

Extended time accommodations in employment testing: A reasonable accommodation, or a detriment to test validity?

Kris Shiflet

Dept. of Psychology, Illinois Tech

Psyc 597: Special Problems

Dr. Kristina Bauer

December 4, 2024

Contents

Introduction	3
Background and Context.....	3
Extended Time Accommodations	6
Differential Prediction	6
Differential Benefits	7
Fairness of the Accommodation.....	8
Additional Considerations	8
Investigating Validity of Extended Time Accommodations	9
Sample.....	9
Method	10
Analysis and Expected Results	10
Limitations and Future Research	11
References	13
Appendix A	16

Introduction

Employers frequently use batteries of standardized tests for the purposes of selection, promotion, and assessment. These test batteries can include cognitive tests, personality tests, job sample tests, etc. If done improperly, especially with intent to discriminate, usage of such tests can break federal law.

The Americans with Disabilities Act of 1990 (ADA) prohibits the use of employment testing procedures that would discriminate against people with disabilities, and stipulates that employers give reasonable accommodations to those who need them. Even though people with disabilities consist of about 20 percent of the workforce in the United States (Bureau of Labor Statistics, 2024), little research has been done on disability discrimination in employment testing, and even less on potential impacts on test validity due to accommodations. Additionally, while organizations such as the Job Accommodation Network (JAN) provide guidance to employers and employees alike on potential accommodations for employment testing, there is limited research behind the appropriateness and utility of these suggested accommodations.

To the author's knowledge, nearly all extant research on how accommodations impact test validity comes from the field of educational testing. There have been a plethora of studies looking at the appropriateness and utility of accommodations in educational testing. Further, while several of the discoveries there are illuminating, the results might not translate to a professional setting. Thus, there is a need to explore the impacts of extended time accommodations in workplace testing to see which results carry over from the field of education to workplace testing, especially given the importance and potential impact of improper employment decisions.

Background and Context

It is only within the past 60 years that testing accommodation has become particularly important to psychometricians and educators in the United States. The passage of the Rehabilitation Act of 1973 implemented requirements for schools to provide accommodations to students with disabilities in the class and testing environments (U.S. Department of Labor, n.d.). A basic perusal of the literature on testing accommodations shows that most of the research really began around the 1980s, with another uptick in interest after the passage of the ADA in 1990. The ADA added further guidelines for accommodations in school, as well as implemented requirements for businesses to accommodate people with disabilities in all steps of the employment cycle (U.S. Department of Labor, 2024).

In the seminal article, *High-stakes testing accommodations: Validity versus disabled rights*, Philips (1994) claims that testing accommodations have the potential to impact the validity and psychometric properties of a test. Therefore, when testing accommodated and non-accommodated individuals, there is a chance that the scores are not truly comparable due to this distortion of psychometric properties. Specifically, Philips (1994) notes that different types of accommodations can impact the validity of a test in different ways. Philips (1994) states that accommodations for physical disabilities are of little debate as one's physical limitations generally do not limit one's cognitive abilities. However, many testing accommodations are requested and made for individuals with cognitive disabilities, such as dyslexia or attention deficit disorder. Accordingly, cognitive disabilities can have a marked impact on certain cognitive and behavioral skills, it is difficult to determine what accommodations might be appropriate, if any. In such cases, accommodations might not just change the delivery of the test, but also the skills being tested, thus "invalidating" the inference made from the test score (Phillips, 1994). How, then, does one appropriately accommodate a person with dyslexia on a

test of reading and verbal reasoning? If test modality is changed from written to oral, then the skill being tested will change from visual reading to understanding of spoken words. Thus, the scores for the accommodated person(s) are not comparable with those of unaccommodated persons.

Conversely, Iliescu & Greiff (2022) posited that standard, unaccommodated tests, will provide results that are not valid for disabled individuals. If done properly, testing accommodations can equalize the playing field and enhance the validity of tests for people with disabilities (Iliescu & Greiff, 2022). This is the key to testing: balancing accessibility and test validity. In education and employment, tests of skill and ability are assumed to be valid predictors of future performance (Iliescu & Greiff, 2022). Consequently, testing accommodations must be used to obtain test scores from which valid inferences can still be drawn (Kettler, 2012; Elliott & Marquart, 2004).

Some may question the efficacy of these accommodation efforts since they appear to impact a seemingly small proportion of the population. However, there are several reasons to be concerned about accommodation. Firstly, as stated previously, disabled people make up about 20 percent of the workforce (U.S. Department of Labor, 2024). That is a sizeable percentage of the workforce that may need to be accommodated. Additionally, more and more people with disabilities are entering higher education, and, thus, the workforce (Ling, Buzick, & Belur, 2020), which means the need for proper testing will only continue in the future. Secondly, the process of employment testing is resource-intensive, and decisions made based on tests with poor validity can have far-reaching and costly impacts (Cascio & Aguinis, 2019). It is imperative to employers and their workers, shareholders, customers, and clients that personnel decisions are made on sound reasoning and meaningful test results.

