

```
# Define data and labels
texts = ["Bu bir örnek cümledir.", "Bu başka bir örnek cümledir.", ...] #
Your texts here
labels = [0, 1, ...] # Your labels here

# Encode data and labels
inputs = tokenizer(texts, padding=True, truncation=True,
return_tensors="pt")
labels = torch.tensor(labels)

# Create data loader
batch_size = 32
data_loader = DataLoader(iter_instances(inputs["input_ids"],
inputs["attention_mask"], labels)), batch_size=batch_size)

# Define model, loss function, optimizer, scheduler
model = ...
loss_fn = ...
optimizer = AdamW(model.parameters(), lr=2e-5)
total_steps = len(data_loader) * epochs
scheduler = get_linear_schedule_with_warmup(optimizer, num_warmup_steps=0,
num_training_steps=total_steps)

# Define training loop
epochs = 4
for epoch in range(epochs):
    # Train model on batches of data
    for batch in data_loader:
        # Get batch input and labels
        input_ids, attention_mask, labels = batch

        # Forward pass
        outputs = model(input_ids=input_ids, attention_mask=attention_mask)
        logits = outputs[0]

        # Compute loss
        loss = loss_fn(logits, labels)

        # Backward pass and update parameters
        loss.backward()
        optimizer.step()
        scheduler.step()

    # Reset gradients
    optimizer.zero_grad()
```

e-ISSN: 2148-7456

a peer-reviewed
online journal

hosted by DergiPark

International Journal of Assessment Tools in Education

Volume: 10

Issue: Special Issue

December 2023

<https://dergipark.org.tr/en/pub/ijate>



e-ISSN: 2148-7456

Volume: 10

Issue: Special Issue

2023

Special Issue

Editor

Dr. Okan BULUT

Address

University of Alberta, Edmonton, Canada

E-mail

bulut@ualberta.ca

ijate.editor@gmail.com

Journal Contact

Address

Dr. Eren Can AYBEK

Department of Educational Sciences, Pamukkale University, Faculty of Education, Kinikli Yerleskesi, Denizli, 20070, Türkiye

Phone

+90 258 296 31050

Fax

+90 258 296 1200

E-mail

erencanaybek@gmail.com

Address

Dr. Anil KANDEMİR

Department of Educational Sciences, Agri Ibrahim Cecen University, Faculty of Education, Agri, Türkiye

akandemir@agri.edu.tr

Publisher

Address

Dr. Izzet KARA

Pamukkale University, Education Faculty, Kinikli Campus, 20070 Denizli, Türkiye

Phone

+90 258 296 1036

Fax

+90 258 296 1200

E-mail

ikara@pau.edu.tr

ijate.editor@gmail.com

Frequency

4 issues per year (March, June, September, December)

Online ISSN

2148-7456

Website

<https://dergipark.org.tr/en/pub/ijate>

<https://ijate.net/index.php/ijate>

Cover Design

Merve SENTURK

International Journal of Assessment Tools in Education (IJATE) is a peer-reviewed and academic online journal. The scientific and legal responsibility for manuscripts published in our journal belongs to the authors(s).

International Journal of Assessment Tools in Education

International Journal of Assessment Tools in Education (IJATE) accepts original research on the design, analysis and use of evaluation along with assessment to enhance comprehension of the performance and quality of stakeholders in educational settings. IJATE is pleased to receive discriminating theoretical and empirical manuscripts (quantitative or qualitative) which could direct significant national and international argumentations in educational policy and practice.

IJATE, as an online journal, is hosted by DergiPark [TUBITAK-ULAKBIM (The Scientific and Technological Research Council of Türkiye)].

In IJATE, there are no charges under any procedure for submitting or publishing an article.

