

ENTREPRENEUR & MANAGEMENT RESEARCH

Instrument Development: Validity and Reliability

Present By
Dr. Tumpal P. Situmorang, ST.,MM



REVIEW OF MEETING 8

DATA COLLECTION IN ENTREPRENEURIAL CONTEXTS

01

Data collection is where theory meets reality (e.g., entrepreneurial resilience only becomes researchable when it is connected to actual experiences, actions, or decisions).

02

Different methods produce different kinds of evidence (e.g., interviews reveal interpretation, while platform data reveal recorded behavior).

03

Entrepreneurial phenomena are often dynamic and processual (e.g., founders may change strategy or reinterpret failure during the study).

04

Good data collection depends on fit between the question and the source of evidence (e.g., crowdfunding legitimacy may be better studied through campaign texts than through memory-based recall alone).

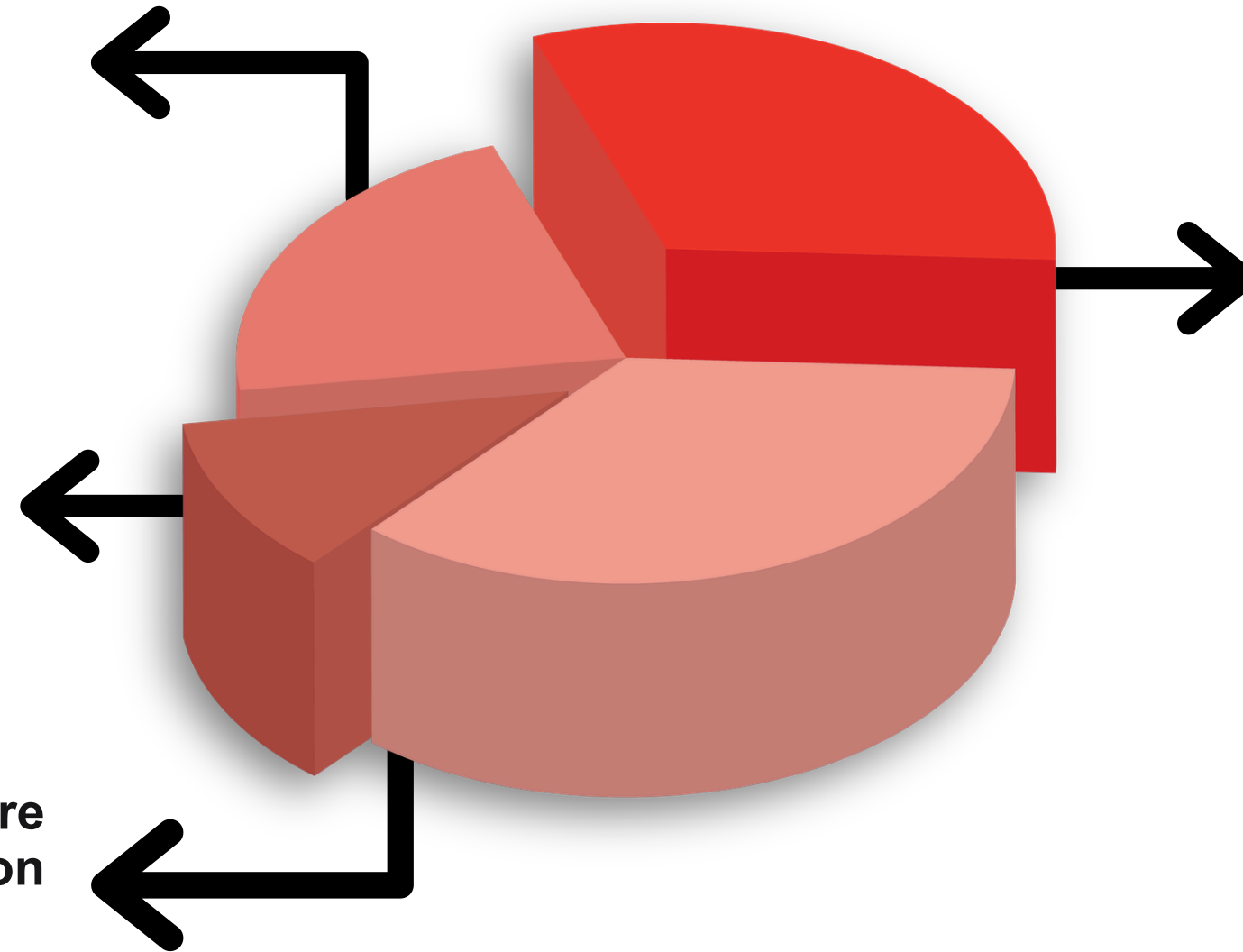
INTRODUCTION

Why Instrument Development Matters

A research instrument connects abstract constructs to measurable evidence (e.g., founder social identity must be translated into observable indicators before it can be analyzed).

