Are individual differences quantitative or qualitative? An integrated behavioral and fMRI MIMIC approach

Zadelaar, J.N.; Weeda, W.D.; Waldorp, L.J.; van Duijvenvoorde, A.C.K.; Blankenstein, N.E.; Huizenga, H.M.

DOI
10.1016/j.neuroimage.2019.116058

Publication date
2019

Document Version
Final published version

Published in
NeuroImage

NeuroImage 202 (2019) 116058

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Are individual differences quantitative or qualitative? An integrated behavioral and fMRI MIMIC approach

Jacqueline N. Zadelaar [a,*], Wouter D. Weeda [b,c], Lourens J. Waldorp [d], Anna C.K. Van Duijvenvoorde [e,f], Neeltje E. Blankenstein [e,f,g], Hilde M. Huizenga [a,h,i]

[a] Department of Developmental Psychology, University of Amsterdam, Nieuwe Achtergracht 129B, Mailbox 15916, 1001 NK, Amsterdam, the Netherlands
[b] Department of Psychology, Methodology and Statistics Unit, Wassenaarseweg 52, 2333 AK, Leiden University, Leiden, the Netherlands
[c] Leiden Institute for Brain and Cognition, Wassenaarseweg 52, 2333 AK, Leiden, the Netherlands
[d] Department of Psychological Methods, University of Amsterdam, Nieuwe Achtergracht 129B, Mailbox 15906, 1001 NK, Amsterdam, the Netherlands
[e] Department of Developmental and Education Psychology, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, the Netherlands
[f] Brain & Development Research Centre, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, the Netherlands
[g] Department of Child and Adolescent Psychiatry, Amsterdam UMC, Vrije Universiteit Amsterdam, Meibergdreef 5, 1105 AZ, Amsterdam, the Netherlands
[h] Amsterdam Brain and Cognition Centre, University of Amsterdam, Nieuwe Achtergracht 129B, Mailbox 15900, 1001 NK, Amsterdam, the Netherlands
[i] Research Priority Area Yield, Research Institute of Child Development and Education, Mailbox 15780, 1001 NG, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Individual differences
MIMIC model
Cognitive neuroscience
Latent variable analysis
vmPFC
Decision strategies

ABSTRACT

In cognitive neuroscience there is a growing interest in individual differences. We propose the Multiple Indicators Multiple Causes (MIMIC) model of combined behavioral and fMRI data to determine whether such differences are quantitative or qualitative in nature. A simulation study revealed the MIMIC model to have adequate power for this goal, and parameter recovery to be satisfactory. The MIMIC model was illustrated with a re-analysis of Van Duijvenvoorde et al. (2016) and Blankenstein et al. (2018) decision making data. This showed individual differences in Van Duijvenvoorde et al. (2016) to originate in qualitative differences in decision strategies. Parameters indicated some individuals to use an expected value decision strategy, while others used a loss minimizing strategy, distinguished by individual differences in vmPFC activity. Individual differences in Blankenstein et al. (2018) were explained by quantitative differences in risk aversion. Parameters showed that more risk averse individuals preferred safe over risky choices, as predicted by heightened vmPFC activity. We advocate using the MIMIC model to empirically determine, rather than assume, the nature of individual differences in combined behavioral and fMRI datasets.

1. Introduction

In cognitive neuroscience, there is a growing interest in individual differences (Barch et al., 2013; Dubois and Adolphs, 2016; Foulkes and Blakemore, 2018; Pfeifer and Allen, 2012; Sherman et al., 2018). This paper focusses on the analysis of individual differences in combined behavioral and functional magnetic resonance imaging (fMRI) data, using an integrated latent variable approach, the so-called multiple indicators multiple causes (MIMIC)[1] model.

The MIMIC model falls into the domain of structural equation modeling (SEM), a statistical technique popular in the social sciences, including confirmatory factor analysis, path modeling, mediation analysis, and latent growth modeling (Kline, 2015). Particularly useful is that it allows for the modeling of directional relationships between quantitative and qualitative, as well as latent and manifest variables (Bollen, 1992; Cooper et al., 2019; Karimi and Meyer, 2014; Tomarken and Waller, 2005; Ullman and Bentler, 2003). SEM has found multiple applications in cognitive neuroscience (Astolfi et al., 2004), among which the study of individual differences in either neuroimaging, or combined behavioral and neuroimaging data (Cooper et al., 2019; Kievit, 2018; Yarkoni and Braver, 2010; Simpson-Kent et al., 2019).

In the MIMIC model a set of manifest (i.e. observable) variables called 'cause indicators' predicts one or more latent (i.e. indirectly observable) variables, which in turn predict(s) another set of manifest variables called

* Corresponding author. Nieuwe Achtergracht 129B, 1018 WS, Amsterdam, the Netherlands.
E-mail addresses: j.n.zadelaar@uva.nl (J.N. Zadelaar), w.d.weeda@fsw.leidenuniv.nl (W.D. Weeda), waldorp@uva.nl (L.J. Waldorp), a.c.k.van.duijvenvoorde@fsw.leidenuniv.nl (A.C.K. Van Duijvenvoorde), n.blankenstein@amsterdamumc.nl (N.E. Blankenstein), h.m.huizenga@uva.nl (H.M. Huizenga).

[1] **Abbreviations**: MIMIC model, Multiple Indicators Multiple Causes model.

https://doi.org/10.1016/j.neuroimage.2019.116058
Received 24 June 2019; Received in revised form 23 July 2019; Accepted 25 July 2019
Available online 26 July 2019

'effect indicators' (Bollen, 1984, 2002; Jöreskog and Goldberger, 1975). Each latent variable represents an underlying construct of interest (i.e. anxiety or temper), variability wherein can be taken to represent individual differences, predicting and predicted by observations (Kievit et al., 2011, 2012), see Fig. 1. If individual differences in said construct are quantitative in nature (i.e. dimensional, spread across a continuum), the latent variable will likewise be quantitative, as represented in what is henceforth referred to as the 'quantitative model', see Fig. 1A. If individual differences are qualitative (i.e. categorical, consisting of multiple homogenous classes), then so is the latent variable, see the 'qualitative model' in Fig. 1B."

Previous applications of the MIMIC model to combined behavioral and neuroimaging data studied quantitative individual differences in intelligence (Kievit et al., 2011, 2012; Ritchie et al., 2015). Herein structural neuroimaging measures (e.g. white matter volume) predicted individual differences in the (latent) general intelligence factor, so-called '*g*', which in turn predicted performance on cognitive tests. The MIMIC model outperformed models with different directional relationships in explaining the data. In the current study, we extend the MIMIC model to study whether individual differences are quantitative of qualitative in nature. We illustrate this using a decision making paradigm as the nature of individual differences therein is still under debate.

Some argue individual decision making differences to be quantitative, represented by a latent variable with numeric values that lie on continuous scale. A well-known example thereof comes from Prospect Theory, which states that decisions are made by constructing a utility value for each available choice option, and choosing the option with the highest value. Individual decision making differences are then explained as quantitative differences in utility value (Kahneman and Tversky, 1979). The viewpoint of individual decision making differences stemming from a quantitative latent variable finds support in a multitude of behavioral and neuroimaging studies (Barberis, 2013; Fox and Poldrack, 2009; Glöckner and Pachur, 2012; Trepel et al., 2005; Zeisberger et al., 2012).

Another possibility is that individual differences in decision making are qualitative, represented by a latent variable containing multiple classes or categories. An example of this lies in heuristics, wherein decisions are based on incomplete or selective pieces of information. Take for instance the 'take-the-best' heuristic, which entails deciding based solely on the most valid attribute that discriminates sufficiently between the different choice options (Dietrich, 2010; Gigerenzer and Gaissmaier, 2011; Graefe and Armstrong, 2012; Hardman and Hardman, 2009; Newell et al., 2007). Individual differences in decision making would come from qualitatively different heuristics being applied. The perspective of qualitative individual decision making differences also finds support in the field of both behavioral and neuroimaging studies (Artinger et al., 2015; Bexkens et al., 2016; Gigerenzer and Gaissmaier, 2011; Jansen et al., 2012; Mata and Nunes, 2010; Mata et al., 2010; Pachur and Galesic, 2013; Van Duijvenvoorde et al., 2016; Volz et al., 2006, 2010). For example, Venkatraman et al. (2009) found that individuals differed qualitatively in their decision strategies (i.e. how many, which, and how available choice attributes are utilized to reach a decision) as predicted by striatal sensitivity.

Light may be shed on the nature of individual decision making differences by comparing two MIMIC models applied to combined behavioral and neuroimaging decision making data – one wherein the latent variable representing individual differences is quantitative and one wherein it is qualitative. The latent variable of the model best explaining the data describes the nature of individual differences.

This approach may have several advantages. First, it negates the need to rely on assumptions concerning the nature of individual differences, instead providing a way to empirically verify this per dataset. Second, this model-based approach requires explicit selection of variables and definition of their effects, thus allowing for both confirmative hypothesis testing and directional relationship inferences as opposed to more exploratory alternatives (Cohen et al., 2017; Forstmann et al., 2011; Kievit et al., 2011). Third, the MIMIC model enables simultaneous analysis of combined behavioral and neuroimaging data, which has been shown to circumvent limitations associated with using either data type individually (Forstmann et al., 2011; Turner et al., 2016a, 2016b).

The current paper extends the MIMIC model approach in the neuroimaging field by Kievit et al. in several ways. First, and most importantly, whereas Kievit et al. assumed a quantitative latent variable, the main purpose of the current paper is provide a method to determine whether a latent variable is quantitative vs. qualitative in nature. Second, whereas Kievit et al. used structural MRI indices as cause indicators, we use functional MRI (fMRI) indices. Finally, whereas Kievit et al. incorporated quantitative behavioral indices as effect indicators, we incorporate qualitative (binary) choices, which are common in decision making research.

The currently proposed method relates to joint modeling, a framework that likewise allows for the simultaneous modeling of behavioral and neural data (Turner et al., 2013, 2015), even directionally as in our SEM model (Palestro et al., 2018). Closest resemblance to the approach is presented in Turner et al. (2017), wherein the neural data also loads onto latent variables representing theory-based mechanisms that drive behavior. While the focus in those papers was to obtain a reasonable joint distribution for both behavioral and neural data, the focus here is on identifying the nature of latent variables (i.e. individual differences), potentially utilizing qualitative latent variables to obtain a joint distribution. To the best of our knowledge the goal of obtaining a joint distribution for behavioral and neural data where the nature of the latent variable is to be identified has not been considered before.

