The Autonomy-Validity Dilemma in Mechanical Judgment Procedures: The Quest for a Compromise

Marvin Neumann, Susan Niessen, Jorge Tendeiro, and Rob Meijer

university of groningen

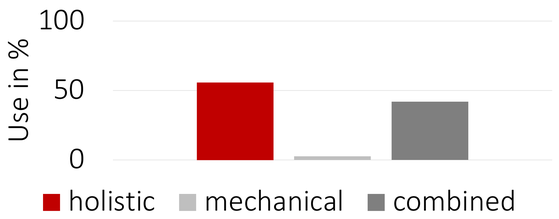
Introduction

Background

- In personnel- and educational selection, information from multiple assessments (e.g., test scores and interview ratings) is often used, which can be combined in two ways[1,2]:
  - Holistic judgment: information is subjectively combined in the mind
  - Mechanical judgment: information is combined with an explicit decision rule
    - Prediction = predictor 1 * w1 + predictor 2 * w2 …
- Mechanical judgment is on average more valid than holistic judgment[1,2]

The problem

Holistic judgment dominates in practice[3,1]



Contribution

- Decision makers may use mechanical judgment more often when they retain autonomy
  - Decision makers could choose predictor weights (w1, w2)[4]
  - Decision makers could holistically adjust predictions[5]
- Research questions:
  1. Do decision makers prefer autonomy-enhancing judgment procedures, compared to strictly using an optimal decision rule?
  2. How does increased autonomy affect predictive validity?

Conclusion

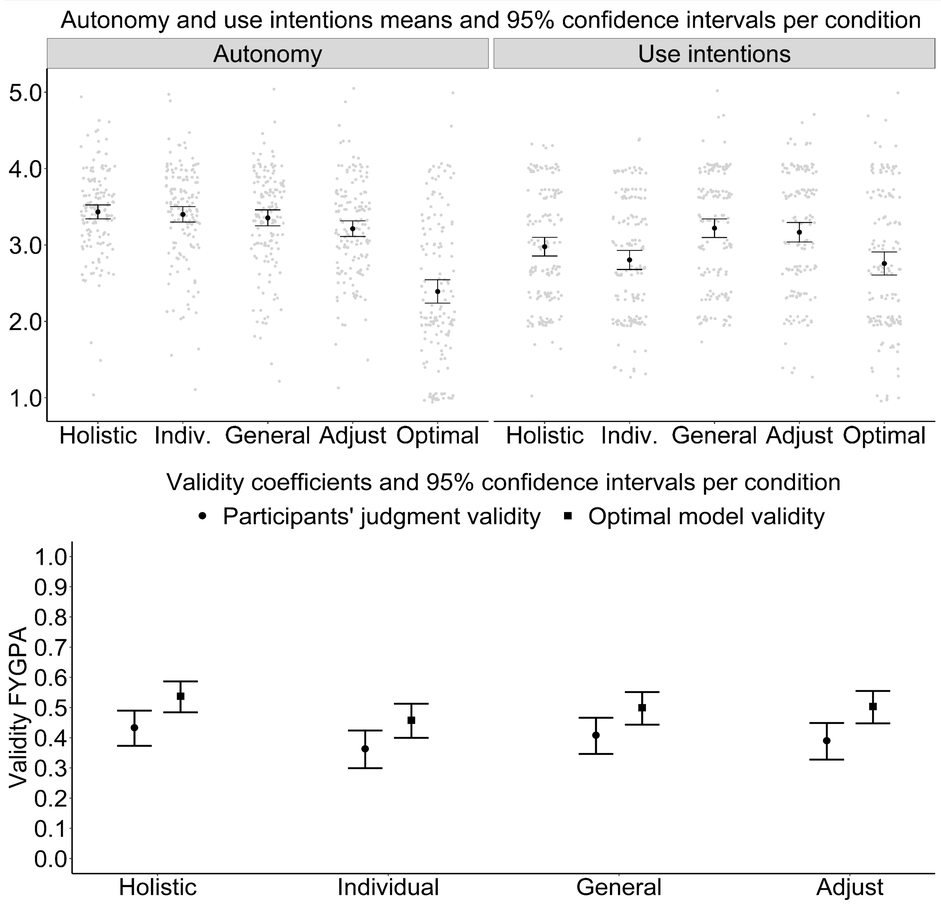
- Two promising procedures in terms of an autonomy-validity tradeoff emerged
  1. Choosing general weights when predictor validity information is available
  2. Holistically adjusting optimal model predictions
- Yet, our results prevent a clear conclusive statement regarding a compromise between autonomy and validity