Poor instruments weaken strong theories (e.g., a vague scale of customer value will produce unclear findings even with a large sample).

Instrument development is conceptual before it is statistical (e.g., item wording depends on how the construct is theoretically defined).



Validity and reliability matter because measurement is the foundation of inference (e.g., inconsistent or invalid measures cannot support strong conclusions).

WHAT IS A RESEARCH INSTRUMENT?

From Abstract Idea to Empirical Tool

- ▶ An instrument is a structured way of capturing a construct (e.g., a founder's social identity can be captured through items that reflect role centrality and self-definition).
- ▶ Instruments may take the form of scales, checklists, indexes, or coded protocols (e.g., an empowerment construct can be operationalized through multidimensional scale items).

source: <https://share.google/sSeYiLLWMBTeMohRx>



source: <https://share.google/ow7Xoi9hi2kL06adH>

source: <https://share.google/9G3p0RqsItCAaxx7R>



source: <https://share.google/p3YCOOV6Ek2f0yYhG>

WHAT IS A RESEARCH INSTRUMENT?

From Abstract Idea to Empirical Tool

- A good instrument does not merely ask questions; it operationalizes theory (e.g., hustle is not measured by generic “hard work” items, but by theory-based urgent and unorthodox action indicators).
- An instrument must be interpretable, not merely answerable (e.g., respondents may answer an item, but the question is whether the answer truly represents the construct).

source: <https://share.google/sSeYiLLWMBTeMohRx>



source: <https://share.google/ow7Xoi9hi2kL06adH>

source: <https://share.google/9G3p0RqsItCAaxx7R>



source: <https://share.google/p3YCOOV6Ek2f0yYhG>



FROM CONCEPT TO INDICATOR

THE FIRST DISCIPLINE OF MEASUREMENT

- 01** Instrument development starts with conceptual clarity (e.g., firm engagement must be clearly distinguished from customer engagement and employee engagement).
- 02** A construct becomes measurable only when its dimensions are articulated (e.g., masstige marketing had to be decomposed into elements such as prestige, self-congruence, and willingness to pay a premium).



FROM CONCEPT TO INDICATOR

THE FIRST DISCIPLINE OF MEASUREMENT

- 03** Indicators should emerge from theory and literature, not guesswork (e.g., digital orientation requires items grounded in the concept of a firm-wide digital configuration).
- 04** The sharper the concept, the sharper the instrument (e.g., vague constructs generate vague items and unstable findings).

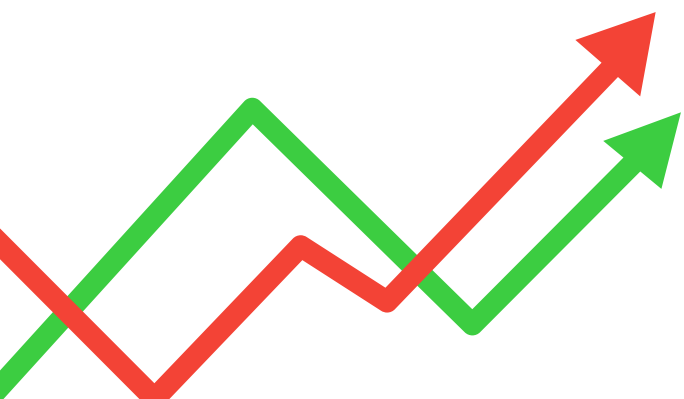


CONTENT VALIDITY



DO THE ITEMS ADEQUATELY REPRESENT THE CONSTRUCT?

01 CONTENT VALIDITY ASKS WHETHER THE INSTRUMENT COVERS THE SUBSTANCE OF THE CONSTRUCT (E.G., ENTREPRENEURIAL HUSTLE SHOULD NOT BE REDUCED TO GENERIC BUSYNESS OR LONG WORKING HOURS).



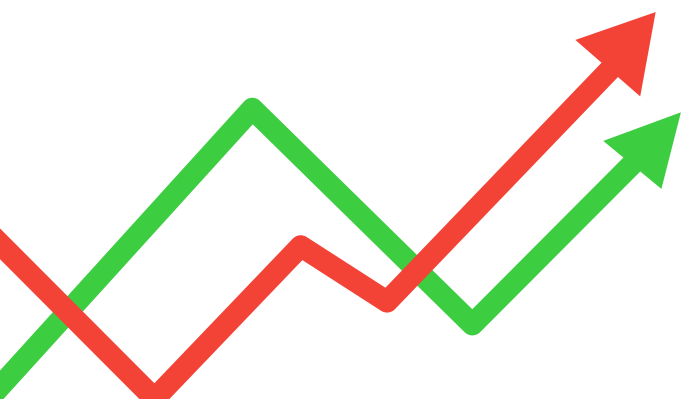


CONTENT VALIDITY



DO THE ITEMS ADEQUATELY REPRESENT THE CONSTRUCT?