Indexes and Platforms:

- Emerging Sources Citation Index (ESCI)
- Education Resources Information Center (ERIC)
- TR Index (ULAKBIM),
- EBSCOhost,
- SOBIAD,
- JournalTOCs,
- MIAR (Information Matrix for Analysis of the Journals),
- idealonline,
- CrossRef,
- ResearchBib,
- Index Copernicus International

Editor

Dr. Okan BULUT, *University of Alberta, Canada*

Section Editors

Dr. Safiye BILICAN DEMIR, *Kocaeli University, Türkiye*

Dr. Selma SENEL, *Balikesir University, Türkiye*

Dr. Esin YILMAZ KOGAR, *Nigde Omer Halisdemir University, Türkiye*

Dr. Sumeyra SOYSAL, *Necmettin Erbakan University, Türkiye*

Editorial Board

Dr. Beyza AKSU DUNYA, *Bartın University, Türkiye*

Dr. Stanislav AVSEC, *University of Ljubljana, Slovenia*

Dr. Kelly D. BRADLEY, *University of Kentucky, United States*

Dr. Javier Fombona CADAVIECO, *University of Oviedo, Spain*

Dr. Seockhoon CHUNG, *University of Ulsan, Korea*

Dr. R. Nukhet CIKRIKCI, *İstanbul Aydın University, Türkiye*

Dr. William W. COBERN, *Western Michigan University, United States*

Dr. Nuri DOGAN, *Hacettepe University, Türkiye*

Dr. Selahattin GELBAL, *Hacettepe University, Türkiye*

Dr. Anne Corinne HUGGINS-MANLEY, *University of Florida, United States*

Dr. Francisco Andres JIMENEZ, *Shadow Health, Inc., United States*

Dr. Nicole KAMINSKI-OZTURK, *The University of Illinois at Chicago, United States*

Dr. Tugba KARADAVUT, *Izmir Democracy University, Türkiye*

Dr. Orhan KARAMUSTAFAOGLU, *Amasya University, Türkiye*

Dr. Yasemin KAYA, *Atatürk University, Türkiye*

Dr. Hulya KELECIOGLU, *Hacettepe University, Türkiye*

Dr. Hakan KOGAR, *Akdeniz University, Türkiye*

Dr. Seongyong LEE, *BNU-HKBU United International College, China*

Dr. Sunbok LEE, *University of Houston, United States*

Dr. Froilan D. MOBO, *Ama University, Philippines*

Dr. Hamzeh MORADI, *Sun Yat-sen University, China*

Dr. Nesrin OZTURK, *Izmir Democracy University, Türkiye*

Dr. Turan PAKER, *Pamukkale University, Türkiye*

Dr. Murat Dogan SAHIN, *Anadolu University, Türkiye*

Dr. Hossein SALARIAN, *University of Tehran, Iran*

Dr. Halil İbrahim SARI, *Kilis 7 Aralık University, Türkiye*

Dr. Ragıp TERZI, *Harran University, Türkiye*

Dr. Turgut TURKDOGAN, *Pamukkale University, Türkiye*

Dr. Ozen YILDIRIM, *Pamukkale University, Türkiye*

English Language Editors

Dr. R. Sahin ARSLAN, *Pamukkale University, Türkiye*

Dr. Hatice ALTUN, *Pamukkale University, Türkiye*

Ahmet KUTUK, *Akdeniz University, Türkiye*

Editorial Assistants

Dr. Asiye BAHTIYAR, *Pamukkale University, Türkiye*

Dr. Ebru BALTA, *Agri Ibrahim Cecen University, Türkiye*

Dr. Neslihan Tugce OZYETER, *Kocaeli University, Türkiye*

PhDc. Ibrahim Hakki TEZCI, *Akdeniz University, Türkiye*

Technical Assistant

Dr. Eren Can AYBEK, *Pamukkale University, Türkiye*

Dr. Anil KANDEMİR, *Agri Ibrahim Cecen University, Türkiye*

From Special Issue Editor:

Dear Readers,

In the ever-evolving realm of education, the landscape of measurement and evaluation has undergone a profound metamorphosis, primarily driven by the relentless advance of digital assessment tools. These tools, ranging from computer-based tests to automated essay scoring systems, have not only revolutionized how we gauge learning outcomes but have also opened new avenues for educators to tailor their approaches to the unique needs of each learner. As we stand at the nexus of technological innovation and educational practice, it is imperative to reflect on the past, acknowledge the present, and envision the future of educational assessment.

