

Date estimation report

Matt Jaquiere (matt.jaquiere@psy.ox.ac.uk)

Contents

Introduction	1
Method	2
Overview	2
Interface	2
Structure	2
Advice	2
Analysed trials	2
Analysis	3
Resources	3
Results	3
Exclusions and sample	3
Task performance	3
Response times	3
Marker usage	4
Error	5
Error by advisor	6
Manipulation checks	7
Dependent variable	7
Hypothesis testing	8
Post-hoc tests	9
Discussion	10
Etc	11
Credits	11
Acknowledgements	11
R Packages	11
Funding	12
Technical details	12

June 2019

[Script run 2019-06-10 13:04:37]

Introduction

People are capable of rationally altering their use of advice in light of information about the quality of advice (Sniezek & Van Swol, 2001) or trial-by-trial feedback which allows them to derive an assessment of advice quality (Yaniv & Kleinberger, 2000). We examine whether they remain capable of this in the absence of any objective cues to advice quality.

The theoretical background states that people make assessments of advice in the absence of feedback using their own metacognitive abilities. For example, when receiving disagreeing advice, a person who is confident in their initial decision will a) adjust their final decision less, and b) downgrade their assessment of the advisor. This theory is due in large part to Niccolo's work.

Here we present participants with advice from one of two advisors, one who indicates the correct answer and the other who indicates support for the participant’s initial decision. We expect that participants provided with trial-by-trial feedback will recognise the value of the accurate advisor, while those who do not receive feedback will be primarily influenced by the agreeing advisor because that advisor is providing the most plausible answers from the participant’s perspective.

Method

Overview

We present participants with a date estimation task within a judge-advisor system where the participant plays the role of the judge, and the advisors are played by virtual agents. Participants make an initial decision by selecting a span of time on a timeline which they believe covers the event in question (e.g. assassination of Archduke Franz Ferdinand). Participants then see a span of time indicated by an advisor, before submitting a final decision.

Interface

Decisions are indicated by selecting one of three markers, each of which covers a different number of years. The thinnest marker covers a single year, a medium marker covers 3 years, while the widest marker covers 9 years. They are worth 27, 9, and 3 points respectively. The timeline covers the years 1890-2010, and correct answers fall in the range 1900-1999.

Structure

Participants were introduced to the experimental interface, and given 10 trials with feedback and without advice to familiarise themselves with the task (we found participants in earlier prototypes altered their behaviour based on feedback about *their own* decisions). They were then briefly introduced to advice for two trials before completing two core blocks of 15 trials.

Core blocks contained 15 trials: 14 genuine trials and 1 attention check trial. Attention checks provided participants with an instruction to ‘use the smallest marker to indicate the year [e.g.] one nine three three’. On genuine trials participants received advice and either did or did not receive feedback according to their experimental condition. Advice was provided by a single advisor on each block, introduced at the beginning of the block and represented visually throughout. On one block participants received advice from an advisor who agreed with their initial decisions, and on another from an advisor who indicated the correct answer.

Advice

The advice is drawn from a small variation around either the correct answer (1914 in the example above) or the centre of the participant’s answer. Advice always covered a period of 6 years. The off-centre adjustment of the advice was calculated by subtracting the expected value of 3 integers drawn from the range [0,5] (with replacement), and adding the sum of those integers. This resulted in a triangular distribution centred around 0.

Analysed trials

Crucially, on 20% of genuine trials, the advisor offered ‘off-brand’ advice. This advice was the same for both advisors, and was calculated by reflecting the participant’s answer around the correct answer. These off-brand trials were the ones analysed in the core statistical analysis, and they are important because the on-brand trials confound advisor identity with advice-initial-decision-distance, which has been shown to alter advice-taking.

Analysis

The core analysis is a 2x2x2 mixed ANOVA exploring the influence of different advisors (within subjects) depending upon which advisor is seen first, and whether feedback is provided or not (both between subjects). The dependent variable, Weight on Advice (WOA) is a commonly-used measure of advice-taking and gives the extent to which the final decision moves away from the initial decision as a proportion of the advice. It is given by $woa = (f - i)/(a - i)$ and is well-defined for values in the range $[0,1]$.

All other analyses are secondary/exploratory.

Resources

This project's resources can all be accessed via its OSF page.

This study was pre-registered. The registration can be viewed at <https://osf.io/fgmdw>.

The experiment itself can be viewed at https://acclab.psy.ox.ac.uk/~mj221/ESM/ACv2/?PROLIFIC_PID=WriteUp, and the code for both the experiment and the analysis are available in the project GitHub repository <https://github.com/oxacclab/ExploringSocialMetacognition>.

Results

Exclusions and sample

Data were collected iteratively until 5 participants passed exclusion checks for each experimental condition. Power analysis for the key interaction in the 2x2x2 mixed ANOVA was used to determine this sample size. The exclusion process can be viewed in detail in the confirmatory analysis script, but a summary is presented here:

```
## Warning: Column `pid` joining factors with different levels, coercing to
## character vector
```

	fb_AgrFirst	fb_AccFirst	¬fb_AgrFirst	¬fb_AccFirst
attnCheckYear	1	0	0	0
attnCheckYear , safari	1	0	0	0
attnCheckYear, attnCheckMarker	0	2	0	0
excess	0	0	2	4
FALSE	5	5	5	5
guessedManipulation	1	3	0	1
responseCorrect	1	0	0	0
safari	0	0	1	0

We do not collect demographic data on our participants.