Extended Time Accommodations

Extended time accommodations are one of the most regularly proffered accommodations on employment tests (Campbell, et al., 2002; Furlano, et al., 2021). The logic behind why extended time accommodations are given is relatively straightforward: disabilities can prevent individuals from demonstrating their skills within the same time limits as others, and testing them under the standard conditions can introduce construct-irrelevant variance in the test scores. The scores will have variance due to the examinee's disability status, not just variance from the construct being measured (Lovett, 2010). In practice, however, the validity, fairness, and utility of extended time accommodations has come into question numerous times within the past 30 years. The research that has come from this questioning has produced varied results. The key facets of extended time accommodation research I wish to focus on here include differential prediction and differential benefits.

Differential Prediction

Research done on samples of high school and college students taking the Scholastic Aptitude Test (SAT), Law School Admissions Test (LSAT), and Graduate Record Examination (GRE) have shown differences in predictive validity between accommodated and non-accommodated students (Lovett, 2010; Zurcher & Bryant, 2001; Lewandowski, Cohen, & Lovett, 2013; Koenig & Bachman, 2004). There are several pieces of research that show that the predictive validity of these standardized exams is weaker for disabled students who took extended time accommodations. For example, a study by Cahalan, Mandinach, and Camara (2002) found that SAT scores of students with disabilities who received extended time correlated less with grade point average (GPA) than the scores of non-disabled, non-accommodated students ($r = .35$ versus $r = .48$, respectively), with the tests overpredicting for students with

disabilities. Interestingly, a study comparing the MCAT scores of non-disabled students and students who have taken extended time accommodations found no significant differences in Medical College Admission Test (MCAT) scores between the groups, but the test still overpredicted for those who took the test with extended time (Ling, Buzick, & Belur, 2020).

These results indicate that, while the values of the scores between the accommodated and non-accommodated groups appear similar, the scores have different predictive validity and utility. In short, it seems that extended time accommodations can artificially inflate test scores such that they are not an accurate predictor of later performance. In the context of employment testing, tests are typically used to find which candidates are most suited for a particular role. If these tests overpredict performance, then it is possible for companies to hire or promote someone to a position who is not suitably qualified for that role.

Hypothesis 1: Individuals who use extended time on an employment test and individuals who do not use extended time will have approximately equivalent scores on the test.

Hypothesis 2: Individuals who use extended time on an employment test will have, on average, lower performance ratings than individuals who did not use extended time.

Hypothesis 3: The employment test scores for individuals who use extended time will be less predictive of performance rating compared to individuals who do not use extended time on employment tests.

Differential Benefits

There are those that argue that accommodations must have differential benefits in order to be acceptable. That is, an accommodation should only provide significant benefit to disabled examinees, not examinees without disabilities (Hollenbeck, 2005; Pitoniak & Royer, 2001; Furlano, et al., 2021). The research regarding differential benefits of extended time accommodations is mixed. Select research shows that disabled examinees and non-disabled examinees both benefit from extended time accommodations (Kettler, 2012), and there is also

research that indicates that non-disabled individuals might benefit more from extended time than disabled individuals (Lewandowski, Cohen, & Lovett, 2013). There is some research which shows that it is only disabled test-takers who receive significant benefit from extended time accommodations, and the samples in this research were usually older students in high school and college settings (Elliott & Marquart, 2004).

Unfortunately, Furlano et al. (2021) found that extended time accommodations benefited both disabled and non-disabled candidates in an employment aptitude test for military personnel. The researchers concluded that organizations and testing groups should re-examine the usage of extended time accommodations to certain groups to ensure that they do not unintentionally disadvantage others. Furlano et al. (2021) was the only piece of research I could find on extended time accommodations in the employment context. This mixed bag of research regarding the benefit of extended time accommodations shows that there is a need for further research in that area, especially in the context of employment testing.

Fairness of the Accommodation

The issues previously discussed have caused some to call into question the fairness of extended time accommodations (Lovett, 2010). In the context of employment, there are several ways that extended time accommodations could lead to perceived unfairness. For instance, the overprediction of performance for accommodated individuals could lead to a disabled individual getting a job or promotion that they are not properly equipped for, which could lead to poor performance and lowered self-esteem. Further, this could be detrimental to the organization. If a person does not do their job right, it can damage the organization's reputation and then finding a suitable replacement can be quite costly.