The second section of this paper introduces the MIMIC model in the context of combined behavioral and fMRI decision making data. The

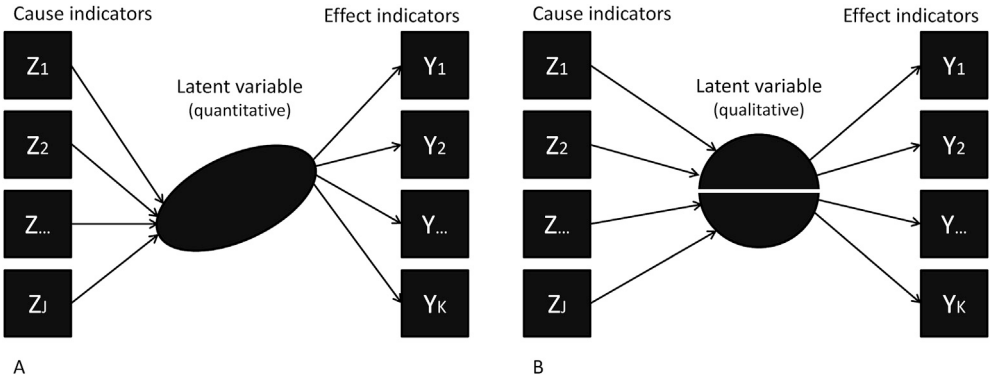

Fig. 1. The MIMIC model with *J* cause indicators (Z_1, Z_2, ..., Z_J) predicting a latent variable, which in turn predicts *K* effect indicators (Y_1, Y_2, ..., Y_K). In this example the cause indicators are neuroimaging data and the effect indicators are behavioral data. The latent variable represents the underlying construct indirectly observed via the aforementioned indicators, variability wherein represents individual differences. Fig. 1A) The quantitative latent model. Fig. 1B) The qualitative latent model.

third section consists of a simulation study assessing the power of the MIMIC model to identify the nature of individual differences, and additionally examine the models' parameter recovery. The fourth and fifth section illustrate MIMIC model application to combined behavioral and fMRI decision making data (Van Duijvenvoorde et al., 2016, and Blankenstein et al., 2018, respectively).

2. The MIMIC model

The behavioral data (Y) of each individual consist of binary responses on K items, each presenting a decision problem with two choice options (Baker et al., 2002; Fülöp, 2005; Hardman and Hardman, 2009; Payne, 1976), providing K manifest, qualitative variables. Corresponding fMRI data (Z) consists of the mean beta values of J pre-specified regions of interest (ROIs), resulting in J manifest, quantitative fMRI variables.

The MIMIC model (see Fig. 2) connects these variables so that J ($J \in \mathbb{N}_1$) quantitative fMRI indices (Z_1, Z_2, ..., Z_J) predict the latent variable (X), which in its turn predicts K ($K \in \mathbb{N}_1$) qualitative behavioral item responses (Y_1, Y_2, ..., Y_K) (Kievit et al., 2011, 2012; Ritchie et al., 2015). The latent variable is quantitative or qualitative (i.e. consisting of two or more latent classes).

The MIMIC model is essentially a combination of a standard formative model (i.e. manifest variables predicting a latent variable) and a standard reflective model (i.e. a latent variable predicting manifest variables) (Bollen, 1984; Jöreskog and Goldberger, 1975; Kievit et al., 2011, 2012). The manifest variables in the formative and reflective part of the model are referred to as 'cause indicators' and 'effect indicators', respectively.

Cause indicator coefficients (β_1, β_2, …, β_j) describe the effect of the cause indictors on the latent variable. If the latent variable is quantitative, these coefficients indicate the change in latent variable value as the corresponding cause indicator increases with one. The cause indicator intercept (β_0, not show in figure) denotes the latent variable score when $Z_1 = Z_2 = \ldots\ Z_j = 0$. If the latent variable is qualitative, the coefficients represent the change in log odds of the latent variable assuming one rather than another (reference) class, as the corresponding cause indicator increases with one. The cause indicator intercept gives this log odds for when $Z_1 = Z_2 = \ldots\ Z_j = 0$. The relationship between log odds and probability is given by: $1/(1 + e^{-\beta})$.

Effect indicator coefficients (λ_1, λ_2, …, λ_k) and intercepts (τ_1, τ_2, …, τ_k) describe the effect of the latent variable on the respective effect indicator. If the latent variable is quantitative the coefficients are log odds of the respective item taking on value 1 as opposed to 0 as the latent variable increases with one. Intercepts indicate these log odds if $X = 0$. Were the latent variable qualitative, then there are only effect indicator coefficients, one set per latent class, each describing the log odds of its item taking on value 1 rather than 0 for the respective class.

Cause indicators may be correlated amongst each other, though too high correlations are advised against (Bollen, 1984; Diamantopoulos and Siguaw, 2006). Effect indicators should be conditionally independent (i.e. uncorrelated) given the latent variable. Violation of this assumption suggests that a relevant latent variable is not yet incorporated in the model (Bollen, 1984; Jöreskog and Goldberger, 1975; Wang and Wang, 2012), which distorts coefficient values.

As effect indicators are predicted, they have error terms (ε_k). Similarly, the latent variable error term (δ) implies that this variable is predicted by the cause indicators, not a mere composite thereof (Bollen, 1984; Diamantopoulos, 2006; Jöreskog and Goldberger, 1975; Wang and Wang, 2012). Details on the mathematics of these models are found in the Supplementary Materials.

Cause (i.e. fMRI) and effect (i.e. item) indicator scores, and latent variable scores are expected to vary across individuals, reflecting individual decision making differences. Model parameters (i.e. coefficients and intercepts) are constant across individuals, representing a common process underlying decision making.

3. The simulation study

3.1. Goals

In a simulation study we determined whether the MIMIC model has adequate power to identify the true nature of a latent variable

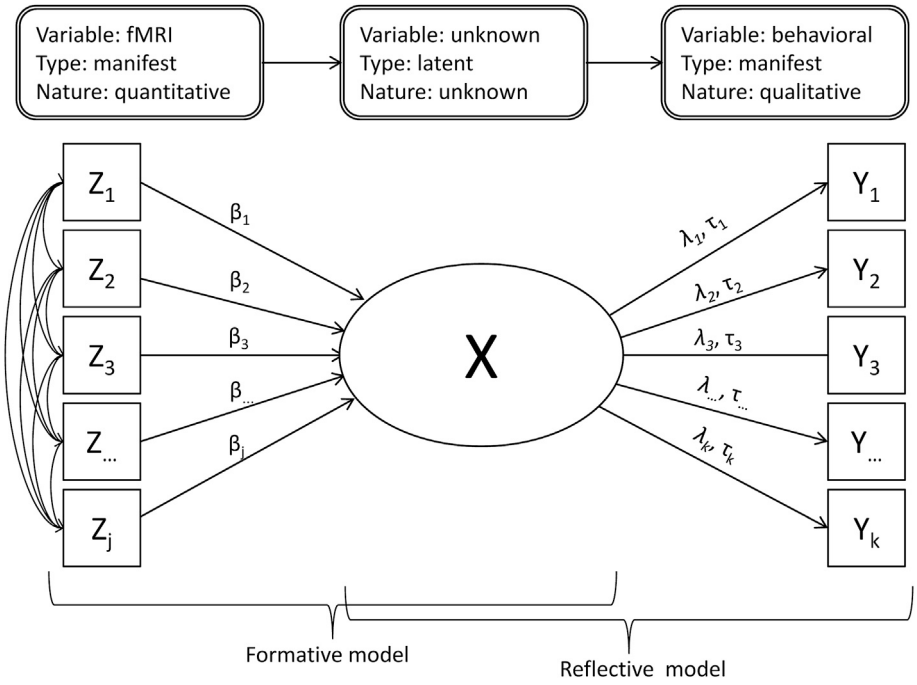

Fig. 2. The MIMIC model consisting of a formative part (left) and a reflective part (right), with J quantitative cause indicators (fMRI: Z_1, Z_2, ..., Z_J) predicting one latent variable of an unknown nature (X), which in turn predicts K qualitative effect indicators (behavior: Y_1, Y_2, ..., Y_K). Unidirectional arrows indicate the influence of one variable on the other. Double arrows indicate covariances between variables. For simplicity the depicted model is that of a single participant ($N = 1$).

representing individual differences. Two types of data were simulated, once using the MIMIC model with a quantitative latent variable (i.e. the 'quantitative model') and once using the MIMIC model with a qualitative latent variable (i.e. the 'qualitative model') with two latent classes. Of each data type 1000 datasets were simulated, which were then analyzed both correctly (i.e. with the same model as data simulation) and incorrectly (i.e. with the other MIMIC model). The proportion of simulations wherein fit measures favored the correct analysis over the incorrect analysis was taken as an estimate of power to detect the true nature of the latent variable. Additionally, parameter recovery of the correct analyses was inspected to examine whether model parameters lend themselves to meaningful interpretation.

3.2. Simulated data mimicking empirical data

Simulated data were made to resemble the empirical data of Van Duijvenvoorde et al. (2016) and Blankenstein et al. (2018), to be analyzed in the next sections, as to ensure comparability.

The Van Duijvenvoorde et al. (2016) study investigated the neural mechanisms underlying decision making in risky choice as assessed by the Gambling Machine Task (GMT); 6 two-choice items with 24 replications each (= 144 items total) were administered during a whole-brain fMRI scan. GMT items consisted of two gambling machines that could vary on the amount of certain gain, the amount of loss, and the probability of loss. Participants were instructed to pick the best machine. There were simple items (i.e. only one attribute varied between machines), conflict items (i.e. multiple attributes varied with different attributes favoring different machines, implementing conflict), and control simple and control conflict items (machines were identical and the correct response was indicated, negating the need to decide). The machine with the highest expected value ($EV = gain + P(loss) \times loss$) was objectively the optimal (i.e. the correct) choice.