Task performance

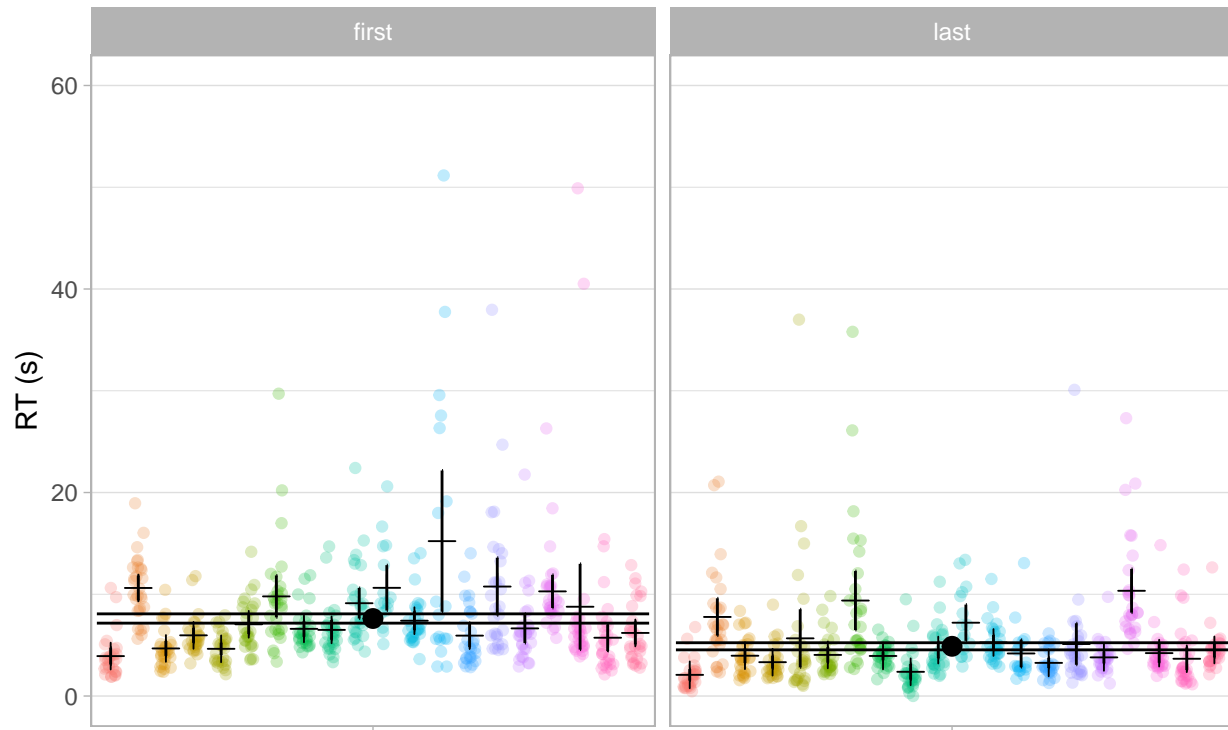
We first want to check whether the participants performed the task in a sensible way. To this end we present a characterisation of task performance for each participant.

```
## Warning: Column `pid` joining factors with different levels, coercing to
## character vector
```

Response times

Participants took an average of 7.82 +/-1.30s to make their initial decisions, and 4.93 +/-1.02s to make their final decisions.

Dots show individual trials, crosses show means and 95%CIs for individuals.



Response times look reasonable. Trials with $RT > 60s$ are excluded ($n = 25$). Final decisions are more rapid than initial decisions, as we might expect given that initial decisions involve an exhaustive search of memory and/or a reasoning process, while final decisions require only a plausibility assessment of the advice.