Additional Considerations

Whether someone has taken an extended time accommodation is not the only factor that might impact Test Score or Performance Rating. Since there is almost no research on this topic, we also ask if Job Type, Test Type, and Disability Type play a role in test and performance outcomes.

Research Question 1: Do Job Type, Test Type, or Disability Type have an effect on the Test Scores and predictive validity of employment tests, and does the impact differ between those who take extended time versus those who do not?

Investigating Validity of Extended Time Accommodations

In order to grapple with the dearth of research on extended time accommodations in an employment setting, I propose a study of current members of the general workforce to examine predictive validity of standardized employment selection tests for disabled and non-disabled individuals.

Sample

The sample for this study would be recruited through networking and snowballing. The research team would reach out to various companies with employment testing and request that they share test scores and performance review scores from consenting employees. Requests will start within-network, with researchers contacting other HR/IO specialists working in firms across the country. The data will be anonymized, with only disability status, age, gender, and the scores available. In order to get ample participation from people with disabilities, we will make sure to reach out to companies with an employee resource group (ERG) for disabled employees. We will target companies who use standardized measures of constructs such as intelligence, personality, general job knowledge, and emotional intelligence, as non-standardized measures would require more specific research. Companies will also be required to use a numeric scale of some sort for their performance reviews.

Method

Once participants have agreed to participate in the study, their information will be sent to the researchers by the company. We will keep the data collection window open for nine months to allow participants to trickle in, as the networking/snowball recruiting method can take a long time to get an ample sample size. Once the data is collected, it will be standardized for analysis and compiled into a spreadsheet where there will be columns for Name, Age, Job Type, Gender, Disability Status, Disability Type, Test Type, Accommodation, Test Score, Performance Rating. Collecting data for these specific variables is important to help control for different interaction effects that might arise that would influence the relationship between accommodation usage and predictive validity.

Test Scores will be converted to a standard 1-100% scale and Performance Rating will be standardized to a 1-5 (1 = poor performance, 3 = adequate performance, 5 = superior performance) scale for ease of analysis. For example, if a company rates its employees on a 1-3 scale, we might change it to be a 1-5 scale with the original values of 1, 2, and 3 taking on values of 1, 3, and 5, respectively. See Appendix A for more detailed information about each variable being studied.

Analysis and Expected Results

To test Hypothesis 1, a t-test will be conducted to determine which group (accommodated vs non-accommodated) has higher scores on the employment test. This process will be repeated for Performance Rating in order to test Hypothesis 2. Cohen's d (effect size) will also be calculated when studying mean differences. Hypotheses 1 and 2 are expected to be supported by this study.

To test Hypothesis 3, a comparison of correlations of Test Score vs Performance Rating between the accommodated and non-accommodated groups will be done. This will be done by conducting a correlation coefficient for both groups and using the Fisher z-transformation to test if the difference between the two correlations is significant. R^2 will be calculated for all correlations. Hypothesis 3 is expected to be supported.

Finally, to explore Research Question 1, hierarchical linear regression will be used to explore if Test Type, Job Type, and Disability Type moderate the relationship between Test Score and Performance Rating. Additional Fisher z-transformations will be conducted, as needed, to do more subgroup comparisons if the hierarchical model shows that the different categorical variables impact the relationship between Test Score and Performance Rating. R^2 will be calculated for all correlations. The results of the test for Research Question 1 are uncertain, as there is no extant literature, to the author's knowledge, that discusses the topic.

Limitations and Future Research

This set of studies would, hopefully, be a useful addition to the sparse literature of accommodations in employment testing, though it would not be without its limitations. First and foremost, it is difficult to get a large sample size when studying individuals with disabilities. It is likely that we would run into this issue while conducting this research, thus minimizing the power and generalizability of our results. Especially since we are going to ask for test results, individuals might be uncomfortable with their scores being shared, despite the anonymity of the data.

Second, since we opted to look at standardized tests, it will be important for future research to investigate the impacts of extended time accommodations on assessments such as situational judgment tests or specific job knowledge tests. Companies that have constructed their

own assessments, or have contracted consultants to construct assessments, would likely benefit from specific research on their tests to make sure that their methods are reliable and valid for those who do use extended time accommodations as well as those who do not.

Third, while we will be using different analyses to study various factors outside of extended time usage, there are plenty of potential variables that could influence our results. If test scores for folks who use extended time accommodations are overpredicting, is it because of actual performance differences, or something else? If the employee does not get extended time on work tasks day-to-day, is a test with extended time really an appropriate predictor of job performance? In the future, it would be pertinent to evaluate whether extended time accommodation makes sense in the context of a job and create a set of guidelines for when such an accommodation is and is not appropriate, if ever.

It will also be pertinent to investigate the validity of extended time accommodation across different industries. While Job Type is accounted for in this study, more detailed examination of trends and phenomena in certain industries should also be studied.

