena of the task, and also provides a useful framework for investigating interesting questions about auction psychology.

How well do people fit to Benford’s law?

Bruce D Burns, Duyi Chi
The University of Sydney

In recent years there has been an explosion of interest in Benford’s law as a property of data with the online Benford’s law bibliography containing more than 500 papers published in mathematics and statistics; 400 in Finance and Accounting; 130 in Science and Psychology; 90 in computer and digital science, 60 in Politics and Economics, and 15 in clinical or medical settings. However few studies have asked whether the numbers people generate fit to Benford’s law? Early studies that asked people to generate random numbers did not find evidence. When Diekmann (2007) and Burns & Krieger (2015) asked people to generate meaningful, nonarbitrary numbers then their first digits were a reasonable fit to Benford’s law. Across a set of studies we found estimates of effect size of .40 to .60 using weighted linear contrast. So Benford’s law pattern of first-digit frequencies explains a substantial amount of variance in the first-digits people generate, but it is evidence of a “Benford bias” rather than a precise fit to Benford’s law. We found the same bias when people tried to predict the outputs of a linear

system suggesting that Benford bias may be a ubiquitous influence on the judgments we base on numbers we generate.

Investigating the time course of feature processing in consumer-like choices

Jon-Paul Cavallaro, Guy Hawkins
University of Newcastle

Hypothetical choice scenarios provide insight into a consumer's decision-making process when considering products or services. The perceptual decision-making literature contains many examples of the importance of decision time and the impact that time pressure has on choice processes. Yet the impact of time on choice is rarely explicitly tested in the consumer choice literature. To address this, we combined methods from experimental psychology and consumer choice research, implementing a speed-accuracy tradeoff (SAT) manipulation in a preferential choice scenario. We extend the use of the response-signal task to a multi-dimensional stimulus in the form of a discrete choice experiment (DCE) for consumer-like products. We modelled SAT functions for the preferential choices to examine when different sources of information become 'cognitively online' to influence choices. We use these results to infer which product attributes influence the choices people make, and when.

Testing the explanations of the first digit phenomenon in number generation

Duyi Chi, Bruce Burns
The University of Sydney

Recent research (e.g., Burns & Krygier, 2015) suggests that people can spontaneously exhibit a strong bias towards the smaller leading digits (i.e., 1, 2) when generating numbers. This approximates Benford's law, a well-established phenomenon that the frequencies of first digits from many naturally occurring datasets approximate a log distribution. We have replicated this finding a number of times when people generate unknown values for meaningful questions they know little about, but why this bias emerges is unclear. We have proposed and tested two hypotheses that may explain the bias: the Recognition Hypothesis and the Integration Hypothesis. The Recognition Hypothesis suggests that implicit learning due to exposure to Benford's law leads us to be biased by it in our own behaviour. However a set of experiments have found no evidence that people prefer numbers with low first digits even when the same participants tend to generate such numbers more often. The Integration Hypothesis is inspired by mathematical conjectures about Benford's law that it is a consequence of aggregating information, so it suggests that the more

information sources processed the stronger the first digit phenomenon psychologically. Consistent with this we have found that when estimating the number of jelly beans in a jar the aggregation estimation showed a stronger bias towards the smaller first digit than individual estimation. Nevertheless, the manipulation of increasing visual information when asking participants to estimate the numbers of dots in a stimulus failed to enhance such the bias. Therefore, more research is needed to fully understand the utility of the Integration Hypothesis as a reliable explanation.

Modelling decision rules in consumer choice with the LBA

Gavin Cooper, Guy Hawkins, Ami Eidels
University of Newcastle, Callaghan, Australia

Systems Factorial Technology (SFT) is a powerful tool for discriminating between processing architectures, serial vs parallel, and decision rules, exhaustive vs self-terminating, in individual participant data. One limitation of SFT is a large number of trials and low error rates necessary for this discrimination. In a recent work, Cox & Criss (2019) presented parametric methods to extend the traditional analyses of SFT. One benefit of this hierarchical Bayesian approach is a greatly reduced number of trials necessary to provide insight into the possible architectures used by participants. I will discuss the application of these parametric extensions of SFT to a dataset of multi-attribute consumer choices.

Using signed difference analysis to differentiate between reasoning processes

Nicole Cruz, Brett K. Hayes, Anagha Kaluve, Jasmine Choi-Christou, John Dunn, Rachel Stephens
UNSW

A major debate in the psychology of reasoning concerns whether people evaluate inferences on a single scale of inference quality (single process theories), or instead use distinct criteria to judge the strength of deductive and inductive inferences (dual process theories). Ways of addressing this debate have traditionally included searching for double dissociations in factorial experiments, in which the relationships between variables are assumed to be linear. However, linearity is not always met in practice. An alternative is to use signed difference analysis (SDA), a generalisation of state-trace analysis (STA) in a signal detection theory framework. SDA makes assumptions only about monotonic relations between variables but not about their distributional forms. Unlike STA, it can also be applied to models with more than one parameter. We introduce SDA and apply it to a

new experiment in which people made explicit judgments about the logical validity of arguments as well as “intuitive” judgments of validity (e.g., rating the perceptual brightness of arguments). We replicate previous findings that people are sensitive to the logical validity of arguments in both explicit and implicit logic tasks. We compare the fit of different SDA models to these data, and discuss the results in the context of what we can learn from SDA when studying the number of cognitive processes that underlie reasoning.

Understanding Memory for WHERE using Smartphone Data

Simon Dennis, Elizabeth Laliberte, Hyungwook Yim, Benjamin Stone
University of Melbourne

In 1984, Ronald Cotton was convicted for rape and burglary. He was sentenced to life + 50 years. In 1995, he was released having served over 10 years in prison. When Cotton was interrogated he provided a false alibi. Rather than report where he had been at the time of the crime, Cotton recalled where he had been the week before. A primary challenge for alibi generation research is establishing the ground truth of the real world events of interest. We used a smartphone app to record data on participants (N=57) for a month prior to a memory test. The app captured their accelerometry continuously and their GPS location and sound environment every ten minutes. After a week retention interval, we presented participants with a series of trials which asked them to identify where they were at a given time from among four alternatives. Participants were incorrect 36% of the time (SD = 16%). Furthermore, our forced choice procedure allowed us to conduct a conditional logit analysis to assess the relative importance of different aspects of the events to the decision process. The Cotton example suggests that participants might also confuse days across weeks and we found strong evidence of this kind of error. In addition, people often confused weeks in general and also hours across days. Similarity of location induced more errors than similarity of sound environments or movement types.

Fit is not a salvageable concept: Twenty years after Roberts and Pashler (2000)

Chris Donkin, Aba Szollosi

UNSW

Twenty years ago, Roberts and Pashler (2000) questioned the utility of a good fit. In the proceeding decades, researchers have proposed a number of solutions that center around statistical implementations of the flexibility (or complexity) of mathematical models. All solutions of this type do not solve

the problem because they focus on the flexibility of a particular model specification, and not the motivating theory. For science to progress, however, it is theories that must be hard to vary. Theories are constrained by their implications. One such implication is which mathematical implementations are appropriate, an issue that is not incorporated into statistical indices such as Bayes factors. Rather than attempting to salvage a statistical solution, we should focus on ensuring that the theoretical motivations for the assumptions in our mathematical models are clear, so that our theories are hard to vary.