This special issue on “*Educational Measurement and Evaluation: Lessons from the Past, Visions for the Future*,” presented by the International Journal of Assessment Tools in Education (IJATE), marks a significant milestone—the 100th anniversary of the Republic of Türkiye. In celebrating this historic occasion, we have curated a collection of ten articles that delve into various facets of educational measurement and evaluation. Each article serves as a testament to the remarkable journey we have undertaken in understanding and enhancing the educational assessment landscape. For this special issue, we specifically asked researchers to discuss the evolution of educational measurement and evaluation concepts and their vision for the future of these concepts. All articles submitted for publication went through a rigorous peer review based on the review standards established by IJATE.

Within the inaugural article of this issue, Zumbo delivers an outstanding exploration of test validity, imparting his insights on methodological innovations in explanation-focused validity. This paper is poised to be a cornerstone resource, equipping both researchers and practitioners with a nuanced understanding of test validity spanning historical roots to contemporary perspectives. Concurrently, the sixth article, penned by Mor and Karatoprak Ersen in this issue, delves into the realm of test validity. The authors meticulously examine the ramifications of prevailing validity frameworks within the context of classroom assessment, enriching the discourse on the crucial subject of test validity.

The second article by Arici and Kutlu (in this issue) focuses on factors related to collaborative problem-solving skills. Leveraging the outcomes of PISA 2015, the authors meticulously scrutinize both direct and indirect factors influencing the collaborative problem-solving aptitude of students in Türkiye. Simultaneously, the eighth article, contributed by He in this issue, also centers on collaborative problem-solving skills within large-scale assessments. The author evaluates the intricacies of item design and scoring and engages in a thoughtful discussion on potential approaches to gauge students’ proficiency more accurately in collaborative problem-solving. Collectively, these articles contribute to the evolving discourse surrounding collaborative problem-solving, offering valuable insights for educators, researchers, and policymakers.

In the third article, Schwarz et al. (in this issue) describe a data pipeline for digital large-scale assessments (the authors refer to this as “e-large-scale assessments”). Employing the versatile R programming language, the authors skillfully showcase the automation of various data analysis steps. These include data transformation, psychometric analyses grounded in Classical Test Theory and Item Response Theory, and the streamlined generation of score reports. The article provides a comprehensive demonstration of how a meticulously designed data pipeline can automate these crucial processes, offering an insightful guide for practitioners and researchers engaged in large-scale assessments.

The fourth and fifth articles center on the automation of two pivotal psychometric tasks—item development and scoring—through the application of advanced computer algorithms. In their contribution to this issue, Sayin et al. showcase the efficacy of template-based automatic item generation in crafting non-verbal items for a visual reasoning test. The study's findings underscore the potential of automatic item generation in expeditiously building a substantial

repository of items. Similarly, Firoozi et al. delve into automated essay scoring in the fifth article, elucidating the application of large-language models. The authors intricately describe and illustrate the framework of automated essay scoring systems within the specific context of the Turkish language. Collectively, these articles illuminate the transformative impact of advanced algorithms in revolutionizing key psychometric processes, offering valuable insights into the future of assessment methodologies.

Taskin Bedizel (in this issue) unfolds the outcomes of a comprehensive bibliometric analysis, spanning publications from 1994 to September 2023, focused on the intersection of artificial intelligence and educational assessment research. The insights derived from this study not only offer a panoramic view of the evolution over time but also provide valuable guidance for researchers and practitioners navigating the trajectory of AI-powered assessment tools in education.