On the three simple item types the optimal choice was characterized by either the highest gain amount (Simple: Gain), the lowest loss amount (Simple: Loss), or the lowest loss probability (Simple: Probability). Two out of three conflict items varied in gain and loss amount, wherein the machine with the lowest gain and lowest loss amount (Conflict: Loss) and the highest gain and highest loss amount (Conflict: Gain) was the optimal choice. On the final conflict item type all attributes varied and the machine with the highest gain and lowest loss amount, and the highest loss probability was optimal (Conflict: Loss-Gain). There were eight variations per item type, each of which was presented thrice, resulting in 24 repetitions per item type.

Van Duijvenvoorde et al. (2016) applied a mixture analysis (see also: Jansen et al., 2012; Van Duijvenvoorde et al., 2010) to only the behavioral GMT responses during the fMRI task, revealing three latent classes of decision makers. The main focus was to study and compare compensatory decision makers (i.e. choice options are compared on integrated values of their attributes, allowing disfavor on one attribute to be compensated for by favor on another) and non-compensatory decision makers (i.e. choice options are compared on individual attributes separately). In the resulting 'compensatory' class the majority of responses on all GMT item types were correct (i.e. optimal). The 'non-compensatory' class, actually consisting of two slightly distinct classes, responded similarly, except for one item type (Conflict: Gain) which was given mostly incorrect (i.e. suboptimal) responses. It was argued that on the simple items both classes decided given single attribute differences, whereas in complex items, the compensatory class chose according to EV differences and the non-compensatory class decided given loss amount differences. In a next step, fMRI data of both classes were examined.

The Blankenstein et al. (2018) study examined the neural mechanisms associated with individual differences in decision making under risk using a wheel-of-fortune task; 46 trials were presented to 198 participants during a whole-brain fMRI scan. Wheel of fortune items had participants choose between two wheels, one of which had a small certain gain ('safe' option) and one had a higher gain but also a chance of winning nothing ('gambling' option). There were 30 items where the chance of gain of the gambling (i.e. risky) option was 50%, 8 items where this was 25%, and 8 items where this was 75%. There were also 46 ambiguous items (i.e. the chance of gain in the gambling option was unknown) but these data are not considered in the current application. Decision outcomes (gain or no gain) were presented after every item. The certain amount of gain was fixed at €3, the risky amount of gain varied between €31 and €34. Individual differences in decision making under risk were assumed to be quantitatively specified by individuals' risk aversion.

In the current study, simulated data properties varied were the number of decision items, fMRI measures, and participants to gage the effects thereof on MIMIC model analysis. Variation ranges were chosen to correspond with the empirical data properties. Thus, data were simulated as if either 40, 120, or 200 participants ($N = 40, 120, 200$) responded to either 3, 6, or 9 decision item types ($K = 3, 6, 9$), repeated 24 times, the latter akin to the Van Duijvenvoorde study. The number of fMRI measures was based on the number of areas selected in the current two empirical applications by entering task-related key words into the online database Neurosynth (https://www.neurosynth.org/), details in sections 4 and 5. This produced 6 fMRI measures for both empirical applications. To introduce variety we chose the range of $J = 3, 6, 9$.

3.3. Methods

Data simulation was performed in R: Version 3.4.4 (RStudio Team, 2016). Data analysis was performed in Mplus: Version 7.31 (Muthén and Muthén, 2012). In the quantitative data simulation and analysis there was one quantitative latent variable, in the qualitative simulation and analysis there was one qualitative latent variable with two classes.

Parameter ranges were chosen to prevent complete separation, a phenomenon wherein an effect indicator (item score) is perfectly predicted (Albert and Anderson, 1984; Heinze and Schemper, 2002; McCullough and Vinod, 2003; Miller and Miller, 2011). In the MIMIC model, this may occur if effect indicator coefficients become too large, resulting in item score probabilities approximating zero or one. Therefore, we drew effect indicator coefficients from [-2,2]. Simulations showed that complete separation was not observed in this range. In the quantitative model complete separation may also occur as a result of extreme cause indicator coefficients, as these increase the range of latent variable scores, which in turn predict the item score probabilities; more extreme latent variable scores result in more extreme item score probabilities (i.e. approximating zero or one). Therefore we drew cause indicators parameters from [-1,1], simulations showed that complete separation was not observed in this range.

Scores of N participants on J fMRI indices were drawn randomly from a multivariate normal distribution with a mean of one and a covariance matrix derived from the empirical data (see Supplementary Materials). Specifically, the covariance matrix was computed from the mean correlation and variance of the Van Duijvenvoorde fMRI data, assuming equal variances and covariances for all fMRI indices. Resulting scores were fixed over simulations and standardized prior to the generation of latent variable scores. Standardized drawn scores ranged from −2.76 to 2.51.

Cause indicator coefficients ($\beta_1, \beta_2, \ldots, \beta_j$) were simulated by drawing them randomly from uniform distribution ranging from −1 to 1. The first coefficient was set to one ($\beta_1 = 1$) (Lang, 2010) and the intercept to zero ($\beta_0 = 0$). Coefficients were fixed over simulations. In the analyses all cause indicator coefficients were freely estimated, except for β_1, fixed at one, and β_0, fixed at zero.

Latent variable scores (quantitative model) and latent variable class probabilities (qualitative model) were calculated from cause indicator scores and coefficients. Noise was added to the latent variable resulting in a signal to noise ratio (SNR) of 0.5. In quantitative model analysis the latent variable residual variance was estimated freely. In the qualitative model this parameter is not estimated (Muthén and Muthén, 2012).

Latent variables scores were used to simulate K effect indicator (i.e.

item) scores, repeated $K\_repeat = 24$ times, by utilizing effect indicator coefficients (λ_1, λ_2, …, λ_k) and intercepts (τ_1, τ_2, …, τ_k), both drawn randomly from a uniform distribution ranging from −2 to 2. These draws were fixed over simulations. In the analyses, coefficients and intercepts were constrained to be the same across the 24 item repetitions as the latent variable was expected to affect these equally.

After simulation, the data were analyzed using both the quantitative and qualitative MIMIC model. Four information criterion available in Mplus are suitable for comparing the goodness of fit of non-nested models. We included the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) (Atkinson, 1981; Burnham and Anderson, 2003; Lin and Dayton, 1997; Posada and Buckley, 2004; Vrieze, 2012). Moreover, we included the corrected AIC (AICc, or Hurvich and Tsai's criterion) and adjusted BIC (aBIC), which counter potential sample size related biases better than their uncorrected or unadjusted counterparts (Burnham and Anderson, 2004; Field, 2013; Hurvich and Tsai, 1989; Nylund et al., 2007; Stoica and Selen, 2004). These four criteria were compared across the varying number of items, fMRI measures, and participants to infer which criterion is most suitable for which data properties. Lower values indicate better model fit.

3.4. Results

First, power was examined, which was defined as the proportion of analyses producing more favourable fit estimates for the correct than for the incorrect model. Recall that 1000 datasets were simulated for analysis per data type. Omitted were simulations wherein both the correct and incorrect analysis produced errors and/or warnings, indicative of estimation problems such as model non-identification. If only one analysis produced errors/warnings, the other model was considered best by default. The number of errors/warnings produced per simulation, per applied model, can be found in the Supplementary Materials.

The power of the MIMIC model is shown in Fig. 3. Note firstly that the power generally approximates values of 0.8 or higher, suggesting that the MIMIC model can accurately identify the true nature of the latent variable, whether it be quantitative or qualitative. Secondly, in the quantitative model, the AIC, BIC, aBIC, and AICc perform equally well across simulation properties. The same is observed in the qualitative model except when the number of items is at its lowest ($K = 3$), when the aBIC appears suboptimal compared to the other fit measures, most notably when the number of participants is low ($N = 40$). Thirdly, in the only cases wherein power of all fit measures fell below 0.8 (i.e. at $N = 120$, $J = 9$, $K = 6$ of the quantitative model and $N = 40$, $J = 9$, $K = 3$ of the qualitative model) this was largely attributable to the correct model producing errors/warnings, rather the fit measures themselves indicating the incorrect model to fit the data best.

The power of the model to identify the number, rather than the nature of latent variables is briefly discussed in the Supplementary Materials.

Next, parameter recovery was assessed by examining the mean difference between simulated and estimated parameters, 'simulated' parameters being the parameter values used to simulate the data and 'estimated' parameters being the parameter values returned by MIMIC model analysis. Good parameter recovery was characterized by mean differences between the simulated and estimated parameters approximating zero. Omitted were simulations wherein an analysis produced an error/warning.

Parameter recovery of both the quantitative and qualitative MIMIC model are shown in Fig. 4.

As can be seen in Fig. 4, both in the correct quantitative and the correct qualitative data analyses, the mean differences between estimated and simulated parameters generally approximated zero, indicating good parameter recovery. However, parameter recovery decreased at a high number of fMRI measures and a low number of items and participants. This was more notable in the qualitative than the quantitative model, and solely affected fMRI parameter estimates, and not the item parameter estimates.

4. Empirical study – qualitative individual differences

4.1. Goals

Both the quantitative and qualitative MIMIC model were applied to the empirical decision making data of Van Duijvenvoorde et al. (2016) to determine if individual differences in this study were quantitative or qualitative. The original study found these to be qualitative. Parameters of the superior (i.e. best fitting) model were interpreted.

4.2. Methods

Of the original 144 GMT items, 50 were omitted as all participants answered these items correctly, causing Mplus to not recognize these items as variables. Parameters of the remaining 94 items were constrained so that repetitions of the same item type had equal parameter estimates. Recall that there were six item types (i.e. effect indicators).

For ROIs (cause indicators) we refrained from using the exact brain areas reported by Van Duijvenvoorde et al. (2016) to avoid double-dipping. Instead, we selected brain areas related to either 'value' or 'conflict', the main aspects varied in the GMT. Entering these keywords into the online database Neurosynth (https://www.neurosynth.org/) yielded masks (association test, FDR, $p < 0.01$) of the left and right Nucleus Accumbens (lNAcc; rNAcc), the ventromedial Prefrontal Cortex (vmPFC), and the dorsal Anterior Cingulate Cortex (dACC), see Table 1. The association test returns activations that occur more consistently for studies mentioning the terms value or conflict than studies that do not mention these terms.

The left and right NAcc were highly correlated. To our knowledge, there is no evidence that these are differently involved in value coding, so we averaged activity across these two regions. This reduced potential effects of multi-collinearity and resulted in three rather than four ROIs.