The Geology of Eyewitness Memory

John Dunn, Kym McCormick, David Kellen, Carolyn Semmler
University of Western Australia

Eyewitness memory is traditionally tested using an n-person lineup task. Alternative ways of testing memory are the m-alternative forced choice (m-AFC) task, where m ranges from 2 to n, and a ranking task in which the n lineup items, including the target, are ranked from most to least likely to be the target. Receiver Operating Characteristic (ROC) curves can be derived from each of these tasks. Our aim was to examine lineup memory in terms of these ROCs. First we describe how ROCs are generated from each of these tasks and then compare them using data from two experiments. In experiment 1, N = 2,026 online participants studied a short video and then given an m-AFC task with each participant allocated to a single m in the range 2 to 7. In experiment 2, participants viewed the same video and presented with an 8-person lineup. For N = 722 participants, the lineup contained the target, for N = 331 participants, the lineup did not contain the target. All participants were first given an identification task and asked to rate their confidence in their decision and were then asked to rank the items in order of familiarity. We constructed ROC curves from each of these data sets and will report on their similarities and differences. We show through simple simulation how the main patterns of these ROCs can be reproduced. We conclude that memory for the target is surprisingly invariant across different retrieval tasks.

Computational cognitive requirements of random decision problems

Pablo Franco, Karlo Doroc, Nitin Yadav, Peter Bossaerts, Carsten Murawski
University of Melbourne

Real life instances of problems can be modelled as emerging from a random generating process. These are referred to as random instances. Previous studies have found that for electronic computers the computational

requirements of solving an instance of a problem are related to a specific set of features of the problem. This mapping has been shown to apply to a multitude of problems and is referred to as Instance Complexity (IC). Overall, there is ample evidence suggesting that this theory is a general framework for electronic computers. Additionally, it has been shown that this pattern holds when humans solve the knapsack problem. However, it remains an open question whether IC is a generalisable measure, for humans, of the expected computational requirements of solving a problem. For this purpose, we ran a set of experiments in which human participants solved a set of instances of one of two, widely studied, computational problems (Traveling Salesperson and the Boolean Satisfiability). Instances varied in their IC. We found that, in line with our hypothesis, IC had a negative effect on human performance in both problems. In particular, on instances with high IC participants exerted more effort and had a lower accuracy rate. Our results suggest that IC is a general measure of expected difficulty of random instances. Since IC might be used as an approximation to the expected computational requirements of the task, IC could be a crucial component in the cognitive resource allocation process in the brain.

Symbolic wheel of fortune: confusion analysis and the mental representations of digits across cultures

Paul Garrett, Zachary Howard, Yu-Tzu Hsieh, Murray Bennett, Cheng-Ta Yeng, Daniel Little, Ami Eidels
University of Newcastle, Australia

People express quantities using a remarkably small set of units and digits. Confusing digits could be costly, and not all confusions are the same; confusing a price tag of 2 dollars with 9 dollars is naturally more costly than confusing 2 with 3. Confusion patterns are intimately related to the distances between our mental representations and hypothetical internal symbols said to represent ‘real’ external stimuli. The distance between the mental representations of two digits could be determined by their numerical proximity or visual similarity, and this distance may be influenced by our expertise. For example, numerical proximity is only relevant after learning the numerical order. Through a set of identification experiments, we applied multidimensional scaling and investigated the mental representations of familiar and unfamiliar digits (4 sets: Arabic, Chinese, Thai, and non-symbolic dots) in a Chinese and an English speaking cohort. We controlled for undesired effects of response bias using Luce’s choice model and compared group-level mental representations using Procrustes analyses. Across cohorts, our findings showed Arabic, Chinese and Thai digits were confused due to perceptual similarities. Non-symbolic dots were confused due to numerical and perceptual similarities. Across cohorts, the mental representations of Arabic, Thai, and dot digits were similar, whereas the

mental representations of Chinese digits differed drastically. This difference appeared to reflect the difference in cohort expertise.

Bayesian Inference for Multidimensional Scaling Representations with Psychologically-Interpretable Metrics

Quentin F. Gronau, Michael D. Lee
Department of Psychological Methods, University of Amsterdam

Multidimensional scaling (MDS) models represent stimuli as points in a space consisting of a number of psychological dimensions, such that the distance between pairs of points corresponds to the dissimilarity between the stimuli. Two fundamental challenges in inferring MDS representations from data involve inferring the appropriate number of dimensions, and the metric structure of the space used to measure distance. We approach both challenges as Bayesian model selection problems. Treating MDS as a generative model, we define priors needed for model identifiability under metrics corresponding to psychologically (partially) separable and psychologically integral stimulus domains. We use differential evolution Markov chain Monte Carlo (DE-MCMC) for parameter inference, and Warp-III bridge sampling for model selection. We apply these methods to a series of previous data sets, which collectively test the ability of the procedure to infer an appropriate dimensionality and to find an appropriate metric structure. We demonstrate that our method produces sensible results, but note a number of remaining technical challenges that need to be solved before the method can easily and generally be applied.

Latent class discovery for cognitive models

Guy Hawkins, David Gunawan, Minh-Ngoc Tran, Robert Kohn, Mark Steyvers, Scott Brown
University of Newcastle

Hierarchical Bayesian analysis permits simultaneous inference about group- and participant-level parameters of cognitive models. The participant-level parameters allow for individual differences though most standard hierarchical analyses assume participants are all sampled from a single population distribution. This overlooks the possibility that latent cognitive processing may differ from one individual to the next in a manner that is better described by the existence of more than one population rather than random parameter variability within a single population. In this project we developed an augmented hierarchical modelling scheme that estimates group- and participant-level parameters while flexibly sampling new population distributions as the data require them, through an infinite mixture model. We demonstrate the modelling approach with an application

of a choice response time model to behavioural data from over 1000 participants performing two working memory tasks in the Human Connectome Project. We discuss the potential of the new approach to explore emergent individual and group differences in cognitive modelling.

When is absence of evidence evidence of absence? And who can tell the difference?

Brett Hayes, Yuki Wen, Dani Navarro
UNSW, Sydney

Inductive inference involves generalizing from what is known about a sample of instances to novel cases. In many environments, the instances we observe are sampled selectively – we are only able to observe some types of instances but not others. The reasons why some instances are excluded can have a major impact on how we generalize from samples. Our work uses experiments on property induction and a Bayesian modeling framework to examine, a) whether people show sensitivity to selective sampling in inductive inference, and b) whether such sensitivity varies between individuals. In several experiments (N = 748), participants observed a sample of instances with a novel property (e.g., “has plaxium”) and then inferred whether the property generalized to transfer items. In each experiment, two groups of participants observed the same training sample but were given different “sampling frames” or descriptions of how sample instances were selected. Each experiment found narrower property generalization when the sample instances were selected because they shared a common property (property sampling) than when they were selected because they belonged to the same category (category sampling). Follow-up studies (N = 122) using a within-subjects sampling frames paradigm, found that although a majority of individuals are sensitive to the effects of selective sampling on property inference, a substantial minority are not. We discuss the implications of these findings for previous claims that people are generally insensitive to sample selection processes (meta-cognitive myopia).

What do confidence and response time tell us about accuracy (and a model).