Additionally, this study focused on the validity of extended time accommodation, and future research should also study whether there is differential benefit of the accommodation for individuals with and without disabilities. This was something we would not be able to research effectively due to the structure of our study, but it is still an important piece of data to help determine if extended time accommodation is appropriate.

Finally, if it does turn out that extended time does, indeed, impair the predictive validity of employment tests, there will need to be research into a) why that is so and b) what alternatives might there be to accommodate people with disabilities.

References

- Campbell, W., Fox, H. R., Bruyère, S., Tippins, N., Rosenfeld, M., Keiser, S., & Goldsmith, S. (2002). V. Employment, Certification, and Licensing. In R. Ekstrom, & D. Smith, *Assessing individuals with disabilities in educational, employment, and counseling settings* (pp. 191-248). Washington DC: American Psychological Association.
- Cascio, W., & Aguinis, H. (2019). *Applied psychology in talent management* (8th ed.). SAGE Publications, Inc.
- Elliott, S., & Marquart, A. (2004). Extended time as a testing accommodation: Its effects and perceived consequences. *Council for Exceptional Children*, 70(3), 349-367.
- Furlano, R., Goodrich, S., Darr, W., Gibbard, K., Ho, J., Wang, Z., . . . Hausdorf, P. (2021). Determining the appropriateness of extended time accommodations in standardized cognitive ability testing. *Canadian Journal of Behavioral Sciences*, 53(2), 152-163. doi:10.1037/cbs0000257
- Hollenbeck, K. (2005). Validity issues and decisions about testing accommodations. *Assessment for Effective Intervention*, 31(1), 7-17.
- Iliescu, D., & Greiff, S. (2022). Some thoughts and considerations on accommodations in testing. *European Journal of Psychological Assessment*, 38(4), 239-242. doi:10.1027/1015-5759/a000732
- Kettler, R. (2012). Testing accommodations: Theory and research to inform practice. *International Journal of Disability, Development, and Education*, 59(1), 53-66. doi:10.1080/1034912X.2012.654952

- Koenig, J. A., & Bachman, L. (2004). *Keeping score for all: The effects of inclusion and accommodation policies on large-scale educational assessments*. Washington, DC: The National Academies Press.
- Lewandowski, L., Cohen, J., & Lovett, B. (2013). Effects of extended time allotments on reading comprehension performance of college students with and without learning disabilities. *Journal of Psychoeducational Assessment*, 31(3), 326-336.
doi:10.1177/0734282912462693
- Ling, G., Buzick, H., & Belur, V. (2020). The GRE and students with disabilities: A validity study at 10 universities. *Exceptional Children*, 86(2), 193-208.
doi:10.1177/0014402919872688
- Lovett, B. (2010). Extended time testing accommodations for students with disabilities: Answers to five fundamental questions. *Review of Educational Research*, 80(4), 611-638.
doi:10.3102/0034654310364063
- Phillips, S. (1994). High-stakes testing accommodations: Validity versus disabled rights. *Applied Measurement in Education*, 7(2), 93-120.
- Pitoniak, M., & Royer, J. (2001). Testing accommodations for examinees with disabilities: A review of psychometric, legal, and social policy issues. *Review of Educational Research*, 71(1), 53-104.
- U.S. Department of Labor. (2024). *Disability employment statistics*. Retrieved from DOL.gov:
<https://www.dol.gov/agencies/odep/research-evaluation/statistics>
- U.S. Department of Labor. (n.d.). *Accommodations*. Retrieved from DOL.gov:
<https://www.dol.gov/agencies/odep/program-areas/employers/accommodations>

Zurcher, R., & Bryant, D. P. (2001). The validity and comparability of entrance examination scores after accommodations are made for students with LD. *Journal of Learning Disabilities, 34*(5), 462-471.

PROPERTY OF KRIS SHIFLET - DO NOT FEED TO AI, DO NOT DISTRIBUTE WITHOUT PERMISSION

Appendix A

Table 1. Variable Guide

Variable	Type	Values
Name	Categorical	Whatever the full names of the participants are
Age	Quantitative	Any value above 18
Gender	Categorical	Female, Male, Other/Non-Binary
Accommodation	Categorical	Yes or No
Test Score	Quantitative	Scores will be standardized to fall on a 1-100% scale
Performance Rating	Quantitative	Scores will be standardized to fall on a 5-point scale
Disability Status	Categorical	Disabled or Not Disabled
Disability Type	Categorical	Physical, Mental, None
Test Type	Categorical	Intelligence, Job Knowledge, Emotional Intelligence, Personality
Job Type	Categorical	Retail, Finance, Legal, HR, Academia, Manufacturing, Distribution, etc.