The general linear model (GLM) of the GMT task included two regressors that coded choices for simple (1) and conflict (2) items respectively, and two regressors that coded control items (presenting no decision) for the simple (3) and conflict (4) items. All regressors were modeled at decision onset for the duration of the decision made. In a separate GLM two additional parametric regressors of absolute EV differences between options were included, separately for the simple and conflict condition. These parametric regressors coded which brain regions vary activation with the EV difference between choice options in simple (5), and conflict (6) items. In all analyses only trials in which participants chose in accordance with their dominant decision strategy were included. Deviating responses from the dominant decision strategy rarely happened and were modeled in an additional regressor of no interest.

Each ROI was associated with two neural contrasts. For the value regions (NAcc, vmPFC) these were brain responses that tracked differences in expected value to simple (1) and conflict (2) items. For the conflict region (dACC) these were brain responses tracking decision conflict in simple (simple > control contrast) and conflict items (conflict > control contrast). Fslmeants was used to extract mean beta values for the contrasts per participant, per ROI. Taken all ROIs together, this resulted in six fMRI measures (i.e. cause indicators).

Model selection was based solely on the AICc (or Hurvich and Tsai's criterion) as these data had a relatively small sample size, and only this criterion corrects for small sample size related biases (Cavanaugh, 1997; Dziak et al., 2017; Hurvich and Tsai, 1989; Stoica and Selen, 2004). Better model fit is indicated by lower AICc values.

The data were analyzed with a quantitative model and a 2-class and a 3-class qualitative model. In qualitative models with >3 classes, parameters outnumbered participants, in which case the AICc favors rather than penalizes model complexity, disallowing meaningful model comparison, see Supplementary Materials.

As in the simulations, the cause indicator intercept and first coefficient were set to zero ($\beta_0 = 0$) and one ($\beta_1 = 1$), respectively. Likewise,

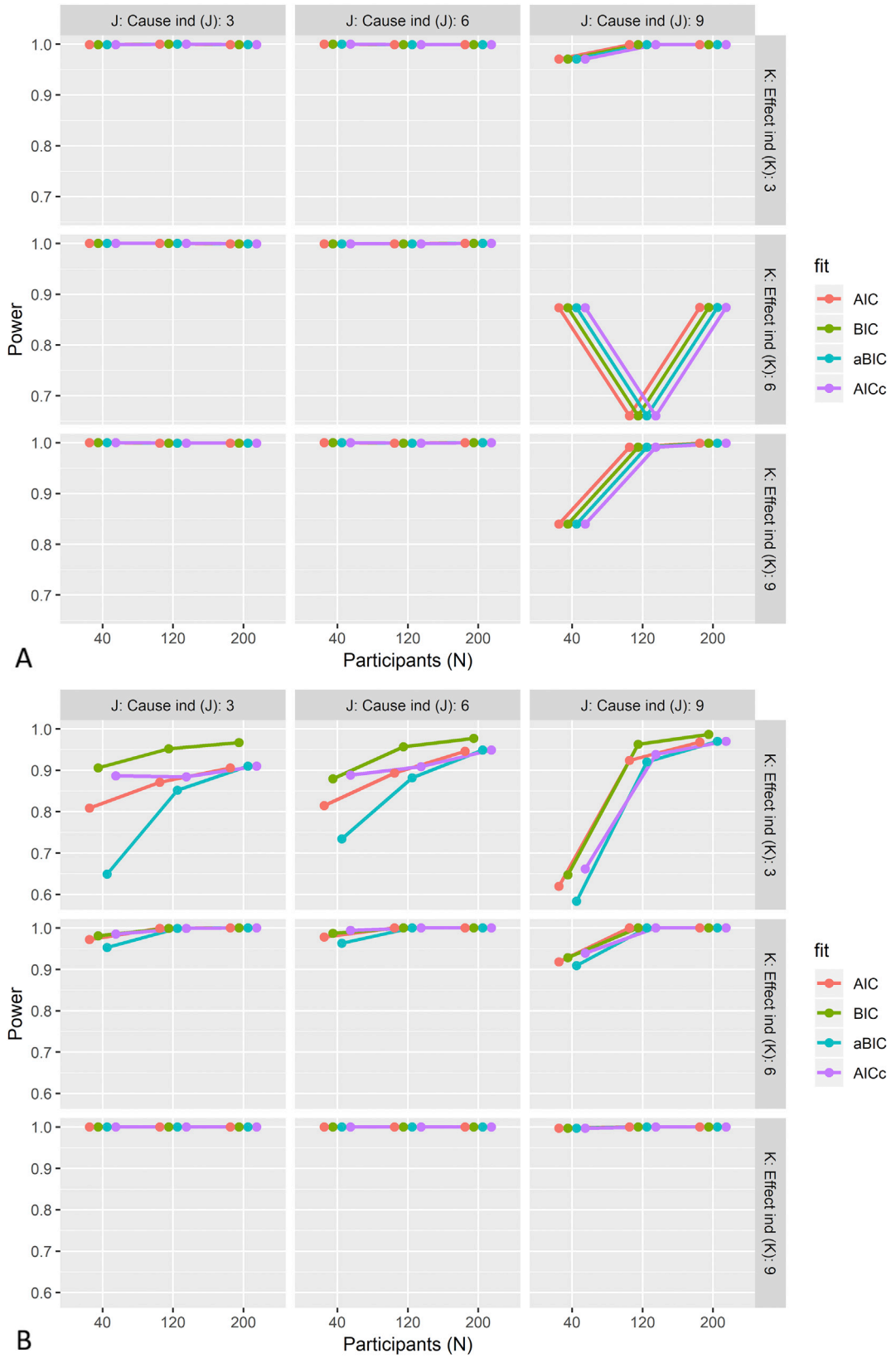

Fig. 3. Power of the MIMIC model analysis for a varied number of participants (*N*, x-axis), fMRI measures (*J*, columns), and items (*K*, rows). The power (i.e. proportion of simulations where the fit measures favoured the correct rather than the incorrect model, excluding simulations wherein both models produced errors/warnings) is shown on the y-axis. Fig. 3A. Power of the MIMIC model applied to quantitative MIMIC model simulated data. Fig. 3B. Power of the MIMIC model applied to qualitative MIMIC model simulated data.

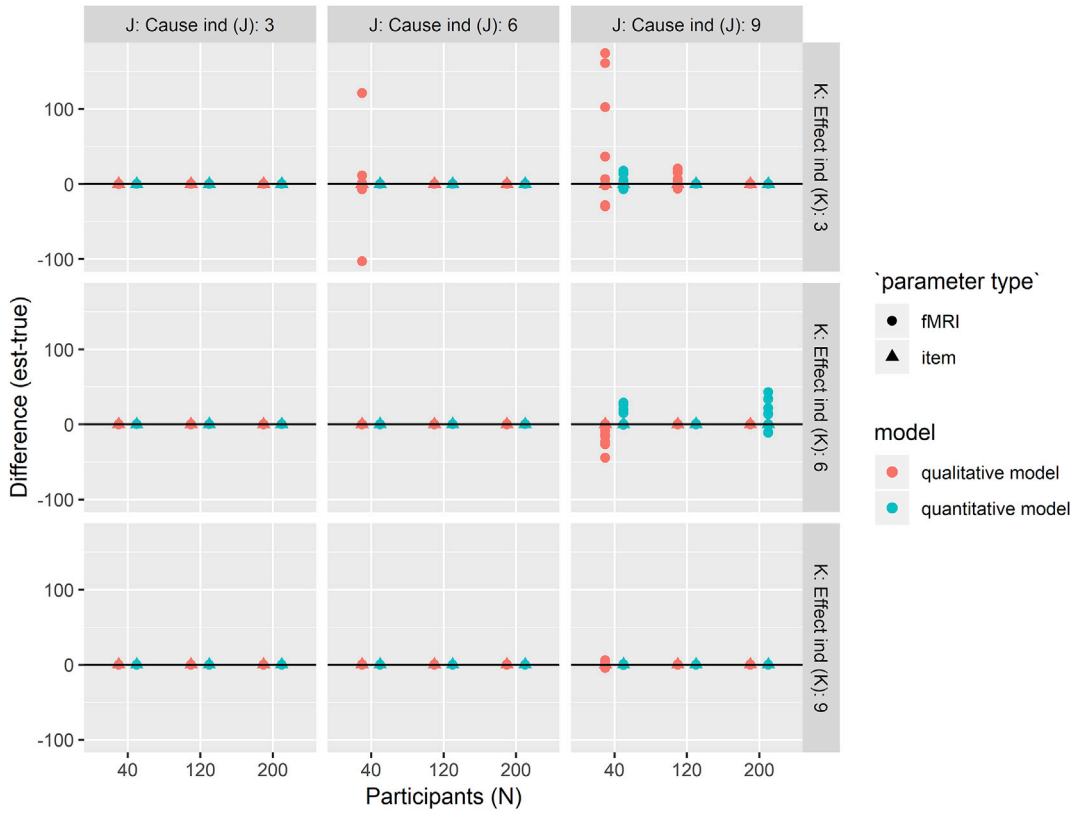

Fig. 4. Parameter recovery of the correctly applied MIMIC model for a varied number of participants (*N*, x-axis), fMRI measures (*J*, columns), and items (*K*, rows). Parameter recovery (i.e. mean difference between simulated and estimated parameter values) is shown on the y-axis. Values closer to zero indicate more accurate parameter recovery.

Table 1
Brain regions used in the empirical data analysis of the Van Duijvenvoorde et al. (2016) data, including the keyword of retrieval in https://www.neurosynth.org/, the size of the masks, and coordinates of the max. z-value in MNI space.

Area	Keyword	Size cluster	MNI coordinates mask		
	Neurosynth	(bilateral)	x	y	z
Left NAcc	Value	607	−10	10	6
Right NAcc	Value		12	12	−8
vmPFC	Value	760	−6	38	−10
dACC	Conflict	558	6	14	42

cause indicator scores were standardized. Other parameters were estimated freely. Data were prepared in R and analyzed in Mplus.

4.3. Results

The AICc of the quantitative and qualitative models were compared, see Table 2. The 2-class qualitative model produced the lowest AICc. We therefore conclude that in these data, individual differences in decision making are qualitative in nature.

Table 2
Fit measures for the quantitative and qualitative analysis of empirical data; * indicates the lowest AICc. The quantitative model has one additional (free) parameter, namely the residual variance of the latent variable.