Andrew Heathcote, Angus Reynolds, Peter Kvam, Adam Osth
University of Tasmania

Fast and confident responses are usually accurate, but not always. How much information do confidence ratings and response time (RT) give us about whether a response is correct, and how does that differ among people

and between choices? We use variants of Receiver Operating Characteristics based on confidence and RT and conditioned on choices to provide a non-parametric answer these questions and present a model that provides a parametric explanation of these relationships.

Questioning the assumptions of the Capacity Coefficient

Zach Howard, Paul Garrett, Daniel Little, Ami Eidels, James Townsend
University of Newcastle

Systems Factorial Technology (SFT) is a popular framework for that has been used to investigate processing capacity across many psychological domains over the past 25+ years. To date, it had been assumed that no processing resources are used for sources in which no signal has been presented (i.e., in a location that can contain a signal but does not on a given trial). Hence, response times are assumed to be driven by only the "signal-containing'' location or locations. This assumption is critical to the underlying mathematics of the capacity coefficient measure of SFT. In this presentation, we show that stimulus locations influence response times even when they contain no signal, and that this influence has repercussions for the interpretation of processing capacity under the SFT framework, particularly in conjunctive (AND) tasks - where positive responses require detection of signals in multiple locations. We propose a modification to the AND task requiring participants to fully identify both target locations on all trials. This modification allows a new coefficient to be derived. We apply the new coefficient to novel experimental data and resolve a previously reported empirical paradox, where observed capacity was limited in an OR detection task but super capacity in an AND detection task. Hence, previously reported differences in processing capacity between OR and AND task designs are likely to have been spurious.

Modelling decisions of the multiple object tracking task

Reilly Innes, Caroline Kuhne
University of Newcastle

Decision making is a vital aspect of our everyday functioning, from simple perceptual demands to more complex and meaningful decisions. The strategy adopted to make such decisions can be simplified as balancing the speed of the decision against the level of caution. Strategic differences, such as a more cautious approach, can be observed via modelling simple

decision making data to tease apart processing differences from strategy. Our study used a multiple object tracking task where student participants and a highly skilled military group were compared on their ability to track several items at once. Using the linear ballistic accumulator (LBA), we show the underpinnings of how two groups differ in performance. Results showed a large difference between the groups on accuracy, with an interaction effect observed between groups and level of difficulty in response times, where RAAF group response times slowed at a greater rate than the student group. Model results indicated the RAAF group were more cautious than students, and tended to have faster processing speeds. Our findings show the strength of sequential sampling models in explaining decision making strategies, as well as a first attempt at modelling decisions from the multiple object tracking task.

A causal Bayes net framework for reasoning from non-random samples

Charles Kemp, Brett Hayes, Danielle Navarro
University of Melbourne

Knowing how a sample was generated affects what inferences can be drawn from it. Previous work suggests that reasoners are able to adjust for the sampling process in some cases but not others. We present a Bayes net framework that can be used to capture many different problems requiring reasoning from non-random samples and to characterize the normative solution to each problem. We compare this normative model to an alternative that computes statistics over the observed data without adjusting for sampling, and present two experiments suggesting that people often adjust their inferences in the direction predicted by the normative model.

Multivariate Normal Priors for Evidence Accumulation Models

Caroline Kuhne, Scott Brown, Andrew Heathcote, Dora Matzke
University of Newcastle

Bayesian estimation is able to naturally take prior knowledge into account. Despite this, the vast majority of applications use “uninformed” priors. Methods that do not include informed priors may waste useful information and lead to implausible parameter estimates. Further, the accessibility of informed priors may be another way to reduce research degrees of freedom. We develop informed priors for evidence accumulation models by collecting hundreds of previous parameter estimates. The prior we have specified is a

multivariate normal, with a mean, standard deviation, and a correlation matrix for parameters at the individual subject level. The multivariate normal addresses the issue of correlation between parameters, allows for a distribution for each parameter, and also opens up the possibility to calculate effects on parameters from experimental manipulations. However, the overarching goal of this project is to provide a growing repository of open science that includes data sets, past posterior estimates, and the distributions of effects on these parameters for different experimental manipulations. The repository will be available for researchers to utilise different research methods to advocate for, and use, different informed priors, with different distributions. As this repository continues to grow, calculated priors will more adequately represent cumulative science.

A mixture of experts in associative generalization

Jessica C. Lee, Peter F. Lovibond, Brett K. Hayes, Stephan Lewandowsky
UNSW Sydney

After learning that one stimulus predicts an outcome (e.g., an aqua rectangle leads to shock) and a very similar stimulus predicts no outcome (e.g., a slightly greener rectangle leads to no shock), some participants generalize the predictive relationship on the basis of similarity, while others generalize on the basis of the relational difference between the stimuli. To date, these individual differences in generalization rules /hypotheses have remained unexplored in associative learning. Here, we present evidence that a given individual simultaneously entertains belief in both “similarity” and “relational” rules, and generalizes using a mixture of these strategies. Using a “mixture of experts” modelling framework, we show that considering multiple rules predicts generalization better than a single rule, and that participants are more likely to switch between rules, rather than average between their implied responses. The results suggest that in humans, classic generalization phenomena such as peak shift may actually reflect a mixture of different individual generalization rules.

A model of triadic comparison applied to the problem of understanding semantic memory impairment

Michael D. Lee, Holly A. Westfall
University of California Irvine

In a triadic comparison task, people are presented with three items and must choose the odd-one-out. This task is often used to understand how people perceive the similarities between the items. One important application is in assessing the effects of cognitive impairment on semantic memory. Some cognitive assessments used in clinical settings involve a triadic comparison task of animal names. In this work we develop a cognitive model of triadic comparison as a means of assessing semantic memory decline in Alzheimer's patients. The model assumes people have a feature-based representation of animals, and assess similarity in terms of the weights of common features. The model also assumes people use a form of the Luce-choice rule to make triadic comparison decisions. We apply our model to a clinical data set including observations from 3,602 clinical patients between 16 and 104 years of age, independently diagnosed in terms of the functional impairment of their memory. In particular, we use our model to compare competing hypotheses about changes in semantic impairment. The three hypotheses are that people's triadic comparison choices change because (a) they use different features to represent the animals as they become impaired, (b) they assign different weights or saliencies to the features, or (c) their access to an intact and stable representation becomes noisier. We find strong evidence against the first hypothesis, but supporting evidence in favor of the other two hypotheses.

To cue or not to cue? Seeking and integrating external information into choice

Garston Liang, Jennifer Sloane, Chris Donkin, Ben Newell
UNSW, Sydney

Decision aids have become near-ubiquitous across a number of applied domains (e.g. finance, medical diagnosis). In this experimental talk, I describe two experiments in which we instantiate a decision-aid as a partially predictive cue that has a 70% chance of providing the correct response. Participants are aware of this performance level and could request the cue on any trial of a binary-choice task. However, the cue's response would only be useful on hard trials and would likely be misleading on easy trials. We manipulated training, block feedback, and exposure to the cue's performance to examine a) when participants request the cue's answer, and b) whether they respond consistently with it. We find that participants distinguished the difficulty levels, and each manipulation increased the number of requests for the harder trials. A basic diffusion model analysis suggests the cue affects the drift rate rather than the start point parameter. Alternative ways of modeling the results will be discussed.