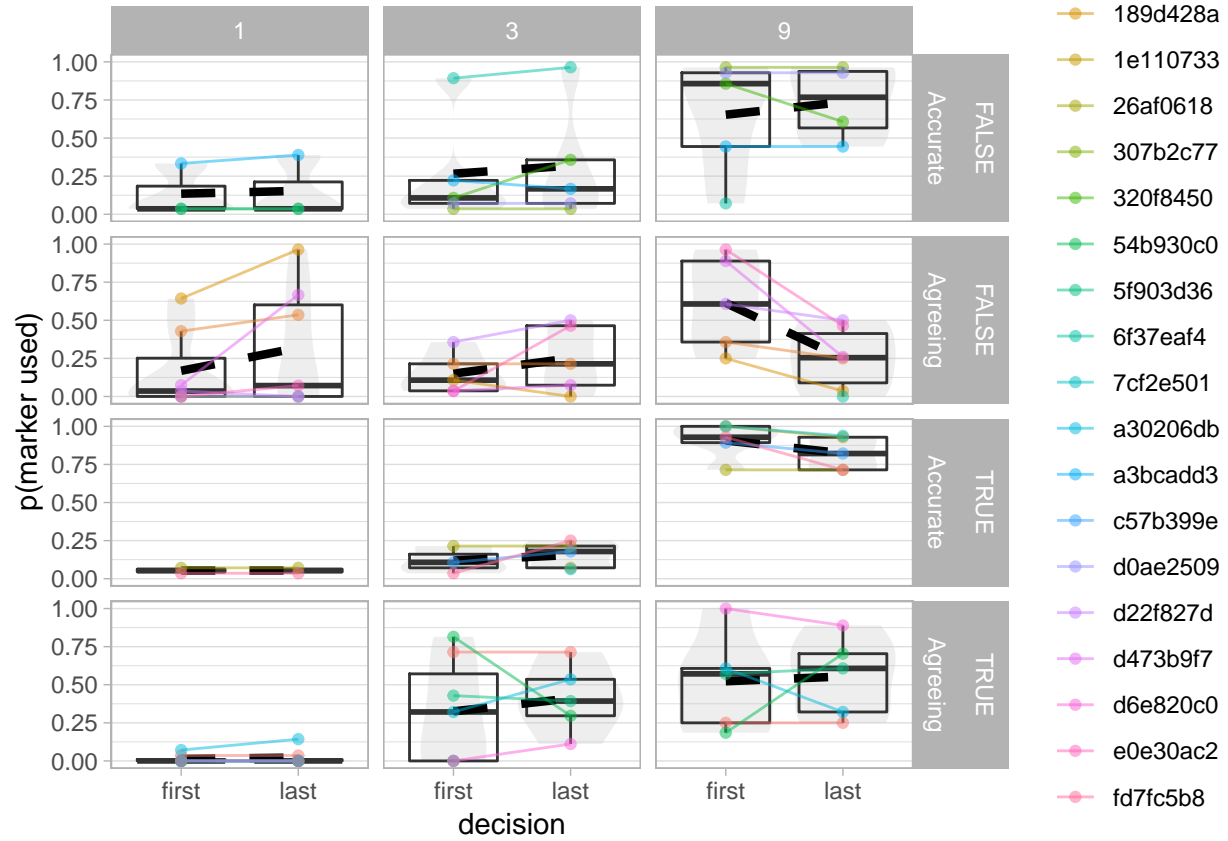
Marker usage

Three different markers were available:

marker	years	points
thin	1	27
medium	3	9
wide	9	3

How the participants used these markers is shown below.

```
## Warning in max(data$density): no non-missing arguments to max; returning -
## Inf
```