Analysis	Latent Classes	Number of Parameters	Number of Cause Parameters	Number of Effect Parameters	AICc
Quantitative	–	18	5	12	1757.505
Qualitative	2-class	17	5	12	1749.746*
Qualitative	3-class	28	10	18	1756.923

Next, parameters of the 2-class qualitative model were interpreted. Effect (item) indicator parameters are given in Table 3. The first latent class ($N = 18$) showed a high probability of responding correctly to all items, suggesting that these participants based decisions on EV. The second class ($N = 22$) distinguished itself by a low probability of correct responses to the conflict item type wherein the choice option with the highest loss amount was correct. This thus suggests that they used a loss minimizing strategy. These results match those of Van Duijvenvoorde et al. (2016), including individual class membership assignment. For brevity the two classes are henceforth referred to as 'compensatory and 'non-compensatory', in accordance with the original study.

The cause (fMRI) indicator parameter estimates and class means are depicted in Table 4.

Cause indicator parameter estimates suggested that the two classes could be distinguished by the EV effect of both simple and conflict items in the vmPFC. Recall that it is likely that participants in the compensatory class based decisions regarding conflict items on EV differences, whereas participants in the non-compensatory class did not. In accordance, class means in conflict items show more pronounced vmPFC coding of EV differences in the compensatory than in the non-compensatory class. Note however that the class mean in the compensatory class was negative (i.e. the effect decreased with increasing EV differences). Therefore we speculate that the vmPFC did not code the difference in EV but rather the difficulty of EV based decisions, as was also proposed by several other authors (FitzGerald et al., 2009; Hare et al., 2011; Lim et al., 2011; Philiastides et al., 2010; Rushworth et al., 2011).

Recall that in simple items, it is likely that participants of both classes based decisions on the sole varying attribute (amount of gain, amount of loss, or the probability of a loss). Thus one would not expect class distinctions in vmPFC coding of EV differences. Surprisingly though, more pronounced vmPFC coding of EV differences was associated with the non-compensatory class. The corresponding class mean was negative,

Table 3
Effect parameters of the 2-class qualitative model. Per class, the estimates and corresponding probabilities of responding correctly and p-values; * indicates significant deviation from zero of the parameter estimate ($p < 0.05$). Note that Mplus estimates give the log odds of responding to an item incorrectly, rather than correctly.

Item type	Class 1 ($N = 18$) "Compensatory"			Class 2 ($N = 22$) "Non-compensatory"		
	Estimate	p-value	P(correct)	Estimate	p-value	P(correct)
Simple: Probability	−4.625	0.000*	0.990	−3.235	0.000*	0.962
Simple: Loss	−3.814	0.000*	0.978	−3.338	0.000*	0.966
Simple: Gain	−3.859	0.000*	0.979	−3.122	0.000*	0.958
Conflict: Loss	−2.789	0.000*	0.942	−3.224	0.000*	0.961
Conflict: Gain	−2.591	0.000*	0.930	2.324	0.000*	0.089
Conflict: Loss-Gain	−3.766	0.000*	0.977	−1.556	0.000*	0.826

Table 4
Cause parameters of the 2-class qualitative model. Estimates represent change in log odds of belonging to the compensatory rather than the non-compensatory class if the corresponding fMRI measure increases with one. Class means indicate the mean change in activity per class, as a function of EV (NAcc and vmPFC) or when presented with a real item compared to a control item (dACC). * indicates significant deviation from zero of the parameter estimate ($p < 0.05$).

fMRI index	Parameters		Class means	
	Estimate	p-value	Comp	Non-comp
Effect EV simple items NAcc	1.000	–	0.637	−0.760
Effect EV conflict items NAcc	0.462	0.334	−0.635	−0.362
Effect EV simple items vmPFC	2.647	0.000*	0.894	−1.235
Effect EV conflict items vmPFC	−1.619	0.000*	−0.912	0.360
Contrast simple > control dACC	0.661	0.349	23.411	20.845
Contrast conflict > control dACC.	0.517	0.367	37.419	36.507

Table 5
Brain regions used in the empirical data analysis of the Blankenstein et al. (2018) data, the size of the masks, and coordinates of the max. z-value in MNI space.

Area	Size cluster	MNI coordinates mask		
	(bilateral)	x	y	z
vmPFC	709	0	42	18
dmPFC	1218	0	32	−18
Left LPFC	1488	−41	22	43
Right LPFC		41	22	43
Left SPL	622	−33	−48	62
Right SPL		33	−48	62
Left Striatum	2847	−16	9	4
Right Striatum		16	9	4
Left Insula	1080	−39	−2	−3
Right Insula		39	−2	−3

again indicating that the vmPFC codes the difficulty of EV based choice. This may seem counterintuitive as the non-compensatory class did not base their decisions on EV, certainly not in simple items. Note however that in simple items, EV differences and single attribute differences are confounded. As such, we speculate again that these results may be taken to suggest that the vmPFC coded for the difficulty of decision in participants using a non-compensatory strategy.

5. Empirical study – quantitative individual differences

5.1. Goals

In this section the MIMIC model is used to determine whether individual differences were quantitative or qualitative in the study by Blankenstein et al. (2018), who assumed these to be quantitative. Parameters of the superior model were interpreted.

5.2. Methods

The 46 risky wheel-of-fortune items were constrained so that items with the same probability of gain (either 25%, 50%, or 75%) had the same parameter estimates. This provides three types of items (i.e. effect indicators).

Similar to the approach in the previous section, ROIs (i.e. cause indicators) were selected from Neurosynth, based on keywords reflecting properties varied during the task. In the wheel of fortune task, this was 'risk'. The resulting masks (association test, FDR, $p < 0.01$) contained the vmPFC, the dorsomedial prefrontal cortex (dmPFC), the lateral prefrontal cortex (LPFC), the superior parietal lobe (SPL), the striatum, and the insula. We aggregated over hemispheres of ROIs as again left and right sides of the same ROIs proved highly correlated ($r > 0.7$), thus likely causing multicollinearity related issues. This resulted in six ROIs, see Table 5.

Details of the fMRI data acquisition of the Blankenstein et al. study are described in the respective study. The GLM included two events: the choice phase (choosing risk or choosing safe) and the outcome phase (which was not of interest in the current study). Events of the choice phase were modeled separately for choosing the risky option ('risk') and for choosing the safe option ('safe'). The least-squares parameter estimates of the height of the best-fitting canonical HRF for each condition separately were used in pairwise contrasts. These pairwise comparisons resulted in subject-specific contrast images averaged across trials. Of interest was the contrast risk > safe. These subject-specific risk > safe contrasts were extracted from the a-priori defined Neurosynth ROI masks. The MarsBaR toolbox (Brett et al., 2002; http://marsbar.sourceforge.net) was used to extract parameter estimates for each ROI and for each participant.

Six ROIs with one contrast per ROI provide 6 fMRI measures (i.e. cause indicators). A total of 24 participants were omitted as these had never once chosen the safe response option, disallowing the intended risk > safe contrast, leaving $N = 174$ participants.

The data were analyzed with a quantitative model and a qualitative model with up to six classes. As before, the cause indicator intercept was set to zero ($\beta_0 = 0$), the first coefficient to one ($\beta_1 = 1$), and other parameters were estimated freely. Cause indicator scores were standardized. As the data has no properties suggesting one fit measure would be better than another, we compare models on the AIC, AICc, BIC, and aBIC. Lower values indicate better fit.

5.3. Results

As the quantitative model produced the lowest AIC, AICc, BIC, and aBIC (see Table 6), we conclude that individual differences in these decision making data are quantitative in nature.

Parameters of the quantitative model were interpreted, starting with the effect (item) indicator parameters, see Table 7. Intercepts show that, when the latent variable is zero, people are highly likely to choose safely in items with a low change of gain (25%), but that they are more likely to choose the risky option in items with higher chances of gain (50% or 75%). In other words, participants made riskier decisions when the gains thereof were more likely. The coefficients, being positive, indicate that individuals with a high latent variable score were more likely to choose safely. This suggests the latent variable to represent risk aversion.

Table 6
Fit measures for the quantitative and qualitative analysis of empirical data; * indicates the lowest estimate of the corresponding fit measure. The quantitative model has one additional (free) parameter, namely the residual variance of the latent variable.

Analysis	Latent Classes	Number of Parameters	Number of Cause Parameters	Number of Effect Parameters	AIC	AICc	BIC	aBIC
Quantitative	–	12	5	6	5179.206*	5181.219*	5216.693*	5178.699*
Qualitative	2-class	11	5	6	5656.048	5657.740	5690.412	5655.584
Qualitative	3-class	19	10	9	5317.921	5323.056	5377.277	5317.118
Qualitative	4-class	27	15	12	5284.053	5292.445	5359.028	5283.038
Qualitative	5-class	35	20	15	5299.980	5312.589	5390.575	5298.754
Qualitative	6-class	43	25	18	5397.975	5415.870	5504.189	5396.538

Table 7
Effect parameters of the quantitative model. Per class, the estimates and corresponding probabilities of choosing the safe rather than the risky option, and significance values; * indicates significant deviation from zero of the parameter estimate ($p < 0.05$). Note that Mplus estimates give the log odds of choosing the safe rather than risky option.

Item type	Intercepts			Coefficients		
	Estimate	p-value	P(safe)	Estimate	p-value	P(safe)
25% chance gain	2.092	0.000*	0.890	0.769	0.002*	0.683
50% chance gain	−2.130	0.000*	0.106	1.042	0.000*	0.739
75% chance gain	−3.593	0.000*	0.027	0.275	0.026*	0.568

Table 8
Cause parameters of the quantitative model. Estimates increases in the quantitative latent variable as the corresponding fMRI measure increases with one; * indicates significant deviation from zero of the parameter estimate ($p < 0.05$).

fMRI index	Parameters	
	Estimate	p-value
Risk > Safe dmPFC	1.000	–
Risk > Safe vmPFC	−0.419	0.036*
Risk > Safe LPFC	−0.025	0.940
Risk > Safe SPL	−0.172	0.428
Risk > Safe striatum	−0.973	0.079
Risk > Safe insula	0.815	0.068

Cause (fMRI) indicator parameter estimates are shown in Table 8. The vmPFC risk > safe contrast was a significant negative predictor of the latent variable, suggesting that increased activity in this area is associated with reduced risk aversion.