Wanting to know, or not wanting to not know? A comparison of information seeking models

Shi Xian Liew, Danielle Navarro, Benjamin Newell
UNSW

The preference for non-instrumental information (e.g., choosing to receive advance information on a gamble after it has been played) has been widely demonstrated in many domains. Key aspects of such behaviour have been explained by anticipation-based models as well as models of uncertainty aversion. At a fundamental level, these models can appear very similar–the former places positive value on knowing and the latter assumes negative value on not knowing. In this talk we explore the extent of the equivalency of these accounts by describing how these models differ at both qualitative and quantitative levels when fit to a series of behavioural experiments. Beyond contrasting their core assumptions, we also discuss how the models vary in their treatment of temporal discounting, outcome probability, and learning, in addition to how future experiments can formally disentangle these competing accounts of information seeking.

Team Coordination: A Multiple Goal Pursuit Problem

Erin Lloyd, Prof Andrew Neal, Dr Timothy Ballard
The University of Queensland

As work becomes more complex, organisations are increasingly making use of teams. Teamwork is typically split, giving each member responsibility for a different component of the overall task. Each member, therefore, has one or more individual goals in addition to the shared team goal, and must continually evaluate how best to allocate their limited resources across these multiple goals. The aim of the current work is to inform our understanding of how teams coordinate their actions in the pursuit of both shared and individual goals to determine prioritisation. These processes are examined using a computer-based experimental task. Set in outer space, the goal of the simulated game is to protect assets from harm posed by incoming threats. Completed in both an individual and a team condition, participants pursue multiple goals concurrently, dividing their attention and resources between them, allowing us to examine patterns of behaviour and prioritisation choices that unfold over time. Preliminary data analysis of Study 1 shows that people behave differently when working as part of a team as opposed to individually, with team members tending to be more strategic in their decisions and considerate of the needs of their teammates. Further analysis will involve computational modelling, with the aim of extending existing models of multiple-goal pursuit, to a team level, producing a process model which describes the mechanisms by which teams coordinate goal prioritisation.

Nudging at drift rates: Using the diffusion model to study effects of probabilistic feedback on perceptual decisions

Steson Lo, Jeremy Ngo, Chris Donkin, Ben Newell
The University of New South Wales

Nudges are commonly defined as small tweaks to the environment causing massive change in human decisions. For perceptual decisions, previous research suggested that stimulus ambiguity and stimulus base-rates can serve as effective behavioural nudges. Extending this earlier work, the present study manipulated base-rate, such that 30%, 50% or 70% of stimuli should be classified as belonged to one or another colour category. The association between stimulus and response was also varied probabilistically, such that 100%, 60% or 50% of feedback agreed, on average, with the colour dominating each stimulus. Based on this design, results from an experiment with undergraduates were analysed using a hierarchical diffusion model. As expected, changes to base-rate impacted start point values by reducing the amount of evidence required for the modal response. Although base-rate also influenced values for drift rate, the effect of this variable was larger when the probability of receiving stimulus congruent feedback was low. Together, these results suggest that information from a

nudge was gradually accumulated as evidence when participants cannot quickly reach a decision using the perceptual stimulus.

Using latent network structures to link neurophysiology and behaviour

Alexander Ly, Udo Boehm, Simon Kucharsky, Botond Szabo, Dora Matzke, Andrew Heathcote
University of Amsterdam and CWI

The goal of model-based cognitive neuroscience is to explain how neurophysiological processes implement latent cognitive processes in the brain. Cognitive models play an important role in this endeavour as they use observable behaviour to derive quantitative descriptions of cognitive processes in the form of model parameters. One major challenge is to discover the latent network structure that links low-dimensional cognitive model parameters to high-dimensional physiological data. The discovery of this latent network structure is complicated by the uncertainty that is involved in the (a) estimation of cognitive model parameters; and (b) generalisation of network structures discovered in a small sample of participants to the underlying population. In this presentation I elaborate on a fast and easily scalable method that allows for the discovery of latent networks that accounts for both sources of uncertainty. This is on-going work to extend the plausible values method, which uses posteriors draws of a parameter as an observation, developed in Ly et al. (2017) to the multivariate set-up.

How recalling values from the edges of the distribution supports evaluation and choice

Alice Mason, Christopher Maddan, Marcia Spetch, Elliot Ludvig
University of Warwick

When making risky decisions based on past experience, people tend to better remember the most extreme outcomes (highest and lowest) in a decision context. This memory bias can lead to people being more risk-seeking for relative gains than relative losses. In an extension of previous work, we examine how memory for individual values influences choice. Each payoff is associated with a distribution of outcomes meaning that we can collect free-recall data for each of the options. We also address to what extent memory for extreme outcomes is dependent upon people making preferential choices. Across experiments, recall for specific values was poor, but participants nonetheless generated values consistent with the overweighting of the extremes. When asked to estimate the value of each outcome, participants underestimated the low risky and overestimated the

high risky options. The results suggests people are not sampling veridically from memory but have generated distributions of plausible outcomes.

Forced choice eyewitness memory: Testing the assumption of random utility

Kym M McCormick, John C. Dunn, Carolyn Semmler, David Kellen
The University of Adelaide

Eyewitness memory research is moving away from intuitive theorising and towards the application signal detection theory (SDT) models of eyewitness identification. These models assume that eyewitness identification decisions are based on evidence generated through a stochastic process, with resulting choice probabilities having a random utility representation. We tested this basic assumption by evaluating a set of eyewitness n- alternative forced-choice (n-AFC) probabilities against the corresponding set of Block-Marschak inequalities, the satisfaction of which is both a requirement for, and an indicator of, random utility representation (Block, H. D. & Marschak, J. (1960). Random orderings and stochastic theories of response. In I. Olkin, S. Ghurye, W. Hoeffding, M. Madow, & H. Mann (Eds.), Contributions to probability and statistics (pp. 97-132). Stanford: Stanford University Press; Falmagne, J.-C. (1978). A representation theorem for finite random scale systems. Journal of Mathematical Psychology, 18(1), 52-72). To achieve this, we employed a 1 x 6 between-subject factorial design, manipulating lineup size (n ={2,…,7}). Participants (N = 2,026) were recruited from Amazon Mechanical Turk and completed a single n-alternative forced choice (AFC) eyewitness identification task after watching a short video. Empirical choice probabilities were not significantly different from those generated by a best-fitting set of conforming points), providing support for the random utility representation account.

On the Link Between Risk Preferences and Cognitive Abilities: A Meta-Analysis

Tehilla Mechera-Ostrovsky, Steven Heinke, Sandra Andraszewicz, Jörg Rieskamp
University of Basel, Switzerland & University of New South Wales, Australia

When making decisions under risk, people must evaluate the consequences of their decisions and the associated probabilities. Past work argues that these evaluations depend on people’s cognitive abilities, which affect risk-taking behavior. We test this using two hypotheses. The first posits that risk preferences are linked to cognitive abilities such that people with higher (or lower) cognitive abilities are less (or more) likely to take risks. The alternative assumes that people with low cognitive abilities are simply more likely to make unsystematic mistakes and therefore show more random

choice behavior. In the majority of studies, we find the architecture of the risk preference elicitation tasks is such that random choice behavior leads to a classification of risk aversion. This means that the observed correlation between cognitive abilities and risk preferences might be spurious and mediated by the bias for random choice behavior in the task architecture. In the current work, we conduct a systematic literature search and run a meta-analysis of 30 studies. We examine different aspects of choice architecture and identify possible causes of previous, divergent findings. We find no link between risk-taking and cognitive abilities. More importantly, we show that the correlation between risk aversion and cognitive abilities is heavily affected whenever the choice set's architecture is biased by random choice behavior. We introduce a generalized scheme to measure how elicitation methods are affected by random choice (i.e. Random-Choice Risk-Taking (RCRT) bias), discuss its implications, and make recommendations for future work utilizing risk elicitation measures.

