The Iowa Gambling Task (IGT) is a popular test for assessing executive functions and decision-making deficits in several clinical populations, especially if the symptomatology includes excessive risk taking. The IGT is a complex task which requires participants to correctly integrate and remember information, evaluate consequences and make a decision. However, there are many papers studying different models of IGT variance for multiple purposes (see overview see Bush & Sahak, 2009). IGT results assess a variety of the test scores constitution. Moreover, the majority of studies use "traditional IGT scores" mainly based on the differences between "good" and "bad" blocks or on the total net scores, although there are other models and up-to-dates methods for IGT scoring.

One of the alternative approaches is Bayesian Cognitive Modeling, for example PVL-delta model (Stinroest et al., 2013; 2014). However, these models are not widely used. High computational demands can be one of the possible reasons, other is a high theoretical nature of cognitive modeling combined with live knowledge about the parameter validity. Although there is a few studies on differences between clinical and normal population (e.g. Stinroest et al., in press), convergent validity to classical test scores or other tests of executive functions is unknown.

The question is if the Bayesian parameters can be used in similar way as the traditional IGT scores for diagnostic purposes. The goal of this exploratory study is to present results of convergent validity analysis of parameters from PVL-delta model to the classical IGT test scores and to other tasks of executive functions.

Participants and data collection

Data were collected in PC lab during approx. 3 hours between other questionnaires in two waves as a part of bigger data. The others)

Data were analyzed with Spearman correlation. Reliability of difference scores (e.g. total IGT score) was estimated using reliabilities of subs tests and their Pearson’s correlation (e.g. Thomas & Zambre, 2012). The estimates can be biased due to non-normal score distributions.

Test Materials

Iowa Gambling Task (IGT): Bayesian Cognitive Modeling Approach

We used classical IGT design (e.g. Bechara et al., 1994) twice longer (200 trials). The PVL-delta model use four parameters to predict block choices:

- Shape parameter (A), which determines the shape of the "utility function". Greater A-parameter means higher respondent's sensitivity to feedback outcomes.
- Loss aversion parameter (w). People with low w-parameter are more sensitive to gains than to the losses, people with high w-parameter are more sensitive to losses than to gains.
- Updating parameter (x), which quantifies memory and learning. Its high value indicates that the recent card outcome has a strong influence on the expected utility of blocks and forgetting is quick.
- Response consistency (y), which captures the exploitation vs. exploration process. Low value indicates more random choices (more exploitative behavior), while the high value indicates strong choices dependency on expected utility of block (more exploitative behaviour).

The PVL-delta model was estimated using Markov chain Monte Carlo (MCMC) sampling algorithm in Stan and R. The procedure is described in our previous work (Šmírek, 2014). For 22 participants, MCMC did not converged (theta > 1.83) and they were excluded. All the scripts are available on-line: https://github.com/nazarke/NICC-IGT.

Reliability estimation of PVL-delta parameters

According to our knowledge, there has not been any published procedure for reliability estimation of Bayesian cognitive modeling parameters. We propose two approaches:

(1.) As reliability is defined as the correlation between two parallel tests, we sampled a pair of each parameter from its posterior distribution for each respondent, and correlated these pairs. We repeated it 1000times for each parameter what results to posterior distributions of reliability estimates. All the distributions were approx. normally distributed with means [and 95% CIs]: rA = -.74 [-.82, -.66]; rW = .84 [0.79, 0.86]; rX = .89 [0.83, 0.94]; rY = .87 [0.81, 0.91].

(2.) For each respondent (and each parameter), we computed variance of posterior distribution. Mean of these error variances is a mean-square error (MSE) estimate. The reliability is then r = 1 - MSE/var(r), where var(r) is a point (means) parameter estimate between subjects. Because this approach assumes normality for posterior distributions, estimates are slightly smaller: rA = -.56; rW = .81; rX = .87; rY = .84.

Iowa Gambling Task (IGT): Traditional IGT scores

LTC (earning of long term consequences) is the difference between number of choices from good and bad decks: (C-D)−(A-B). IRLBS (bias for infrequent losses) is the difference (B+D)−(A+C). PGE is the preference of good deck among high frequent losses deck: (D-B). TOTAL is a final total net score.

Delay Discounting Task (DDT)

We used two conditions, 900 CZK (37 EUR) or 24000 CZK (950 EUR), and week/month/three month/one year delays. For both conditions we estimated original rate (e.g. Mazur, 1987) and AUC was estimated according to r = bP (of k parameter to the data; higher value suggests lower fit). Higher values of k and lower values of AUC suggest preference of smaller amount now compared to higher reward later. Total score were means of parameters from both conditions.

Go/No-Go Task (GNG)

We used 192 trials with two conditions divided into four blocks. Scores: relative frequency wrong on go (GOweong) and no-go tasks (NOGOweong); mean reaction time on go (GORT) and no-go (NOGORT) tasks.

Stop Signal Task (SST)

192 trials divided into four blocks were used. Scores: relative frequencies of commission (COM), omission (OM) and correct (COR) responses alongside with mean reaction time (RT) on go task, stop-signal delay (SD) and stop-signal reaction time (RSST = RT-SD). We worked also relative with frequency of fail on no-go tasks, but its blocks were not positively correlated (assoc. = 0.61) and therefore this score was excluded.

N-Back

We used 2-back and 3-back tasks, both divided into four blocks with 24 tasks each (6 ‘yes’ conditions). Score (N-back) was defined as the total correct of the ‘yes’ trials.

Spatial Stroop Task (Sstroop)

Position on the screen was combined with words up/down/ left/right, respondents used arrow keys. There were four blocks: neutral, spatial, neutral, congruent and incongruent, where respondents had to mark the sense of the word, however its position was different. The total score (Sstroop) was estimated as the difference between the mean reaction time in the 4th and 2nd block.

Number-Letter Task (NL)

Respondents were exposed to the pairs of a numbers and a letters. Based on a top/down position, they had to decide if the letter is consonant/vowel, or if the number is odd/even. One third of 48 trials was congruent (both led to the same key). Scores: total correct (NCorr) and the difference of the mean reaction time between switch and repeat trials (NL).

Results and Discussion

Left-bottom table contains Spearman correlations and reliabilities of all the scores. We estimated also network model (chart). We used regularized partial correlation network (lASSO algorithm, ERIC minimized) in bootstrapped R package (Epskamp, Borsboom, & Fried, 2017) with nonparametrical transformation to Gaussian distribution (large package).

We believe that the origin correlation matrix is not fully interconnected and correlations between the two blocks of methods are just a product of transformation. These results are appropriate for exploratory purposes only.

Bayesian cognitive scores correlates highly with classical test score, but not classical nor Bayesian IGT score are well correlated to other cognitive tasks; however, correlations between different tests are generally low. Average of Fisher r-transformation absolute values of correlations with other tests was higher for Bayesian scores (.104) than for traditional scores (.021), though the effect was small, Cohen’s d ≈ 0.39.

Conclusion: Bayesian scores seem to have slightly higher concurrent validity to other cognitive tasks, however correlations are generally very small. Also tests which are supposed to measure the same traits, e.g. NL and Stroop or SST and GNG, share only a small amount of variance. However, Bayesian IGT cognitive parameters correlated with classical IGT scores and can be probably used in similar way. Finally, a and c parameters were highly correlated (especially if we take into account their reliabilities), they can probably measure the same traits. We advice to use their mean (after reiterating one of them) for practical purposes and probably merge them in next generation of PVL-delta model.

References