The seventh and tenth articles delve into the analysis of process data derived from international large-scale assessments. In this issue, Yilmaz Kogar and Soysal leverage item response time to investigate the impact of factors like item difficulty, content, and cognitive domain on problem-solving duration in TIMSS 2019. The authors draw upon samples from 4th-grade students in Singapore and Turkey, conducting a comparative analysis of aberrant response behaviors, including rapid guessing, between the two nations. In a parallel exploration, Ersan and Parlak utilize TIMSS 2019 data to scrutinize the influence of on-screen calculators on students' performance in the Problem Solving and Inquiry tasks. Their findings reveal a positive association between the use of on-screen calculators and the likelihood of correct item responses, highlighting the value of process data in understanding students' response behaviors in digital assessments.

I hope that the articles featured in this special issue will inspire further dialogue, spark new ideas, and contribute to the ongoing evolution of educational assessment as we embark on the next century of progress. As we navigate the complexities of the digital age, the featured articles will offer insights into the ways technology has reshaped our approaches to educational assessment. From exploring validity frameworks that underpin assessment methodologies to investigating the intricacies of automated essay scoring, automatic item generation, collaborative problem-solving, and digital large-scale assessments, these contributions encapsulate the breadth of advancements that have shaped the past and continue to mold the future of educational measurement. The lessons gleaned from the past and the visions articulated for the future converge in this special issue, offering a comprehensive exploration of the multifaceted world of educational measurement and evaluation.

In drawing this editorial summary to a close, I wish to express my heartfelt gratitude to the esteemed authors whose expertise and dedication have profoundly enriched this special issue. A special acknowledgment extends to all the diligent reviewers who generously shared their valuable insights, contributing significantly to the refinement of the submitted papers. Additionally, my sincere appreciation goes to the editors of IJATE, Dr. Omer Kutlu and Dr. Izzet Kara, for entrusting me with the privilege of curating this special issue. The realization of this endeavor would not have been possible without the unwavering support and meticulous efforts of the IJATE Editorial Team. Their commitment has been instrumental in bringing this special issue to fruition.

Assoc Prof. Okan Bulut

Measurement, Evaluation, and Data Science

Faculty of Education, University of Alberta
6-110 Education Centre North, 11210 87 Ave NW,
Edmonton, AB T6G 2G5 CANADA

E-mail: bulut@ualberta.ca

CONTENTS

Invited Article

[A dialectic on validity: Explanation-focused and the many ways of being human](#)

Page: 1-96 [PDF](#)

Bruno D. ZUMBO

Research Articles

[An investigation of factors related to collaborative problem-solving skills with mediation models](#)

Page: 97-115 [PDF](#)

Ozge ARICI, Omer KUTLU

[A data pipeline for e-large-scale assessments: Better automation, quality assurance, and efficiency](#)

Page: 116-131 [PDF](#)

Ryan SCHWARZ, Hatice Cigdem BULUT, Charles ANIFOWOSE

[Automatic item generation for non-verbal reasoning items](#)

Page: 132-148 [PDF](#)

Ayfer SAYIN, Sabiha BOZDAG, Mark J. GIERL

[Language models in automated essay scoring: Insights for the Turkish language](#)

Page: 149-163 [PDF](#)

Tahereh FIROOZI, Okan BULUT, Mark GIERL

[Implications of current validity frameworks for classroom assessment](#)

Page: 164-173 [PDF](#)

Ezgi MOR, Rabia KARATOPRAK ERSEN

[Examination of response time effort in TIMSS 2019: Comparison of Singapore and Türkiye](#)

Page: 174-193 [PDF](#)

Esin YILMAZ KOGAR, Sumeyra SOYSAL

[Collaborative problem-solving design in large-scale assessments: Shedding lights in sequential conversation-based measurement](#)

Page: 194-207 [PDF](#)

Qiwei HE

[Evolving landscape of artificial intelligence \(AI\) and assessment in education: A bibliometric analysis](#)

Page: 208-223 [PDF](#)

Nazli Ruya TASKIN BEDIZEL

[The use of on-screen calculator as a digital tool in technology-enhanced items](#)

Page: 224-242 [PDF](#)

Ozge ERSAN, Burcu PARLAK