The Effect of Distractors and Bilateral Presentation on Processing Architecture in Categorization Decisions

Sarah Moneer, Eleanore Williams, Daniel Little
Melbourne School of Psychological Sciences, University of Melbourne

Several studies have demonstrated that the ease with which attention can be deployed to relevant locations or features of objects varies as a function of target location and the presence of distractors; however, it is unclear how the presentation of distractors between two targets affects information-processing architecture. Further, although differences in attentional deployment have been found for objects separated along vertical or horizontal meridians, it is unknown whether these hemifield effects extend to other types of decisions. We have previously shown that simple separable dimensions (saturation of a red hue and an oriented line, presented in different semicircles) can be processed in parallel even when they are spatially separated. This is in contrast to separable dimensions presented in schematic illustrations of lamps, which are processed in serial. We therefore adapted our design to present saturation and orientation in schematic illustrations of rockets and varied the distractors presented between them across three conditions. Processing architecture was diagnosed using Systems Factorial Technology and parametric model fitting. We also varied the rotation of the rockets between blocks to determine whether bilateral presentation (i.e., presentation in opposite hemifields) confers an advantage to processing in categorization decisions. We predominantly found serial processing when the dimensions were vertically separated, and a mixture of serial and parallel processing when they were presented bilaterally. These results indicate differences in processing architecture between alignments

that is more pronounced in the presence of distractors, and importantly, suggest that a bilateral presentation advantage exists in categorization.

The Effects of Competition Type on Goal-Directed Decision Making

Andrew Morgan, Timothy Ballard, Andrew Neal
University of Queensland

While competition has been a focus of psychology for over 100 years, we still have a lot to learn about it. The way people respond to different types of competition and how their responses can change over time have been overlooked in past research. Different types of competitions could lead us to expend more effort or enlist different strategies as we try to win. We have taken a dynamic approach to assess how these behaviours change over time as different types of competitions progress. Data was collected using a simple cognitive task that had participants classify a cloud of dots as moving left or right. Participants received a point for each correct response and lost a point for each incorrect one. In the race condition, participants competed against a simulated opponent to reach a certain score first, while in the tournament condition, they competed to have the higher score by the end of a time limit. In the goal pursuit condition, participants had no opponent, and had to reach a certain score before the end of a time limit. The “do your best” condition tasked participants with scoring as many points as they could before time ran out. Preliminary data analysis shows differences in accuracy between the goal pursuit and do-your-best conditions. When in the individual goal pursuit condition, participants became less accurate over time when their scores were higher. There was no such change in accuracy in the “do your best” condition. Response time increased over time in both conditions when scores were low, but were quicker in the individual goal pursuit condition when scores were high. Between the race and tournament conditions, there was a difference in response time, but not accuracy. In the tournament condition, people made more rapid decisions in response to increases in the opponent’s score, as well as their own. In the race condition, people responded more strongly to the discrepancy between the participant and opponent scores. Further analysis will involve fitting the data to the Linear Ballistic Accumulator to examine how the rate of evidence accumulation and response threshold are impacted by competition type. This research unifies previously separate bodies of research within social and cognitive psychology and advances our knowledge about people respond to competition.

A Novel Methodology For Manipulating The Speed-Accuracy Trade-Off

Thomas Narraway, Dr Guy Hawkins, Professor Scott Brown,
University of Newcastle

The current literature on decision making suggests that emphasis on speed or accuracy is change in decision threshold. Current methodologies use explicit verbal and written instructions, regular feedback, and time-based or visual penalties to manipulate the speed-accuracy trade-off. In this project, we explore current manipulations of the SAT, making use of a novel methodology. We manipulate the SAT without using explicit verbal instructions or regular feedback, instead using a random deadline approach.

Using the Timed Racing Diffusion Model to account for responses to stimuli with limited information

Jeremy Ngo, Chris Donkin
UNSW, Sydney

There are two main types of bias in simple decision tasks, prior bias and dynamic bias. Prior bias is an expectancy to make a certain response that is present before the accumulation of evidence, and dynamic bias is the unbalanced accumulation of evidence in favour of a biased response. We presented participants with a two-alternative forced-choice brightness discrimination task in which we varied the presentation length of the stimuli and the relative frequency of each stimulus. Behavioural measures and parameter estimates from a Bayesian hierarchical version of the simple diffusion model suggest that prior bias is present throughout all the stimulus presentation lengths, but dynamic bias becomes more prevalent as stimulus presentation length decreases. We compare the Diffusion Decision Model's qualitative account of our data with an account offered by the Timed Racing Diffusion Model (TRDM) and discuss the advantages and disadvantages of the TRDM with respect to our data.

Integration of semantic and orthographic information in episodic memory

Adam F. Osth, Daniel Wraith-Franck
University of Melbourne

Memory models posit that retrieval in recognition operates by globally matching the probe cue against each of the representations in memory to produce a measure of global similarity. However, to date many models have not adopted principled representations of items and have been tested with manipulations of study list composition. While some models have used semantic representations from semantic space models, little work has explored the consequences of employing perceptual representations, and it

remains unknown how perceptual and semantic representations are combined, which is the focus of the present work. To accomplish this goal, I first constructed orthographic representations of words using holographic reduced representations (HRRs) constructed from positions, bigrams, or hybrid representations that combine the two. I tested these representations by computing the global similarity, the average of the probe-item similarities from the orthographic representations, and generated predictions about choice and latency using the diffusion decision model (Ratcliff, 1978) implemented in a hierarchical Bayesian framework. The winning representations were hybrid representations, which often outperformed inverse Levenshtein distances in explaining perceptual confusions in recognition memory. Second, I combined such orthographic representations with three different semantic representations for individual words: those from the BEAGLE, GLoVe, and Word2Vec models. Using each of these semantic representations, I explored whether orthographic and semantic similarities were combined additively (which would result from concenation or vector addition), multiplicatively (which would result from outer products, convolutions, or multiplicative feature matching), or were added together and subsequently cubed (as in the Minerva 2 model, Hintzman, 1988). A total of ten datasets revealed strong preference for additive combination of semantic and perceptual similarities. Parameter estimates revealed that orthographic representations were weighted about equally to semantic representations.

Polarisation, learning, and trust in populations of Bayesian agents

Amy Perfors, Keith Ransom
University of Melbourne

Simulating information flow through populations of Bayesian agents is a powerful technique for understanding some of the environmental and cognitive mechanisms that underlie aspects of information flow and cultural change. In this work I build on previous research in which we showed that the formation of “echo chambers” (subpopulations of agents who do not trust and do not talk to each other) is nearly inevitable given sufficient population heterogeneity, even when the agents are fully Bayesian and constrained not to lie (Perfors & Navarro, 2019). Here I present follow-up work suggesting that polarisation is very robust and emerges under a wide set of assumptions about the goals and cognitive abilities of the agents. Moreover, while polarisation can be “broken” if the population has access to an agreed-upon ground truth, this does not occur if access is too unevenly distributed. I will conclude with (very preliminary) work exploring how social media architectures may affect these dynamics.

