

A Reproducibility Study on User-centric MIR Research and Why it is Important

Peter Knees^{1,3}, Bruce Ferwerda², Andreas Rauber¹, Sebastian Strumbelj¹, Annabel Resch¹, Laurenz Tomandl¹, Valentin Bauer¹, Tang Fung Yee¹, Josip Bobinac¹, Amila Ceranic¹ & Riad Dizda¹

¹TU Wien (Austria), ²Jönköping University (Sweden) & ³Georgia Institute of Technology (USA)

Research Goal: reproducibility of user-centric studies using the PRIMAD model

Take Away: using the PRIMAD model of reproducibility, reproducibility could be systematically examined. In general, reproducibility was achieved with minor differences. These differences found could be accounted to cultural differences. To improve reproducibility in the future, studies and methodologies should be documented better.

INTRO

Reproducibility in MIR common practice

- System-centric
- Test performance consistency
- Documentation of workflows and processes
- Access to data

Important to reproduce user studies

- User-centered approaches towards systems
- Basis of requirements engineering
- Foundation and justification of algorithmic choices

REPRODUCIBILITY

Kim, H. J., Park, S. Y., Park, M., & Lee, K. (2020).

Do Channels Matter? Illuminating Interpersonal Influence on Music Recommendations. In *Fourteenth ACM Conference on Recommender Systems*

- Implications for development and understanding music recommender systems
- Well documented study and methodology
- Web-based survey

PRIMAD MODEL OF REPRODUCIBILITY

Different degrees of reproducibility to vary in:

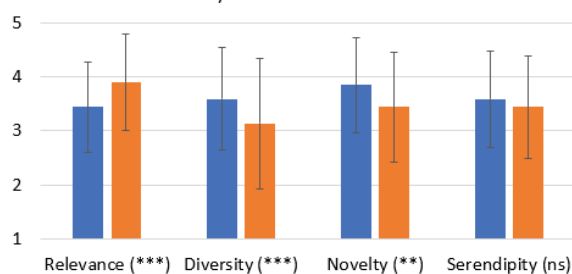
- Platform
- Research Objective
- Implementation
- Method
- Actor
- Data

A simple replication of experiments under completely identical conditions does not lead to any knowledge gains

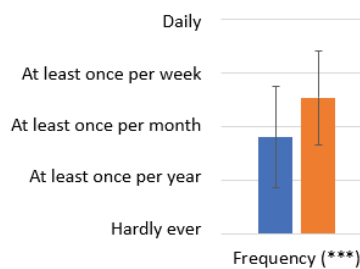
CONCLUSION

- PRIMAD model of reproducibility is an effective model to systematically test reproducibility
- Reproducibility was achieved:
 - Major findings are the same
 - Minor differences due to cultural differences
 - Differences in adaption rates between channels and factors influence adaption rate
- Urgent need to start reproducing user-centric MIR research
- More thorough and detailed documentation is needed on how studies were conducted

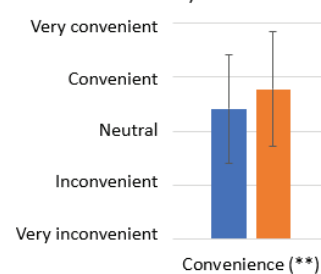
A) Evaluation Criteria



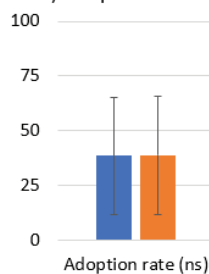
B) Frequency



C) Convenience



D) Adoption Rate



■ Interpersonal ■ Non-interpersonal