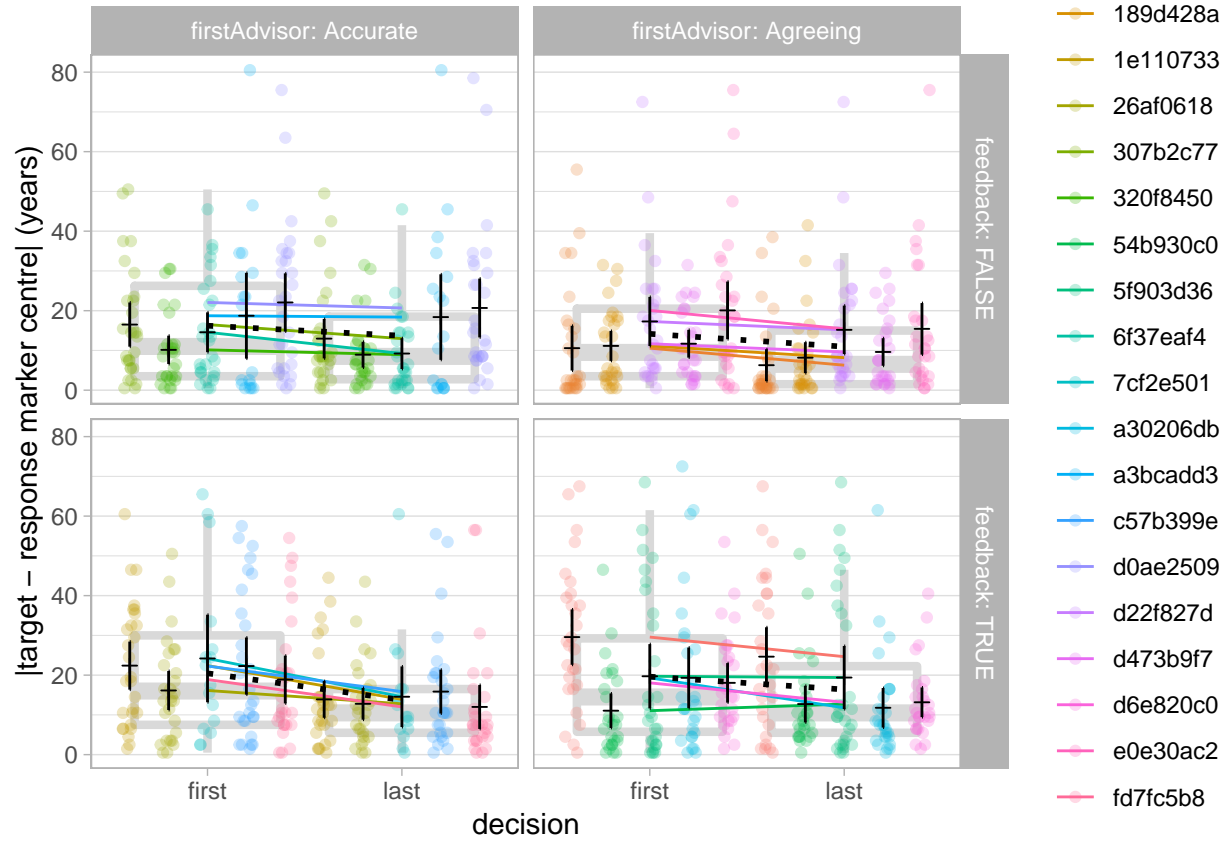


In most conditions, participants tended to use the wider markers more frequently. For initial decisions, few participants used the exact marker (1 year), although use of this marker was substantially higher in the no-feedback condition, suggesting that even after practicing 10 trials without advice participants believed themselves to be better at the task than warranted. This effect is potentially produced by the Agreeing advisor - perhaps the Agreeing advisor acts in place of feedback to encourage participants as to their accuracy (marker usage is highest where the Agreeing advisor is faced first).

Final decisions show a small tendency to increase the usage of thinner markers, suggesting that participants feel more confident on average following advice, although this is largely driven by participants in the no-feedback condition facing the Agreeing advisor first, suggesting that these participants felt the advice enabled them to make more robust decisions.

Error

The **error** is calculated as the distance from the centre of the answer marker to the correct year. It is thus possible for correct answers to have non-zero error, and it is likely that the error for correct answers scales with the marker size.

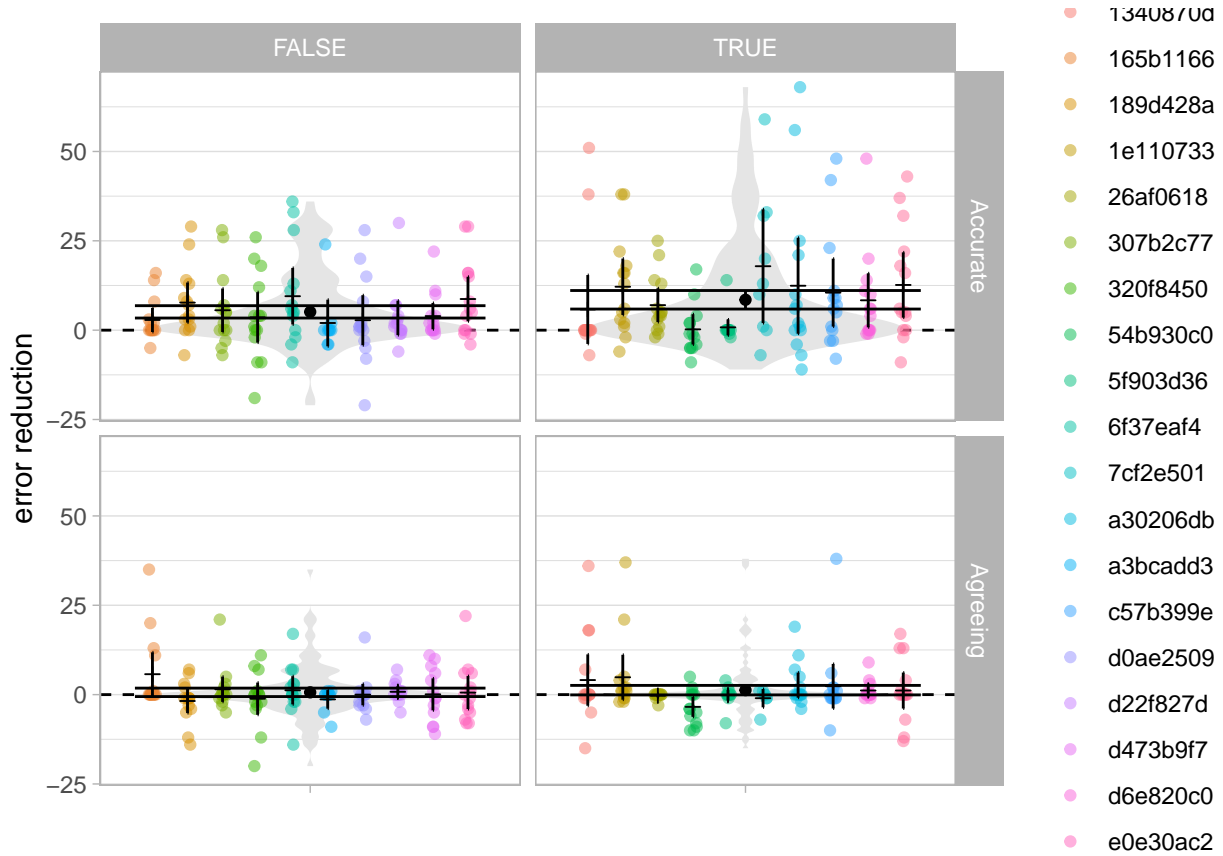


Overall mean error (17.72yrs for initial decisions, and 13.78yrs for final decisions) was reasonably consistent across groups, and all groups showed a slight improvement from initial to final decisions. Individual participants all showed a range of error, for most participants spanning about 0-60 years, or about half the width of the response scale.