The role of utility in semantic categorisation

Keith Ransom, Charles Kemp, Amy Perfors
University of Melbourne

When it comes to semantic categorisation, “carving nature at is joints” is more complex than the allusion suggests. Semantic joints may not be easily discernible, and sub-division along continuous dimensions may be warranted. We examine these difficulties in an experimental setting, where participants are required to communicate about two perceptual object classes using three category labels. By manipulating the observed utility of objects, we ask whether people incorporate such information into category design. And if so, how do they go about it? We consider this latter question by analogy with information theory (and rate distortion theory, in particular). We conceptualise the location of category boundaries as the outcome of an iterative algorithm whereby speakers and listeners evolve a stable balance between the brevity and utility of communication. We present a mathematical formulation of this model and explore its predictions in light of our empirical work. We find that while people are sensitive to the shape of an explicit utility function governing the outcome of imprecise communication, such concerns may be subordinate to the implicit utility function that overall generalisability suggests, as well as the pragmatic concerns of speaker and listener coordination. We consider the implications of our findings for theories of semantic categorisation.

Identifying Sources of Noise in Decision-Making

Roger Ratcliff, Philip L. Smith
The Ohio State University

The random dot motion discrimination task has become one of the most-used tasks for examining decision-making. In modeling decision-making using the motion discrimination task, there has been a bias, especially in some parts of the neuroscience community, to assume that the evidence used to drive the decision process is identical on every trial. This assumes that all variability in processing comes from within-trial noise. In contrast, the more standard psychological diffusion model assumes that the various components of processing vary from trial to trial. It is ironic that the neuroscience community embraces signal detection theory (also strongly supported in psychology) which assumes variability in perceptual strength across trials. In this presentation, we present data from a motion discrimination task which shows an unusual pattern of results. We claim that the only way to address these data is by assuming across-trial variability in drift rate. We relate these results to double pass manipulations

and EEG data that also show the need for trial-to-trial variability in drift rate.

The Use of Response Times as a Critical Test of Risky Decision-making Theories

Jörg Rieskamp, Ashley Luckman
University of Basel, Switzerland

Sequential sampling models have a long tradition in cognitive psychology including decision making under risk. They have the advantage of predicting not only people's choice behaviour but also provide a description of the dynamics of the decision process and the resulting response times. Despite this, research into risky choices still focuses overwhelmingly on preferences and ignores response time. In the current work we build several sequential sampling models of risky decision making that can predict both preferential choices and response times. The models differ by various cognitive components such as utility functions, probability weighting functions, and distraction processes. We test the different models with two experimental studies in which people make risky decisions with different context. The results show that when only focusing on choice data, the models can hardly be discriminated. In contrast the conjoint use of choice and response times lead to a clearer picture of which model is most suitable to describe the decision making process. The results show that focusing not only on choices but taking response times into account can lead to major advances in theory development and testing.

Shall we risk it? Modeling risky choice in groups and individuals

Christin Schulze, Juliane E. Kämmer, Mehdi Moussaid, Wolfgang Gaissmaier, Thorsten Pachur
Max Planck Institute for Human Development

How do people make decisions under risk when deciding collectively as a group or individually by themselves? We examined the decisions of collaborating four-person groups and their individual members in a two-part risky choice task. First, participants (N = 120) individually and independently made a series of risky choices between 92 lottery pairs that included gain-, loss-, and mixed-gambles. Second, participants were randomly assigned to groups of four people and freely interacted to make choices in the same set of risky gambles. We found that interacting groups selected the gamble with the higher expected value more often than did the second-best and average individuals, and as often as did their best

individual members. Risk-seeking behavior, as measured by participants' choices of the option with the higher coefficient of variation, by contrast, did not differ between groups and individuals. Drawing on cumulative prospect theory to model the decisions of interacting groups as well as their individual members, we found evidence that groups made more deterministic choices reflecting less error-prone behavior but did not differ from individuals in loss aversion, sensitivity to outcome differences, or sensitivity to probability differences in the gain and loss domains. These results suggest that groups' information processing differs from that of independent individuals but that the preferences revealed in group decision making are comparable to those elicited in individual choice. This work thus extends cumulative prospect theory to collective decision making arguably the most widely used theoretical framework for modeling individual choice in decision making under risk.

Modeling switch costs in a rule-plus-exception categorization task

David K Sewell, Rebecca Jaymes, Thomas J Palmeri
The University of Queensland

Simple rules are a highly efficient way of representing category information (e.g., birds can fly). However, such rules are often incomplete, necessitating memorization of exceptions to the rule (e.g., bats can fly, but they are not birds). Hybrid categorization theories have been developed to reflect these assumptions, but the costs of coordinating access to multiple representations remains underexplored. We report the results of a categorization experiment where participants were first trained on a simple rule, after which exceptions were gradually introduced. We model the costs of switching between rule- and exception stimuli using the diffusion model. We find that switching is best characterized in terms of time-based access costs rather than interference from competing representations. Our analysis also suggests the presence of a response binding process that operates across successive trials, producing in a pattern of fast errors that is not present in a similar task involving trial-by-trial switches between multiple rules.

Beyond Pattern Completion with Short-term Plasticity

Kevin D. Shabahang, Hyungwook Yim, Simon Dennis
PhD Candidate at University of Melbourne

In a linear associative net, all input eventually settles to a single pattern. Anderson, Silverstein, Ritz, and Jones (1977) introduced saturation to force the system to reach other settling points in the Brain-State-in-a-Box (BSB). Unfortunately, the BSB is limited in its ability to generalize because its

responses are restricted to previously stored patterns. We present a set of simulations showing how a Dynamic Eigen Net, a linear associative net augmented with short-term plasticity (STP), overcomes the single-response limitation in addition to accommodating novel patterns. It accommodates novel patterns by aligning them with encoded structure as opposed to simply retrieving a previously stored pattern. We then present results from a two-slot Dynamic Eigen Net, trained on a text corpus, that shows sensitivity to syntactic violations introduced in novel bi-grams. We propose Dynamic Eigen Nets as a new class of associative nets with greater promise for generalization than the classic alternatives.

Propositions versus Associations in Memory for Sentences

Kevin D. Shabahang, Hyungwook Yim, Simon Dennis
PhD Candidate at the University of Melbourne

Propositional accounts of organization in memory have dominated theory in compositional semantics but their underlying empirical basis remains equivocal. We present data from a set of sentence memory experiments, designed to distinguish between a propositional encoding account and a purely associative account based on the Syntagmatic-Paradigmatic model capable of encoding proposition-like representations based on two simple types of association (Dennis, 2005; SP). We manipulated propositional-overlap by including a distractor sentence in the study set that shared a verb with a target sentence. We manipulated paradigmatic-overlap by including two distractor sentences, one of which contained a name from a target sentence. We embedded the study sentences into narratives in three experiments and examined the probability of correct recall, given a question probing for four target names that were the subject of four target sentences. Whereas the first experiment provided evidence for the SP account, the second provided evidence for a propositional account. In a final experiment, we found mixed results. Evidence for the propositional account was more consistent across the three experiments, however, a propositional account cannot readily explain some of the results.