Cause indicator parameter estimates suggested that activation in the vmPFC while choosing risk versus safe was a negative predictor of the latent variable, which in turn predicted safe (over risky) choices. This suggests that the latent variable represents risk aversion. In other words, individuals who made relatively more safe choices (effect indicator) were characterized by heightened risk aversion (latent variable) and heightened activation in the vmPFC when choosing safe versus risky (cause indicator). This fits well with prior work showing that less risk-seeking (i.e. greater risk averse) behavior is associated with attenuated activation in the vmPFC, and neighboring valuation regions such as the OFC and striatum[2] (e.g. Blankenstein et al., 2017; Sherman, Steinberg and Chein, 2018).

6. Discussion & conclusions

6.1. Discussion

In this study we introduced the MIMIC model to determine, given both behavioral and fMRI data, whether individual differences are quantitative or qualitative in nature. A simulation study showed the MIMIC model to have adequate power to detect the true nature of individual differences. Additionally, the model proved to have generally good parameter recovery, allowing for conceptual interpretation of said individual differences. Application to empirical decision making data of Van Duijvenvoorde at al. (2016) and Blankenstein et al. (2018) illustrated this for qualitative and quantitative individual differences, respectively. We conclude the MIMIC model a promising method of determining the nature of individual differences underlying combined fMRI and behavioral data.

Several aspects deserve consideration. First, note that the fact that individual differences are qualitative in nature by no means implies that the processes within each class are qualitative. For example, the empirical data of Van Duijvenvoorde et al. (2018) was found to consist of two qualitatively different classes of participants, each using a different strategy, yet one class was concluded to use an EV maximizing strategy, in which options are compared on quantitative EV differences. This could be inferred from the MIMIC model parameter estimates.

Second, the MIMIC model relates both data types to the latent variable simultaneously rather than sequentially. In the latter approach, behavioral observations determine the nature of individual differences, and the involvement of fMRI indices is estimated in a subsequent analysis (Bolck et al., 2004; Vermunt, 2010). We chose a simultaneous approach as this has the advantage of circumventing limitations associated with analyzing either data type separately (Forstmann et al., 2011; Turner et al., 2016a, 2016b). Additionally, it allows for contribution of fMRI data in determining the nature of individual differences, giving a more complete account thereof. Finally, compared to sequential analysis the MIMIC model shows increased cause indicator (fMRI) parameter recovery when latent classes are similar (i.e. when qualitative individual differences are small) (Asparouhov and Muthén, 2014). However, because the MIMIC model includes both fMRI and behavioral data in the same analytical model it uses more parameters than sequential alternatives, thus requiring more observations (Jacobucci et al., 2018). In the current study this proved not to be problematic but it may be if, for example, the phenomenon of interest is related to many different brain regions.

Third and relatedly, in both the simulation and the empirical parts of the current study, effect (item) indicator parameters were constrained so that the latent variable representing individual differences had the same effect on all repetitions of an item type. A more flexible approach would have been to estimate these parameters freely. This was not done however, as it would, in some datasets, have resulted in the number of parameters exceeding the number of participants, causing unreliable parameter estimates and biased fit measures (Wolf et al., 2013; Supplementary Materials).

Fourth and also relatedly, some studies might require the inclusion of

[2] Note that the direction of effect of the striatum – although not significant – coincides with this rationale, and confirms the findings by Blankenstein et al. (2018) who showed that more safe choices were related to lowered striatum activation (based on a whole-brain regression). The direction of the – non significant – effect of the insula (a region commonly implicated in risk processing) also fits well with prior results showing that insula activation is heightened with more risk aversion (Sherman, Steinberg and Chein, 2018). Although these effects align well with prior work, these failed to reach significance and should therefore be confirmed in future studies.

a large number cause (fMRI) or effect (behavioral) indicators, which may lead to the necessity to estimate more parameters than the sample size allows for. However, in such circumstances an approach may be used in which it is tested which fMRI indices are required to model the data adequately. In doing so, it might be useful to adopt a regularization approach (Jacobucci et al., 2016, 2018). This may not only increase suitability to small sample data, but also to exploratory studies wherein brain regions of interest are not specified.

Fifth, in the current paper we compare two potential models: a quantitative and a qualitative model, akin to a factor model (Harman, 1976; Kline, 2014; Thurstone, 1947) and a mixture model (McLachlan and Basford, 1988; McLachlan and Peel, 2000; McLachlan et al., 2019), respectively. Individual differences may however be simultaneously quantitative and qualitative (i.e. there being qualitatively distinct classes of individuals, wherein there is quantitative spread). Thus, a potential extension would be to combine these, akin to a factor mixture model (Clark et al., 2013; Lubke and Muthén, 2005, 2007), in which within qualitatively different classes, quantitative differences may exist. This, however, is beyond the scope of the current manuscript.

Our approach here to obtain evidence for latent quantitative or latent qualitative individual differences is to fit a model to the data and to determine the value for each model in terms of an information criterion. A potential problem therein is that of local minima. However, Mplus by default uses 10 random sets of starting values in the initial stage and 2 optimizations in the final stage to minimize the risk of local minima. This likely reduces the issue of local minima, although self-evidently no guarantees can be given as is shown, for instance, in Jahn (2007).

Related, an important assumption of our approach is that the likelihood space is smooth and will not change too much with small changes in the parameters. Since this is not the focus of our paper, we did not pursue this further. However, one could check for stability across different parameter ranges near the optimum for the evidence of one model over another.

Also, parameter interpretation of the proposed method requires caution when the number of participants is small compared to the number of parameters, specifically when the number of fMRI measures (i.e. cause indicators) is large and the number of behavioral measures (i.e. effect indicators) is small. These results confirm a known limitation of SEM: the requirement of relatively large sample sizes (Cooper et al., 2019). We thus caution against using the MIMIC model for conceptual interpretation in datasets with small sample sizes, especially when a large number of fMRI measures is of interest. In the latter case, the aforementioned regularization approach could be of use.

Finally, it must also be noted that in the simulations, some MIMIC model analyses produced error and/or warning messages, indicative of estimation issues like model non-identification. Thus, interpretation of results including such messages is discouraged. Causes of errors and warnings may be remedied by changing starting values or increasing the number of iterations (Muthén and Muthén, 2012). This could not be applied in simulations of the current study though, as these included 27×1000 analyses, rendering manual inspection and adjustment unfeasible. As such, analyses that produced errors and/or warnings therein were simply omitted.

6.2. Conclusions

In conclusion, current results support the MIMIC model as a means to identify the nature of individual differences as quantitative or qualitative in combined fMRI and behavioral data. In doing so, the nature of individual differences needn't be assumed, but can be based on empirical evidence. Conceptual interpretation of individual differences is enabled by the model parameters. While this MIMIC model approach was illustrated in decision making data, it is also suitable for other types of data in which the nature of quantitative vs. qualitative individual differences need to be determined (Barch et al., 2013; Osaka et al., 2003; Ochsner et al., 2006).

Declaration of interest

None; the authors report no conflicts of interests related to this manuscript.

Ethics statement

The Van Duijvenvoorde et al. (2016) study was approved by the local ethics committee, and participants provided written informed consent. Details can be found in the original manuscript.

The Blankenstein et al. (2018) study was approved by the institutional review board of the University Medical Center. Adult participants and parents of underage participants provided written informed consent, and underage participants provided written assent. Details are found in the original manuscript.

No additional participants were collected during the current study.

Research data for this article

Codes, data, and output can be found on Mendeley Data at https://doi.org/10.17632/jd2v73gk6d.2. The empirical data analyzed in this studies belong to Van Duijvenvoorde et al. (2016) and Blankenstein et al. (2018).

Author contributions

HMH, WDW, LJW and JNZ developed the study concept. Empirical data were provided by WDW, ACKVD, and NEB. Data simulation, analysis, and reporting were done by JNZ. The manuscript was drafted by JNZ, with critical revisions by HMH. All authors approved the final version of the manuscript for submission.

Funding details

This work was supported by a VICI grant [number 453-12-005] and the Netherlands Organization for Scientific Research (NWO), the Netherlands. The funding source was not in any way involved with the research.

Acknowledgements

The authors would like to thank Helen Steingroever for checking equations and Joost Agelink van Rentergem Zandvliet for R and Mplus assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116058.

References

Albert, A., Anderson, J.A., 1984. On the existence of maximum likelihood estimates in logistic regression models. Biometrika 71 (1), 1–10. https://doi.org/10.1093/biomet/71.1.1.

Artinger, F., Petersen, M., Gigerenzer, G., Weibler, J., 2015. Heuristics as adaptive decision strategies in management. J. Organ. Behav. 36 (S1), S33–S52. https://doi.org/10.1002/job.1950.

Asparouhov, T., Muthén, B., 2014. Auxiliary variables in mixture modeling: three-step approaches using M plus. Struct. Equ. Model.: A Multidiscip. J. 21 (3), 329–341. https://doi.org/10.1080/10705511.2014.915181.

Astolfi, L., Cincotti, F., Mattia, D., Salinari, S., Babiloni, C., Basilisco, ., Babiloni, F., 2004. Estimation of the effective and functional human cortical connectivity with structural equation modeling and directed transfer function applied to high-resolution EEG. Magn. Reson. Imag. 22, 1457–1470. https://doi.org/10.1016/j.mri.2004.10.006.

Atkinson, A.C., 1981. Likelihood ratios, posterior odds and information criteria. J. Econ. 16 (1), 15–20. https://doi.org/10.1016/0304-4076(81)90072-5.

Baker, D., Bridges, D., Hunter, R., Johnson, G., Krupa, J., Murphy, J., Sorenson, K., 2002. Guidebook to Decis.-making Methods. Dep. of Energy, USA. WSRC-IM-2002-00002.