There were no obvious differences in either actual scores or improvement from initial to final decisions across conditions, which is reasonable given that these conditions are supposed to affect advice utilisation rather than response error specifically. We might, however, have expected to see greater improvements in the feedback groups, who should be able to identify the Accurate advisor and use their advice to reduce the error of their final responses. This pattern did not seem to occur, at least not strongly enough to be obvious in the graph.

Error by advisor

We expect participants to benefit from the advice of Accurate advisors, but not Agreeing advisors.



Indeed, errors only appear to be reduced by advice from the accurate advisor, and they appear to be more reduced in the feedback condition.

Manipulation checks

For our experiment to be a fair test of the theory, key constraints have to be met. Firstly, the advisors have to offer characteristically different advice on average. Accurate advisors should have lower error on their advice than Agreeing advisors, and Agreeing advisors should offer advice which is closer to participants' initial responses. This was indeed the case. Accurate advisors had lower error (), while Agreeing advisors had a lower distance from participants' initial decisions ().

Secondly, there should not be large differences in the advisors' off-brand advice. This constraint was also met, with both advisors exhibiting similar error () and distance () on off-brand trials.

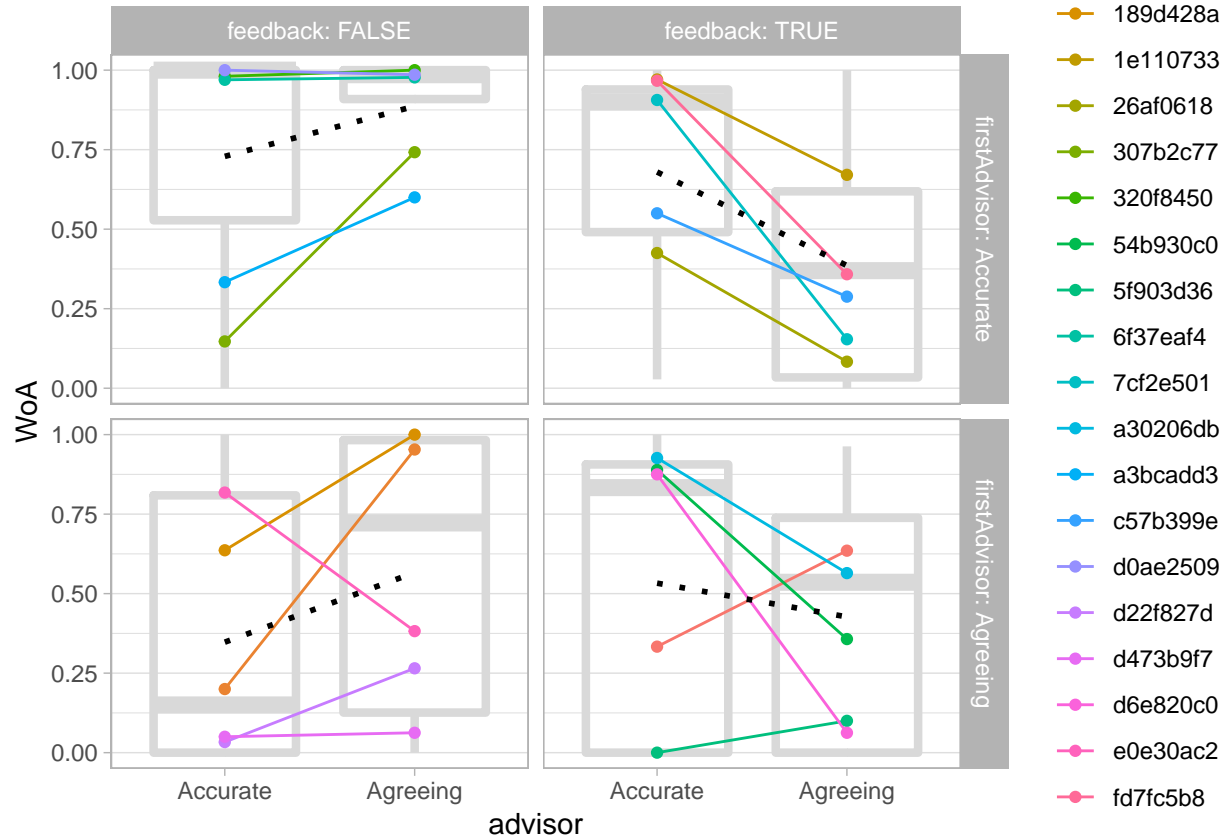
Dependent variable

The dependent variable used in the key analysis is weight-on-advice. This is well-defined for values between 0 and 1 (truncated otherwise), and is

$$\text{WoA} = (\text{final} - \text{initial}) / (\text{advice} - \text{initial})$$

, or the degree to which the final decision moves towards the advised answer. Ill-defined values (14% of responses, 10% of offbrand trials.) are truncated to 0 or 1 for analysis.

Influence is the primary outcome measure, and is thus expected to differ between advisors and feedback conditions.