“Trait” Risk Attitudes as a Result of Response Bias?

Yiyun Shou, Guangyu Zhu, Michael Smithson
The Australian National University

Risk taking propensity has been perceived as a person attribute and is commonly assessed by self-report scales in fields such as clinical, health, and social psychology. The Domain Specific Risk Taking (DOSPERT) scale, for example, is one of the most popular scales used by researchers to assess

individual differences in domain-specific risk attitudes. Yet, not many researchers pay attention to its psychometric properties when using this scale. Recently, a meta-analysis of studies over the past two decades found that the internal consistency of several DOSPERT subscales can be problematic, putting its internal structure validity into question. The present study is an investigation of the internal structure of self-report risk attitudes scales, especially the DOSPERT, with a focus on the impact of response bias. I examine the pattern of response bias in these scales and illustrate how response bias can influence the reliability and factor structure of the scales. I will also discuss methods that can mitigate the response bias, and implications for the theory and measurement of risk attitudes.

Effects of Item Relatedness on Output Interference

Jennifer Sloane, Corey White, Chris Donkin, Ryan Curl
UNSW

In recognition memory tasks, output interference (OI) effects manifest as a decrease in performance over the course of a test list. While this interference effect has been shown in many different experimental contexts, it is unclear how it is influenced by different properties of study and test items. This work investigates the relationship between semantic similarity and OI in memory, comparing semantically related (animal names or emotion words) and unrelated items to better understand memory and decision processes. The experiments used a single item recognition task and showed OI for each stimulus type and differential effects for related words that were emotional in content compared to animal names, such that emotion words showed an overall higher hit rate and false alarm rate compared to the unrelated words, whereas animal names and unrelated words did not. A diffusion model analysis was implemented to ameliorate our understanding of how specific decision making parameters change across conditions and over the course of the test list.

Vision for the Blind: Visual Psychophysics and Blinded Inference for Decision Models

Philip Smith, Simon Lilburn
The University of Melbourne

Evidence accumulation models like the diffusion model are increasingly used by researchers to identify the contributions of sensory and decisional factors to the speed and accuracy of decision making. Drift rates, decision criteria, and nondecision times estimated from such models provide meaningful estimates of the quality of evidence extracted from the stimulus,

the bias and caution in the decision process, and the durations of predecisional and postdecisional processes. Recently, Dutilh et al. (Psychonomic Bulletin & Review, 2019) carried out a large-scale, blinded validation study of decision models using the random dot kinematogram (RDK) task. They found that the parameters of the diffusion model were generally well recovered, but there was a pervasive failure of selective influence, such that manipulations of evidence quality, decision bias, and caution also affected estimated nondecision times. This failure casts doubt on the psychometric validity of such estimates. We argue that the RDK task has unusual perceptual characteristics that may be better described by a model in which drift and diffusion rates increase over time rather than turn on abruptly. We reanalyze the Dutilh et al. data using models with abrupt and continuous onset drift and diffusion rates and find that the continuous onset model provides a better overall fit and more meaningful parameter estimates, which accord with the known psychophysical properties of the RDK task. We argue that further selective influence studies that fail to take into account the visual properties of the evidence entering the decision process are likely to be unproductive.

Understanding Middle-Response Bias

Michael Smithson, Yiyun Shou, Guangyu Zhu
The Australian National University

Researchers using popular rating scales (e.g., with Likert formats) routinely ignore the occurrence of a big “neutral” spike in the middle of the distribution. This phenomenon sometimes is called “middle-response bias” (MRB), and it can distort both measurements and models. To date, the literature on MRB is scanty and provides relatively limited insight into its nature or causes. Using a 20-nation data-set containing 59 Likert-scaled items with pronounced MRB, this talk addresses the following questions: - Is MRB categorical or dimensional? Multi-dimensional? - How can MRB be measured? And does MRB on one set of items predict MRB on another set? - Is MRB factorable? What are its structural and psychometric properties? - To what extent is MRB attributable to the individual responder’s characteristics, versus the type of item? - Do socio-cultural factors influence MRB?

The role of eye gaze in gaze contingent visual working memory tasks

Ed Stewart, Chris Donkin, Mike Le Pelley
UNSW, Sydney

In previous experiments we have examined the role of eye gaze on performance in visual working memory (VWM) tasks. When analysing their

eye movements, we found that participants had a tendency to stare at the centre of the display rather than look at individual items. As a result, our data didn’t show a strong link between looking at an item and being able to recall it. In two new experiments we used a gaze contingent paradigm to force participants to fixate particular items to see if we could observe an affect of fixation on task performance. The first experiment was a repeat of a previous task but now gaze contingent. In order to see the stimuli to be remembered, participants needed to fixate on an item’s location. The second experiment paired items together so that when an item’s location was fixated, that item would be revealed as well as an additional stimulus in the periphery. The preliminary results of these experiments show a clearer link between fixation and task performance. Importantly, the second experiment’s results show that seeing an item in the periphery improves recall. This can explain why passive eye tracking may not be useful in standard VWM tasks.

Modelling how humans use decision aids in simulated air traffic control

Luke Strickland, Bowden, Boag, Heathcote
Curtin University

Air traffic controllers must often decide whether pairs of aircraft will violate safe standards of separation in the future, a task known as conflict detection. Recent research has applied evidence accumulation models (e.g., the linear ballistic accumulator; Brown & Heathcote, 2008) to simulated conflict detection tasks, to examine how the cognitive processes underlying conflict detection are affected by workplace factors such as time pressure and multiple task demands (e.g., Boag, Strickland, Loft & Heathcote, 2019). To meet increasing air traffic demands in future, controllers will increasingly require assistance from automation. Although automation can increase efficiency and overall performance, it may also decrease operator engagement, leading to potentially dire consequences in the event of an automation failure. In the current study, we applied the linear ballistic accumulator model to examine how humans adapt to automated decision aids when performing simulated conflict detection. Participants performed manual conditions, in which they made conflict detection decisions with no assistance. They also performed automated conditions, in which they were provided an (accurate but not perfect) decision aid that recommended a decision on each trial. We found that decision aids improved performance, primarily by inhibiting evidence accumulation towards the incorrect decision. Similarly, incorrect decision aids (i.e., automation failures) impaired performance because accumulation to the correct decision was inhibited. To account for these findings, we develop a framework for understanding human information integration with potentially broad

applications. Future research should investigate how cognitive processes are affected by differing levels of automation reliability, and test whether our model applies to other important task contexts.

People as intuitive scientists: Reconsidering statistical explanations of human cognition

Aba Szollosi, Ben Newell
UNSW

A persistent metaphor in the cognitive sciences casts people as intuitive statisticians. Popular explanations of human learning and decision making based on this metaphor assume that the way in which people represent the environment is specified a priori. Here we argue that the major flaw in this account is that it is not clear how (inevitable) errors in such representations of the environment can be detected and corrected. Instead we suggest that people should be considered intuitive scientists who attempt to detect and correct errors in their representations. Adopting this view would reorient focus from the documentation of presumably static representations to the study of the flexible ways in which representations can be applied, generated, and improved.