Barberis, N.C., 2013. Thirty years of prospect theory in economics: a review and assessment. J. Econ. Perspect. 27 (1), 173–195. https://doi.org/10.1257/jep.27.1.173.
Barch, D.M., Burgess, G.C., Harms, M.P., Petersen, S.E., Schlaggar, B.L., Corbetta, M., et al., 2013. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 80, 169–189. https://doi.org/10.1016/j.neuroimage.2013.05.033.
Bexkens, A., Jansen, B.R., Van der Molen, M.W., Huizenga, H.M., 2016. Cool decision-making in adolescents with behavior disorder and/or mild-to-borderline intellectual disability. J. Abnorm. Child Psychol. 44 (2), 357–367. https://doi.org/10.1007/s10802-015-9996-8.
Blankenstein, N.E., Peper, J.S., Crone, E.A., van Duijvenvoorde, A.C., 2017. Neural mechanisms underlying risk and ambiguity attitudes. J. Cogn. Neurosci. 29 (11), 1845–1859. https://doi.org/10.1162/jocn_a_01162.
Blankenstein, N.E., Schreuders, E., Peper, J.S., Crone, E.A., van Duijvenvoorde, A.C., 2018. Individual differences in risk-taking tendencies modulate the neural processing of risky and ambiguous decision-making in adolescence. Neuroimage 172, 663–673. https://doi.org/10.1016/j.neuroimage.2018.01.085.
Bolck, A., Croon, M., Hagenaars, J., 2004. Estimating latent structure models with categorical variables: one-step versus three-step estimators. Political Anal. 12 (1), 3–27. https://doi.org/10.1093/pan/mph001.
Bollen, K.A., 1984. Multiple indicators: internal consistency or no necessary relationship? Qual. Quantity 18 (4), 377–385. https://doi.org/10.1007/BF00227593.
Bollen, K.A., 2002. Latent variables in psychology and the social sciences. Annu. Rev. Psychol. 53, 605–634. https://doi.org/10.1146/annurev.psych.53.100901.135239.
Bollen, K.A., Long, J.S., 1992. Tests for structural equation models: introduction. Sociol. Methods Res. 21 (2), 123–131. https://doi.org/10.1177/0049124192021002001.
Brett, M., Anton, J.L., Valabregue, R., Poline, J.B., 2002. Region of interest analysis using an SPM toolbox. In: In 8th International Conference on Functional Mapping of the Human Brain, 16(2), p. 497.
Burnham, K.P., Anderson, D.R., 2003. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer Science & Business Media.
Burnham, K.P., Anderson, D.R., 2004. Multimodel inference: understanding AIC and BIC in model selection. Sociol. Methods Res. 33 (2), 261–304. https://doi.org/10.1177/0049124104268644.
Cavanaugh, J.E., 1997. Unifying the derivations for the Akaike and corrected Akaike information criteria. Stat. Probab. Lett. 33 (2), 201–208. https://doi.org/10.1016/S0167-7152(96)00128-9.
Clark, S.L., Muthén, B., Kaprio, J., D'Onofrio, B.M., Viken, R., Rose, R.J., 2013. Models and strategies for factor mixture analysis: an example concerning the structure underlying psychological disorders. Struct. Equ. Model.: A Multidiscip. J. 20 (4), 681–703. https://doi.org/10.1080/10705511.2013.824786.
Cohen, J.D., Daw, N., Engelhardt, B., Hasson, U., Li, K., Niv, Y., et al., 2017. Computational approaches to fMRI analysis. Nat. Neurosci. 20 (3), 304. https://doi.org/10.1038/nn.4499.
Cooper, S.R., Jackson, J.J., Barch, D.M., Braver, T.S., 2019. Neuroimaging of individual differences: A latent variable modeling perspective. Neuroscience & Biobehavioral Reviews Volume 98, March, Pages 29-46 98, 29–46. https://doi.org/10.1016/j.neubiorev.2018.12.022.
Diamantopoulos, A., 2006. The error term in formative measurement models: interpretation and modeling implications. J. Model. Manag. 1 (1), 7–17. https://doi.org/10.1108/17465660610667775.
Diamantopoulos, A., Siguaw, J.A., 2006. Formative versus reflective indicators in organizational measure development: a comparison and empirical illustration. Br. J. Manag. 17, 263–282. https://doi.org/10.1111/j.1467-8551.2006.00500.x.
Dietrich, C., 2010. Decision making: factors that influence decision making, heuristics used, and decision outcomes. Inq. J./Stud. Pulse 2 (02). http://www.inquiriesj..com/a?id=180. (Accessed 10 February 2018).
Dubois, J., Adolphs, R., 2016. Building a science of individual differences from fMRI. Trends Cogn. Sci. 20 (6), 425–443. https://doi.org/10.1016/j.tics.2016.03.014.
Dziak, J.J., Coffman, D.L., Lanza, S.T., Li, R., 2017. Sensitivity and specificity of information criteria. Peer J. PrePrints. https://doi.org/10.7287/peerj.preprints.1103v3.
FitzGerald, T.H.B., Seymour, B., Dolan, R.J., 2009. The role of human orbitofrontal cortex in value comparison for incommensurable objects. J. Neurosci. 29 (26), 8388–8395. https://doi.org/10.1523/JNEUROSCI.0717-09.2009.
Field, A., 2013. Discovering Statistics Using IMB SPSS Statistics, fourth ed. SAGE.
Forstmann, B.U., Wagenmakers, E.J., Eichele, T., Brown, S., Serences, J.T., 2011. Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? Trends Cogn. Sci. 15 (6), 272–279. https://doi.org/10.1016/j.tics.2011.04.002.
Foulkes, L., Blakemore, S.J., 2018. Studying individual differences in human adolescent brain development. Nat. Neurosci. 1. https://doi.org/10.1038/s41593-018-0078-4.
Fox, C.R., Poldrack, R.A., 2009. Prospect theory and the brain. Neuroeconomics: Decis. Mak. Brain 145, 174. https://doi.org/10.1016/B978-0-12-374176-9.00011-7.
Fülöp, J., 2005. Introduction to decision making methods. Laboratory of operations research and decision systems. Comput. Autom. Inst. Hung. Acad. Sci. 1.
Graefe, A., Armstrong, J.S., 2012. Predicting elections from the most important issue: a test of the take-the-best heuristic. J. Behav. Decis. Mak. 25 (1), 41–48. https://doi.org/10.1002/bdm.710.
Gigerenzer, G., Gaissmaier, W., 2011. Heuristic decision making. Annu. Rev. Psychol. 62, 451–482. https://doi.org/10.1146/annurev-psych-120709-145346.
Glöckner, A., Pachur, T., 2012. Cognitive models of risky choice: parameter stability and predictive accuracy of prospect theory. Cognition 123 (1), 21–32. https://doi.org/10.1016/j.cognition.2011.12.002.
Harman, H.H., 1976. Modern Factor Analysis. University of Chicago press.
Hardman, D., Hardman, D.K., 2009. Judgment and Decision Making: Psychological Perspectives. Blackwell, Oxford, ISBN 978-1-4051-2398-3.
Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., Rangel, A., 2011. Transformation of stimulus value signals into motor commands during simple choice. Proc. Natl. Acad. Sci. U. S. A. 108 (44), 18120–18125. https://doi.org/10.1073/pnas.1109322108.
Heinze, G., Schemper, M., 2002. A solution to the problem of separation in logistic regression. Stat. Med. 21 (16), 2409–2419. https://doi.org/10.1002/sim.1047.
Hurvich, C.M., Tsai, C.L., 1989. Regression and time series model selection in small samples. Biometrika 76 (2), 297–307. https://doi.org/10.1093/biomet/76.2.297.
Jacobucci, R., Grimm, K.J., McArdle, J.J., 2016. Regularized structural equation modeling. Struct. Equ. Model.: A Multidiscip. J. 23, 555–566. https://doi.org/10.1080/10705511.2016.1154793.
Jacobucci, R., Brandmaier, A.M., Kievit, R.A., 2018. Variable Selection in Structural Equation Models with Regularized MIMIC Models. https://doi.org/10.31234/osf.io/bxzjf. Pre-print.
Jahn, J., 2007. Introduction to the Theory of Nonlinear Optimization. Springer Science & Business Media.
Jansen, B.R., van Duijvenvoorde, A.C., Huizenga, H.M., 2012. Development of decision making: sequential versus integrative rules. J. Exp. Child Psychol. 111 (1), 87–100. https://doi.org/10.1016/j.jecp.2011.07.006.
Jöreskog, K.G., Goldberger, A.S., 1975. Estimation of a model with multiple indicators and multiple causes of a single latent variable. J. Am. Stat. Assoc. 70 (351a), 631–639. https://doi.org/10.1080/01621459.1975.10482485.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica: J. Econ. Soc. 263–291. https://doi.org/10.2307/1914185.
Karimi, L., Meyer, D., 2014. Structural equation modeling in psychology: the history, development and current challenges. Int. J. Psychol. Stud. 6 (4), 123–133. https://doi.org/10.5539/ijps.v6n4p123.
Kievit, R.A., Brandmaier, A.M., Ziegler, G., Van Harmelen, A.L., de Mooij, S.M., Moutoussis, M., et al., 2018. Developmental cognitive neuroscience using latent change score models: a tutorial and applications. Dev. Cogn. Neurosci. 33, 99–117. https://doi.org/10.1016/j.dcn.2017.11.007.
Kievit, R.A., Romeijn, J.H., Waldorp, L.J., Wicherts, J.M., Scholte, H.S., Borsboom, D., 2011. Mind the gap: a psychometric approach to the reduction problem. J. Psychol. Inq. 22, 67–87. https://doi.org/10.1080/1047840X.2011.550181.
Kievit, R.A., van Rooijen, H., Wicherts, J.,M., Waldorp, L.,J., Kan, K.,J., Scholte, H.,S., Borsboom, D., 2012. Intelligence and the brain: a model-based approach. J. Cogn. Neurosci. 3, 1–9. https://doi.org/10.1080/17588928.2011.628383.
Kline, P., 2014. An Easy Guide to Factor Analysis. Routledge.
Kline, R.B., 2015. Principles and Practice of Structural Equation Modeling. Guilford publications.
Lang, M., 2010. Tax Treaties: Building Bridges between Law and Economics. IBFD, ISBN 978-90-8722-085-3.
Lim, S.-L., O'Doherty, J.P., Rangel, A., 2011. The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. J. Neurosci. 31 (37), 13214–13223. https://doi.org/10.1523/JNEUROSCI.1246-11.2011.
Lin, T.H., Dayton, C.M., 1997. Model selection information criteria for non-nested latent class models. J. Educ. Behav. Stat. 22 (3), 249–264. https://doi.org/10.3102/10769986022003249.
Lubke, G.H., Muthén, B., 2005. Investigating population heterogeneity with factor mixture models. Psychol. Methods 10 (1), 21. https://doi.org/10.1037/1082-989X.10.1.21.
Lubke, G., Muthén, B.O., 2007. Performance of factor mixture models as a function of model size, covariate effects, and class-specific parameters. Struct. Equ. Model. 14 (1), 26–47. https://doi.org/10.1080/10705510709336735.
Mata, R., Nunes, L., 2010. When less is enough: cognitive aging, information search, and decision quality in consumer choice. Psychol. Aging 25 (2), 289. https://doi.org/10.1037/a0017927.
Mata, R., von Helversen, B., Rieskamp, J., 2010. Learning to choose: cognitive aging and strategy selection learning in decision making. Psychol. Aging 25 (2), 299. https://doi.org/10.1037/a0018923.
McCullough, B.D., Vinod, H.D., 2003. Verifying the solution from a nonlinear solver: a case study. Am. Econ. Rev. 93 (3), 873–892. https://doi.org/10.1257/000282803322157133.
McLachlan, G.J., Basford, K.E., 1988. Mixture Models: Inference and Applications to Clustering, vol 84. M. Dekker, New York.
McLachlan, G.J., Peel, D., 2000. Finite Mixture Models. Wiley Series in Probability and Statistics. A Wiley-Interscience Publication.
McLachlan, G.J., Lee, S.X., Rathnayake, S.I., 2019. Finite mixture models. Ann. Rev. Stat. Appl. 6, 355–378. https://doi.org/10.1146/annurev-statistics-031017-100325.
Miller, J.M., Miller, M.D., 2011. Handling quasi-nonconvergence in logistic regression: technical details and an applied example. Interstat 15, 1–22.
Muthén, L.K., Muthén, B.O., 2012. Mplus User's Guide, seventh ed. Muthén & Muthén, Los Angeles, CA.
Newell, B.R., Lagnado, D.A., Shanks, D.R., 2007. Straight Choices: the Psychology of Decision Making. Psychology Press, London.
Nylund, K.L., Asparouhov, T., Muthén, B., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. Struct. Equ. Model. 14, 535–569. https://doi.org/10.1080/10705510701575396.
Osaka, M., Osaka, N., Kondo, H., Morishita, M., Fukuyama, H., Aso, T., Shibasaki, H., 2003. The neural basis of individual differences in working memory capacity: an fMRI study. Neuroimage 18 (3), 789–797. https://doi.org/10.1016/S1053-8119(02)00032-0.
Ochsner, K.N., Ludlow, D.H., Knierim, K., Hanelin, J., Ramachandran, T., Glover, G.C., Mackey, S.C., 2006. Neural correlates of individual differences in pain-related fear and anxiety. Pain 120 (1–2), 69–77. https://doi.org/10.1016/j.pain.2005.10.014.