The results show the expected pattern: where feedback is provided, almost all participants make greater use of the Accurate advisor's advice when the specific properties of the advice are controlled for. Likewise, where feedback is not provided, participants make greater use of the Agreeing advisor's advice. Some participants in the no-feedback, Accurate-first condition exhibit ceiling effects, presumably satisfying themselves of the utility of the Accurate advisor's advice and becoming accustomed to following advice from there.

Hypothesis testing

A 2 (Accurate vs Agreeing advisor; within subjects) x 2 (Accurate vs Agreeing advisor first, between subjects) x 2 (feedback vs no feedback, between subjects) mixed ANOVA provides a formal test of the patterns observed above:

## \$ANOVA							
##		Effect	DFn	DFd	SSn	SSd	F
## 1		(Intercept)	1	16	12.37880299	2.445165	81.0010108
## 2		feedback	1	16	0.10207848	2.445165	0.6679531
## 3		firstAdvisor	1	16	0.39305980	2.445165	2.5719968
## 5		advisor	1	16	0.07823166	1.032816	1.2119359
## 4		feedback:firstAdvisor	1	16	0.18303338	2.445165	1.1976836
## 6		feedback:advisor	1	16	0.72051670	1.032816	11.1619780
## 7		firstAdvisor:advisor	1	16	0.02556949	1.032816	0.3961130
## 8		feedback:firstAdvisor:advisor	1	16	0.02063789	1.032816	0.3197146
##		p < .05					ges
## 1	1.165026e-07	*			0.780662904		
## 2	4.257693e-01				0.028513067		
## 3	1.283252e-01				0.101538534		
## 5	2.872359e-01				0.021998588		


```

## 4 2.899874e-01      0.049995265
## 6 4.145602e-03      * 0.171612983
## 7 5.379872e-01      0.007298164
## 8 5.796226e-01      0.005898869
##
## $aov
##
## Call:
## aov(formula = formula(aov_formula), data = data)
##
## Grand Mean: 0.5563003
##
## Stratum 1: pid
##
## Terms:
##              feedback firstAdvisor feedback:firstAdvisor Residuals
## Sum of Squares 0.1020785    0.3930598            0.1830334 2.4451651
## Deg. of Freedom      1          1              1          16
##
## Residual standard error: 0.3909256
## 3 out of 6 effects not estimable
## Estimated effects may be unbalanced
##
## Stratum 2: pid:advisor
##
## Terms:
##              advisor feedback:advisor firstAdvisor:advisor
## Sum of Squares 0.0782317    0.7205167            0.0255695
## Deg. of Freedom      1          1              1
##              feedback:firstAdvisor:advisor Residuals
## Sum of Squares              0.0206379 1.0328158
## Deg. of Freedom              1          16
##
## Residual standard error: 0.2540689
## Estimated effects may be unbalanced

```

As expected, the key interaction is observed - feedback:advisor ($F(1, 16) = 11.16$, $p = .004$, $\eta^2 = .172$). The effect size is quite reasonable. This interaction effect is broken down in post-hoc tests by providing measures of greater reliance on the Accurate vs the Agreeing advisor, quantified as WoA(Accurate) - WoA(Agreeing).

Post-hoc tests

We want to know preference for accurate over agreeing:

By feedback

This is the key interaction of interest:

$t(17.98) = 3.47$, $p = .003$, $d = 1.55$, $BF = 13.61$; $M|fb = 0.36$ [0.11, 0.61], $M|\neg fb = -0.18$ [-0.42, 0.06]

As expected, this shows that Accuracy preference is stronger in the feedback group.

By first advisor:

$t(17.4) = 0.51$, $p = .617$, $d = 0.23$, $BF = 0.44$; $M|Acc = 0.14$ [-0.15, 0.43], $M|Agr = 0.04$ [-0.31, 0.38]

This result indicates that there is not enough information to make a clear pronouncement, although insofar as there is information there does not appear to be an overall difference consequent on which advisor participants saw first.

By feedback and first advisor:

$t(5.7) = 0.85$, $p = .430$, $d = 0.54$, $BF = 0.61$; $M|fb, Acc = 0.45$ [0.19, 0.72], $M|fb, Agr = 0.26$ [-0.31, 0.83]

$t(6.51) = 0.04$, $p = .965$, $d = 0.03$, $BF = 0.49$; $M|-fb, Acc = -0.17$ [-0.50, 0.15], $M|-fb, Agr = -0.19$ [-0.73, 0.36]

Neither of these tests indicate a large enough sample size to determine the presence or absence of an effect, but both are directionally indicative of the absence of effects, suggesting that the difference made by feedback is not affected by which advisor participants saw first.

Discussion

Participants playing the role of judge in a judge-advisor system received advice on a general-knowledge task from advisors who predominantly agreed with their initial decisions (Agreeing) or indicated the correct answer (Accurate). Participants who received feedback on a trial-by-trial basis showed greater reliance on the Accurate advisor’s advice, while those who did not receive this feedback showed a greater reliance on the Agreeing advisor’s advice.

Where feedback is provided, the task is somewhat similar to (Yaniv & Kleinberger, 2000), and similar results are observed: participants rationally adjust their reliance on advice as a consequence of information obtained about the utility of the advisor’s advice. The presumed mechanism for this is a sequential updating of the perception of an advisor’s usefulness following each presentation of feedback. This updating could be a simple point-estimate or a probability distribution, and could be updated in a Bayesian manner or heuristically. This perception of the advisor’s usefulness is likely only one input determining the weight on advice for any given trial: the participant’s own sense of confidence plays a role, as do properties of the advice itself (e.g. distance from the initial estimate, which is probably at least in part a proxy for plausability).