Method

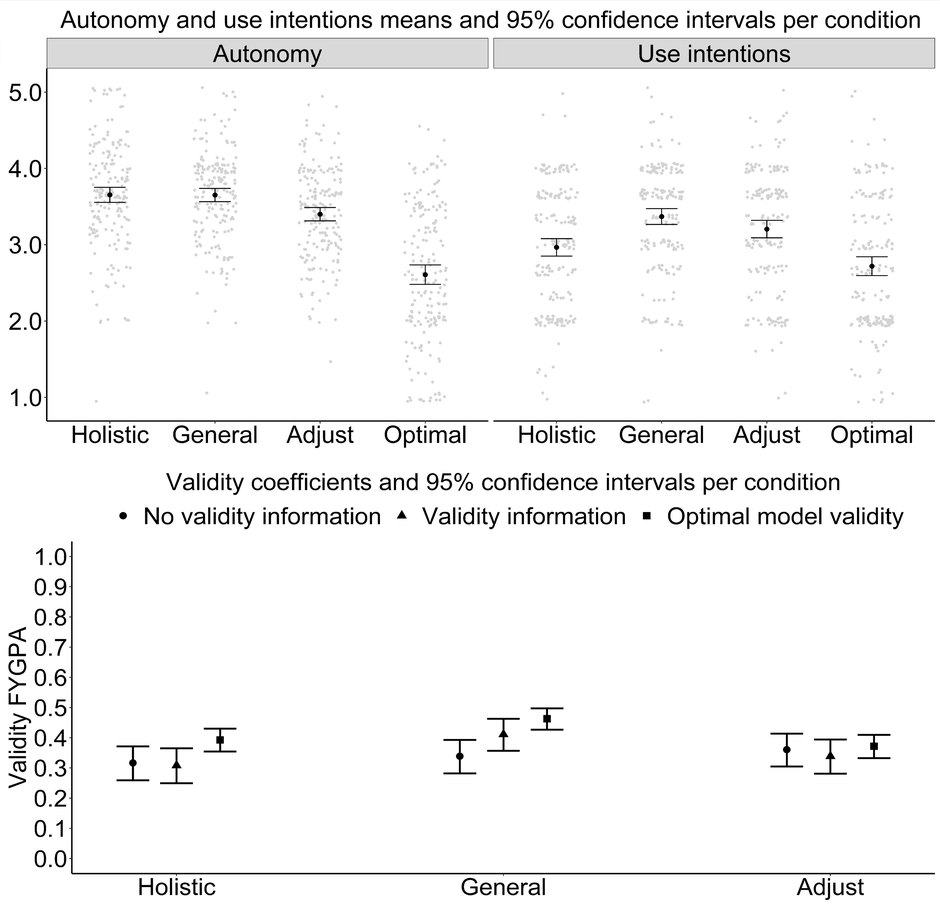
- Prediction task: Predict first-year GPA (FYGPA) of 5 (10 in Study 2) applicants using high school GPA, admission test scores, and personal statements. Participants (students) were informed about predictor validities
- Study 1 (*N* = 150): within-subjects design. Autonomy in making predictions was varied in five conditions
  1. Holistic: Holistic (subjective) predictions based on the predictor scores
  2. Individual: Assign percentage predictor weights for each of the applicants judged
  3. General: Assign one set of percentage predictor weights for all applicants
  4. Adjust: Participants adjusted the predictions of a statistical model unrestrictedly
  5. Optimal: Participants imagined a statistical model would make predictions that they could not adjust
- Study 2 (*N* = 192): mixed design
  - Same within-subjects factor as in Study 1. The “individual” condition was dropped because Study 1 results were not promising. Furthermore, participants could only restrictedly adjust model predictions in the “adjust” condition
  - Between-subjects factor: A random half of participants was not informed of predictor validities

Results and Discussion

Study 1



Study 2



- Perceived autonomy: similar across conditions, but much lower in the “optimal” condition (e.g., general vs. optimal, $d = 1.17$ and $d = 1.35$ in Study 1 and 2, respectively)
- Use intentions: higher in all autonomy-enhancing conditions than in the “optimal” condition (e.g., general vs. optimal, $d = 0.54$ and $d = 0.81$ in Study 1 and 2, respectively)
- Predictive validity: similar across conditions, but optimal model predictions were always better than participants’ predictions. Knowing predictor validities only slightly increased predictive validity in the “general” condition

Key references

- [5]Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. *Management Science*, *64*, 1155–1170. https://doi.org/10.1287/mnsc.2016.2643
- [1]Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis. *Journal of Applied Psychology*, *98*, 1060–1072. https://doi.org/10.1037/a0034156
- [2]Meehl, P. E. (1954). Empirical comparisons of clinical and actuarial prediction. In Clinical versus statistical prediction: A theoretical analysis and a review of the evidence (pp. 83–128). Minneapolis, MN: University of Minnesota Press. https://doi.org/doi:10.1037/11281-008
- [4]Nolan, K. P., & Highhouse, S. (2014). Need for autonomy and resistance to standardized employee selection practices. *Human Performance*, *27*, 328–346. https://doi.org/10.1080/08959285.2014.929691
- [3]Ryan, A. M., & Sackett, P. R. (1987). A survey of individual assessment practices by I/O psychologists. *Personnel Psychology*, *40*, 455–488. https://doi.org/http://dx.doi.org/10.1111/j.1744-6570.1987.tb00610.x