Pachur, T., Galesic, M., 2013. Strategy selection in risky choice: the impact of numeracy, affect, and cross-cultural differences. J. Behav. Decis. Mak. 26 (3), 260–271. https://doi.org/10.1002/bdm.1757.

Palestro, J.J., Bahg, G., Sederberg, P.B., Lu, Z.-L., Steyvers, M., Turner, B.M., 2018. A tutorial on joint models of neural and behavioral measures of cognition. J. Math. Psychol. 84, 20–48. https://doi.org/10.1016/j.jmp.2018.03.003.

Payne, J.W., 1976. Task complexity and contingent processing in decision making: an information search and protocol analysis. Organ. Behav. Hum. Perform. 16 (2), 366–387. https://doi.org/10.1016/0030-5073(76)90022-2.

Pfeifer, J.H., Allen, N.B., 2012. Arrested development? Reconsidering dual-systems models of brain function in adolescence and disorders. Trends Cogn. Sci. 16 (6), 322–329. https://doi.org/10.1016/j.tics.2012.04.011.

Philiastides, M.G., Biele, G., Heekeren, H.R., 2010. A mechanistic account of value computation in the human brain. Proc. Natl. Acad. Sci. 107 (20), 9430–9435. https://doi.org/10.1073/pnas.1001732107.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53 (5), 793–808. https://doi.org/10.1080/10635150490522304.

Ritchie, S.J., Booth, T., Hernández, M.D.C.V., Corley, J., Maniega, S.M., Gow, A.J., et al., 2015. Beyond a bigger brain: multivariable structural brain imaging and intelligence. Intelligence 51, 47–56. https://doi.org/10.1016/j.intell.2015.05.001.

RStudio Team, 2016. RStudio. Integrated Development for R. RStudio, Inc., Boston, MA. URL. http://www.rstudio.com/.

Rushworth, M.F.S., Noonan, M.A.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. Neuron 70 (6), 1054–1069. https://doi.org/10.1016/j.neuron.2011.05.014.

Sherman, L., Steinberg, L., Chein, J., 2018. Connecting brain responsivity and real-world risk taking: Strengths and limitations of current methodological approaches. Developmental Cognitive Neuroscience 33, 27–41. https://doi.org/10.1016/j.dcn.2017.05.007.

Simpson-Kent, Ivan L., Fuhrmann, Delia, Bathelt, Joe, Achterberg, Jascha, Borgeest, Gesa Sophia, the CALM Team, Kievit, Rogier A, 2019. Neurocognitive reorganization between crystallized intelligence, fluid intelligence and white matter microstructure in two age-heterogeneous developmental cohorts. biorxiv. https://doi.org/10.1101/593509. (Accessed 5 April 2019).

Steinberg, L., Chein, J., 2018. Connecting brain responsivity and real-world risk taking: strengths and limitations of current methodological approaches. Dev. Cogn. Neurosci. 33, 27–41. https://doi.org/10.1016/j.dcn.2017.05.007.

Stoica, P., Selen, Y., 2004. Model-order selection: a review of information criterion rules. IEEE Signal Process. Mag. 21 (4), 36–47. https://doi.org/10.1109/MSP.2004.1311138.

Thurstone, L.L., 1947. Multiple Factor Analysis. University of Chicago Press, Chicago.

Tomarken, A.J., Waller, N.G., 2005. Structural equation modeling: strengths, limitations, and misconceptions. Annu. Rev. Clin. Psychol. 1, 31–65. https://doi.org/10.1146/annurev.clinpsy.1.102803.144239.

Trepel, C., Fox, C.R., Poldrack, R.A., 2005. Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. Cogn. Brain Res. 23 (1), 34–50. https://doi.org/10.1016/j.cogbrainres.2005.01.016.

Turner, B.M., Forstmann, B.U., Love, B.C., Palmeri, T.J., Van Maanen, L., 2016a. Approaches to analysis in model-based cognitive neuroscience. J. Math. Psychol. https://doi.org/10.1016/j.jmp.2016.01.001.

Turner, B.M., Forstmann, B.U., Wagenmakers, E.J., Brown, S.D., Sederberg, P.B., Steyvers, M., 2013. A Bayesian framework for simultaneously modeling neural and behavioral data. Neuroimage 72, 193–206. https://doi.org/10.1016/j.neuroimage.2013.01.048.

Turner, B.M., Van Maanen, L., Forstmann, B.U., 2015. Informing cognitive abstractions through neuroimaging: the neural drift diffusion model. Psychol. Rev. 122, 312–336. https://doi.org/10.1037/a0038894.

Turner, B.M., Rodriguez, C.A., Norcia, T.M., McClure, S.M., Steyvers, M., 2016b. Why more is better: simultaneous modeling of EEG, fMRI, and behavioral data. Neuroimage 128, 96–115. https://doi.org/10.1016/j.neuroimage.2015.12.030.

Turner, B.M., Wang, T., Merkle, E., 2017. Factor analysis linking functions for simultaneously modeling neural and behavioral data. Neuroimage 153, 28–48. https://doi.org/10.1016/j.neuroimage.2017.03.044.

Ullman, J.B., Bentler, P.M., 2003. Structural equation modeling. In: Handbook of Psychol, pp. 607–634. https://doi.org/10.1002/0471264385.wei0224.

Van Duijvenvoorde, A.C., Jansen, B.R., Visser, I., Huizenga, H.M., 2010. Affective and cognitive decision-making in adolescents. Dev. Neuropsychol. 35, 539–554. https://doi.org/10.1080/87565641.2010.494749.

Van Duijvenvoorde, A.C., Figner, B., Weeda, W.D., Van der Molen, M.W., Jansen, B.R., Huizenga, H.M., 2016. Neural mechanisms underlying compensatory and noncompensatory strategies in risky choice. J. Cogn. Neurosci. https://doi.org/10.1162/jocn_a_00975.

Venkatraman, V., Payne, J.W., Bettman, J.R., Luce, M.F., Huettel, S.A., 2009. Separate neural mechanisms underlie choices and strategic preferences in risky decision making. Neuron 62 (4), 593–602. https://doi.org/10.1016/j.neuron.2009.04.007.

Vermunt, J.K., 2010. Latent class modeling with covariates: two improved three-step approaches. Political Anal. 18 (4), 450–469. https://doi.org/10.1093/pan/mpq025.

Volz, K.G., Schooler, L.J., Schubotz, R.I., Raab, M., Gigerenzer, G., Von Cramon, D.Y., 2006. Why you think Milan is larger than Modena: neural correlates of the recognition heuristic. J. Cogn. Neurosci. 18 (11), 1924–1936. https://doi.org/10.1162/jocn.2006.18.11.1924.

Volz, K.G., Schooler, L.J., von Cramon, D.Y., 2010. It just felt right: the neural correlates of the fluency heuristic. Conscious. Cognit. 19 (3), 829–837. https://doi.org/10.1016/j.concog.2010.05.014.

Vrieze, S.I., 2012. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol. Methods 17 (2), 228. https://doi.org/10.1037/a0027127.

Wang, J., Wang, X., 2012. Structural Equation Modeling: Applications Using Mplus. John Wiley & Sons, New York, ISBN 978-1-119-97829-9.

Wolf, E.J., Harrington, K.M., Clark, S.L., Miller, M.W., 2013. Sample size requirements for structural equation models: an evaluation of power, bias, and solution propriety. Educ. Psychol. Meas. 73 (6), 913–934. https://doi.org/10.1177/0013164413495237.

Yarkoni, T., Braver, T.S., 2010. Cognitive neuroscience approaches to individual differences in working memory and executive control: conceptual and methodological issues. In: Handbook of Individ. Differ. In Cogn., 87-107. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1210-7_6.

Zeisberger, S., Vrecko, D., Langer, T., 2012. Measuring the time stability of prospect theory preferences. Theory Decis. 72 (3), 359–386. https://doi.org/10.1007/s11238-010-9234-3.