Assessing Workload Capacity in a Continuous Task

Alexander Thorpe, Sylvia Evans, Keith Nesbitt, Ami Eidels
University of Newcastle

Designers of user interfaces must consider the demands placed on the users' cognitive capacity. The detection response task (DRT) has been used to assess operators' capacity in real-world, continuous tasks such as driving. However, this task does not assess primary task performance, instead relying on data from a secondary task. In contrast, the capacity coefficient, part of the systems factorial technology (SFT) modelling framework, directly assesses the workload capacity of a user's cognitive processes. Unlike real-world tasks, which are often continuous, SFT requires discrete data derived from trial-by-trial tasks. The current study investigated a novel way of applying SFT to data from a continuous target-tracking task across two experiments. Each experiment was made of two sub-tasks, performed concurrently: a primary task, comprising either a simple signal detection task or a mouse tracking task, and a secondary task in the form of a computer-based DRT. Task load varied trial-by-trial, with either one or two targets presented. In Experiment 1, a redundant target paradigm was used, in which participants only needed to respond to one target (OR rule), while in Experiment 2, participants were required to respond to each target presented (modification of the AND rule). DRT responses did not show

significantly higher task load in two-target trials than in one-target trials, but SFT capacity analysis showed limited-capacity processing, except for the exhaustive signal detection task, which showed super capacity. The results suggest people normally operate under limited capacity, unless otherwise influenced by specific task demands.

Top-down and Bottom-up Processes in Risky Choice

Yonatan Vanunu, Jared M. Hotaling, Ben R. Newell
UNSW

When facing a gamble-choice, presented as an array of possible payoffs, people tend to sample only part of the displayed information in respect to task demands (Vanunu, Hotaling & Newell, in press). However, it is yet unclear whether this mechanism is governed solely by a top-down goal-driven process, or by a bottom-up stimulus-driven course as well. We assume that while attention might be stimulus-driven, sampling should be goal-driven. Hence, together they should determine the probability of each item to be sampled into integration and its impact on choice. In two experiments we tested this hypothesis by modifying the array gambling task to include gains and losses framing (top-down) and different font-sizes within the array (bottom-up). Findings show that people were both driven by goal and font-size in their choices. However, when the manipulation of font-size was prominent, bottom-up attention dominated goal. To disentangle these effects, we compared three versions of the Selective-Sampling and Integration Model (SSIM), where sampling carried-out according to the items' value (i.e. goal), font-size or both. Findings show that while both governed sampling, font-size was often more influential than goal when they competed on resources.

The state of statistics in Knowledge Exchange Transfer questionnaire construction

Erin Walsh, Ginny Sargent
PHXchange (Population Health Exchange), RSPH, Australian National University

Knowledge exchange transfer (KET) is the process by which knowledge is synthesized, disseminated, exchanged, and applied. As institutional and national metrics of success are shifting from quantity of output to impact, the desire to quantify KET is gaining attention in business, policy, and academia. This relevance across disciplines and sectors has led to a highly heterogenous landscape of KET measurement tools, particularly in the past decade.

This talk draws from a comprehensive review of the literature (93,000+ manuscripts screened, 30 instruments identified) and summarises the predominant statistical procedures for the development and evaluation of self- or informant- report KET questionnaires. The focus will be strengths, and areas for improvement, to inform future KET questionnaire design.

Bayes for New and Old: Modelling type and token effects in induction

Belinda Xie, Danielle Navarro, Brett Hayes
UNSW

When making inductive inferences outside the laboratory, learners encounter different forms of evidence. Evidence may take the form of novel facts (unique “types”) or repetitions of previously-learned facts (“token” repetitions). The current work examined how people use these different forms of evidence in property inference tasks. We presented participants with a sample of birds or flowers known to have a novel property then probed whether this property generalised to novel items varying in similarity to the initial sample. We found that increasing the number of novel types (e.g., new birds with the target property) tightened inference, by increasing generalization to highly similar stimuli and decreasing generalization to less similar stimuli. On the other hand, increasing the number of tokens (e.g., repeated presentations of the same bird with the target property) had little effect on generalization. In a separate experiment, we found that repeated tokens are encoded and can benefit recognition, but are subsequently given little weight when inferring property inference. We therefore modified an existing Bayesian model of induction (Navarro, Dry, & Lee, 2012) by assigning less informational value to tokens as compared to types, and show that it can capture these empirical effects.

A systematic re-examination of the list-length effect in recognition memory

Hyungwook Yim, Adam F. Osth, Simon J. Dennis
The University of Melbourne

The list-length effect (LLE) in recognition memory refers to the phenomenon where performance decreases as the length of the to-be-remembered list increases. The phenomenon has been theoretically important in the literature concerning the sources of forgetting since the existence of the LLE entails that memory interference stems from each

individual item (i.e., item-noise). Although initial studies have found evidence for LLE (e.g., Gillund & Shiffrin, 1984, Ratcliff, Clark & Shiffrin, 1990), experimental design confounds regarding retention interval, attention, and contextual reinstatement have been identified, which led to additional studies that provided evidence for a null LLE (Dennis, Lee and Kinnell, 2008). The evidence for a null LLE also suggested that recognition memory interference does not majorly stem from items-noise and that other sources may play a major role (i.e., context noise; Dennis & Humphreys, 2001).

Recently, Brandt, Zaiser, & Schnuerch (2019) pointed out that the experimental designs that support a null LLE suffer from confounds regarding the ordering of experimental conditions. As the LLE is measured by comparing recognition performance in a long and a short list, a within-subject design has been used in which the order of the two different conditions was randomized. However, when Brandt et al. analyzed just the first study test cycles, they found evidence for a LLE. The results were interpreted to mean that proactive interference accumulates across the experiment (interlist interference), so that at the second test, regardless of whether it is a short or long list, the amount of accumulated interference will be equated – leading to a null LLE.

Therefore, with new evidence for a LLE, in the current study, we re-examined the LLE more systematically -manipulating different parameters. We tested five list lengths (i.e., 8, 16, 32, 64, 80), two delay lengths (i.e., immediate, 2min), two stimuli type (i.e., words, fractals), and three study times (i.e., 1, 2, 4 sec) resulting in 60 conditions. To avoid the interlist interference confound created by using a within-subject design, we used a between-subjects design, and to accommodate the amount of individual variance we recruited 3,600 participants through mTurk (60 participants per condition). Results show evidence for a LLE with a different amount of interference showing across different stimulus type and conditions. We also decomposed the sources of interference (i.e., item, context, and background noise) using a computational model (Osth & Dennis, 2015) to compare the relative amount of interference affecting recognition memory across conditions.

Extreme Response Style: Asymmetry and Domain Specificity

Guangyu Zhu, Yiyun Shou, Michael Smithson
Australian National University

Response style (RS) refers to the situation where respondents favor certain responses regardless of question content. Extreme response styles (ERS), which refers to a person’s tendency to use the lowest or highest responses regardless of question content, has the most controversy surrounding its

nature compared to other response styles. It can lead to biased results, threatening the internal validity of studies. However, Current debates about the nature of ERS influence the choice of various methods for controlling it. Using the World Values Survey dataset, this talk aims to clarify the following controversies about ERS: 1. Can the tendencies to choose the lowest response and highest response on a scale be categorized as a single response style called “extreme response style”? Or are these tendencies separate response styles? 2. Is extreme response style (or two separate response styles) a personal trait or an interaction between individual and topic? We discuss our findings and recommendations of choosing methods to deal with ERS.