Where feedback is not provided, participants cannot use the rational updating strategy above. We propose, following (Pescetelli, 2017, On the Use of Metacognitive Signals to Navigate the Social World (DPhil)) that the participant’s own sense of certainty in their initial decision is used as a proxy for feedback. This has potential utility in the real world, where people demonstrate a negative correlation between subjective confidence and error magnitude for estimation tasks (cite), but is useless in this task, where the Agreeing advice offers no new information concerning the correct answer. In the present task, we argue that participants’ initial decisions offer their estimate of the most likely answer, meaning that advice which is closer to this estimate will be regarded as more likely to be correct, and which will therefore lead to greater increases in the perceived usefulness of the Agreeing advisor’s advice.

These data constitute evidence against an alternate mechanism whereby advice prompts participants to generate reasons in support of the advised estimate, and the number and quality of these reasons is used to update the perceived usefulness of the advisor. It is possible, perhaps likely, that on any given trial the advice may induce participants to generate reasons in support of the advice, which will presumably favour the Accurate advisor because the Agreeing advisor’s advice has already exhausted its rational basis in the initial decision phase. If this were the predominant mechanism, however, the Accurate advisor would not emerge as the preferred advisor for participants in the no-feedback condition.

Modelling of trial-by-trial responses could help provide evidence for these mechanisms specifically.

The evidence presented here in favour of a metacognitive account of the utility of information sources (in the absence of feedback) suggests consequences for people’s information-gathering in the real world. Feedback in the real world is typically absent or extremely sparse, only partially-attributable, and almost never epistemically privileged in the way that objective feedback in experimental setups is. In these environments, people may be using their own sense of certainty to filter information. Their sense of certainty is partially

driven by their own biases, and these will prompt them, on average, to make greater use of sources which reflect these biases. Such selective information processing has been shown to produce echo-chamber-like effects among rational agents (Madsen, Bailey, & Pilditch, 2018).

The selective weighting of information is one part of the confirmation bias whereby pre-existing views can become entrenched; selective seeking of information is another. Madsen et al.’s models include selective seeking, as do others (Dandekar, Goel & Lee, 2013), although the extent to which it is a feature of actual human behaviour has been disputed for decades (Sears & Freedman, 1967, Selective exposure to information: A critical review. *Public Opinion Quarterly*, 31(2), 194–213; Garrett, 2009; Colleoni, Rozza & Arvidsson, 2014; Bakshy, Messing, & Adamic, 2015, Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132; Barberá, Jost, Nagler, Tucker & Bonneau, 2015, Tweeting from left to right: Is online political communication more than an echo chamber? *Psychological Science*, 26(10), 1531–1542; Jasny, Waggle & Fisher, 2015; Nelson & Webster, 2017). Experiments in the paradigm used here, offering participants a choice between different advisors, could contribute useful insights to this debate, especially where the utility of advice is placed in conflict with its agreement (in many cases of real-life advice, agreement and accuracy of advice are correlated to some extent).

Advisors whose advice changes as a function of the participant’s initial confidence (Pescetelli, 2017; Jaquiere & Yeung, 2018) will be useful for elucidating whether confidence is indeed important in scaling the updating of perceived advisor usefulness. In order for the paradigm to be suitable for these kind of advisors, marker widths and values need to be selected such that participants tend to use them at similar rates.

Some participants exhibit extreme patterns of advice-taking, e.g. taking no advice. In this study these participants were largely excluded, but a full examination of advice-taking should include a characterisation of these participants. Such an examination should be able to differentiate participants who do not engage with the task from those who engage with the task but not the advice.

Etc

Credits

Acknowledgements

Thanks as always to Nick Yeung and the other folks at the ACC Lab.

R Packages

Package	Citations
base	R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
prettyMD	Matt Jaquiere (2018). prettyMD: Format Test Outputs with Markdown. R package version 0.1.0.
knitr	Yihui Xie (2018). knitr: A General-Purpose Package for Dynamic Report Generation in R. R package version 1.22.0.
ez	Michael A. Lawrence (2016). ez: Easy Analysis and Visualization of Factorial Experiments. R package version 0.5-1.
BANOVA	
BayesFactor	Richard D. Morey and Jeffrey N. Rouder (2018). BayesFactor: Computation of Bayes Factors for Common Distributions. R package version 0.9.12.
Matrix	Douglas Bates and Martin Maechler (2018). Matrix: Sparse and Dense Matrix Classes and Methods. R package version 1.2-10.
coda	Martyn Plummer, Nicky Best, Kate Cowles and Karen Vines (2006). CODA: Convergence Diagnosis and Output Analysis. R package version 0.5-2.
lsr	Navarro, D. J. (2015) Learning statistics with R: A tutorial for psychology students and other beginners. (Version 7.0.1) John Wiley & Sons.
curl	Jeroen Ooms (2018). curl: A Modern and Flexible Web Client for R. R package version 3.2. https://CRAN.R-project.org/package=curl .
forcats	Hadley Wickham (2018). forcats: Tools for Working with Categorical Variables (Factors). R package version 0.5.1.
stringr	Hadley Wickham (2018). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0.
dplyr	Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2018). dplyr: A Grammar of Data Manipulation. R package version 0.8.0.
purrr	Lionel Henry and Hadley Wickham (2018). purrr: Functional Programming Tools. R package version 0.2.5.
readr	Hadley Wickham, Jim Hester and Romain Francois (2017). readr: Read Rectangular Text Data. R package version 1.3.1.
tidyr	Hadley Wickham and Lionel Henry (2018). tidyr: Easily Tidy Data with ‘spread()’ and ‘gather()’ Functions. R package version 0.8.1.
tibble	Kirill Müller and Hadley Wickham (2018). tibble: Simple Data Frames. R package version 1.4.2. https://CRAN.R-project.org/package=tibble .

Package	Citations
ggplot2	H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
tidyverse	Hadley Wickham (2017). tidyverse: Easily Install and Load the ‘Tidyverse’. R package version 1.2.1. https://
testthat	Hadley Wickham. testthat: Get Started with Testing. The R Journal, vol. 3, no. 1, pp. 5–10, 2011

Funding

Matt Jaquierey is funded by a studentship from the Medical Research Council (reference 1943590) and the University of Oxford Department of Experimental Psychology (reference 17/18_MSD_661552).

Technical details

```
## Time stamp: 2019-06-10 13:05:46
## Runtime
##   user  system elapsed
##  66.23    2.82    70.00
##
##
## R version 3.5.1 (2018-07-02)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 17134)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United Kingdom.1252
## [2] LC_CTYPE=English_United Kingdom.1252
## [3] LC_MONETARY=English_United Kingdom.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United Kingdom.1252
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods    base
##
## other attached packages:
## [1] prettyMD_0.1.0      knitr_1.20          ez_4.4-0
## [4] BANOVA_1.1.2        BayesFactor_0.9.12-4.2 Matrix_1.2-14
## [7] coda_0.19-1         lsr_0.5             curl_3.2
## [10] forcats_0.3.0       stringr_1.3.1       dplyr_0.7.8
## [13] purrr_0.2.5         readr_1.1.1         tidyr_0.8.1
## [16] tibble_1.4.2        ggplot2_3.1.0       tidyverse_1.2.1
## [19] testthat_2.1.1
##
## loaded via a namespace (and not attached):
## [1] minqa_1.2.4         colorspace_1.3-2    rio_0.5.10
## [4] htmlTable_1.12      base64enc_0.1-3     rstudioapi_0.7
## [7] rstan_2.18.2        MatrixModels_0.4-1 mvtnorm_1.0-8
## [10] lubridate_1.7.4     xml2_1.2.0          splines_3.5.1
## [13] Formula_1.2-3       jsonlite_1.5        nloptr_1.0.4
## [16] broom_0.5.0         cluster_2.0.7-1     rjags_4-8
## [19] compiler_3.5.1      httr_1.3.1          backports_1.1.2
## [22] assertthat_0.2.0    lazyeval_0.2.1      cli_1.0.0
## [25] acepack_1.4.1       htmltools_0.3.6     prettyunits_1.0.2
```

## [28] tools_3.5.1	bindrcpp_0.2.2	gtable_0.2.0
## [31] glue_1.3.0	reshape2_1.4.3	Rcpp_1.0.0
## [34] carData_3.0-2	cellranger_1.1.0	nlme_3.1-137
## [37] openxlsx_4.1.0	lme4_1.1-17	rvest_0.3.2
## [40] runjags_2.0.4-2	gtools_3.8.1	MASS_7.3-50
## [43] scales_1.0.0	hms_0.4.2	parallel_3.5.1
## [46] inline_0.3.15	RColorBrewer_1.1-2	yaml_2.2.0
## [49] pbapply_1.3-4	gridExtra_2.3	loo_2.1.0
## [52] StanHeaders_2.18.0	rpart_4.1-13	latticeExtra_0.6-28
## [55] stringi_1.1.7	highr_0.7	checkmate_1.8.5
## [58] pkgbuild_1.0.2	zip_1.0.0	rlang_0.3.1
## [61] pkgconfig_2.0.1	matrixStats_0.54.0	evaluate_0.12
## [64] lattice_0.20-35	bindr_0.1.1	labeling_0.3
## [67] htmlwidgets_1.2	processx_3.1.0	tidyselect_0.2.4
## [70] plyr_1.8.4	magrittr_1.5	R6_2.2.2
## [73] Hmisc_4.1-1	pillar_1.3.0	haven_1.1.2
## [76] foreign_0.8-70	withr_2.1.2	mgcv_1.8-24
## [79] survival_2.42-3	abind_1.4-5	nnet_7.3-12
## [82] modelr_0.1.2	crayon_1.3.4	car_3.0-2
## [85] rmarkdown_1.11.2	grid_3.5.1	readxl_1.1.0
## [88] data.table_1.12.0	callr_2.0.4	digest_0.6.15
## [91] stats4_3.5.1	munsell_0.5.0	