02 IT DEPENDS ON THEORY, LITERATURE REVIEW, AND EXPERT JUDGMENT (E.G., ITEM POOLS ARE SCREENED TO ENSURE THEY REFLECT THE INTENDED CONCEPTUAL DOMAIN).





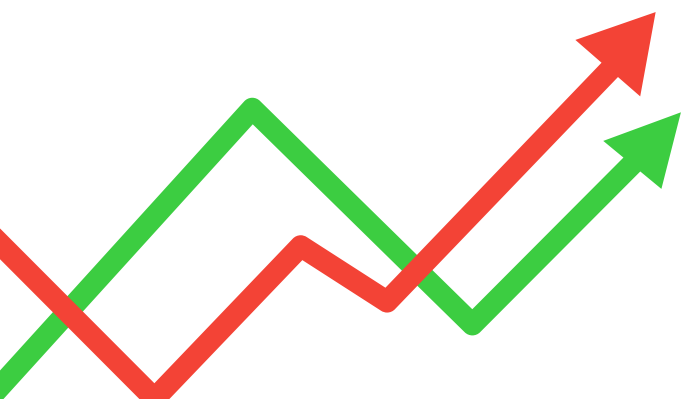
CONTENT VALIDITY



DO THE ITEMS ADEQUATELY REPRESENT THE CONSTRUCT?

03

GOOD CONTENT VALIDITY BEGINS BEFORE ANY STATISTICAL TEST (E.G., WEAK ITEMS CANNOT BE FULLY REPAIRED LATER BY IMPRESSIVE FACTOR LOADINGS).



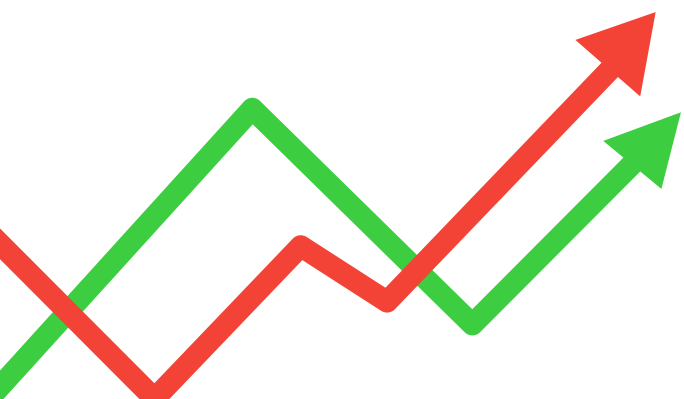


CONTENT VALIDITY



DO THE ITEMS ADEQUATELY REPRESENT THE CONSTRUCT?

04 IT IS ESPECIALLY CRUCIAL FOR NEW OR EMERGING CONSTRUCTS (E.G., FIRM ENGAGEMENT AND DIGITAL ORIENTATION REQUIRE CAREFUL CONCEPTUAL COVERAGE BECAUSE THEIR BOUNDARIES ARE STILL EVOLVING).





CONSTRUCT VALIDITY

**DOES THE INSTRUMENT REALLY MEASURE
THE INTENDED CONSTRUCT?**



**CONSTRUCT VALIDITY ASKS WHETHER THE
INSTRUMENT TRULY MEASURES THE TARGET
CONCEPT (E.G., SOCIAL IDENTITY OF
ENTREPRENEURS SHOULD CAPTURE SELF-
DEFINITION, NOT JUST CONFIDENCE OR
AMBITION).**



CONSTRUCT VALIDITY



**DOES THE INSTRUMENT REALLY MEASURE
THE INTENDED CONSTRUCT?**



**IT USUALLY INVOLVES DIMENSIONAL STRUCTURE
AND EMPIRICAL COHERENCE (E.G., THE
PROPOSED DIMENSIONS SHOULD BE SUPPORTED
BY THE DATA RATHER THAN REMAIN
THEORETICAL ASSUMPTIONS).**



CONSTRUCT VALIDITY

**DOES THE INSTRUMENT REALLY MEASURE
THE INTENDED CONSTRUCT?**



**CONSTRUCT VALIDITY INCLUDES CONVERGENT
AND DISCRIMINANT LOGIC (E.G., RELATED
CONSTRUCTS MAY CORRELATE, BUT DISTINCT
CONSTRUCTS MUST STILL REMAIN
DISTINGUISHABLE).**



CONSTRUCT VALIDITY

**DOES THE INSTRUMENT REALLY MEASURE
THE INTENDED CONSTRUCT?**



**WITHOUT CONSTRUCT VALIDITY,
INTERPRETATION BECOMES UNSTABLE (E.G., A
HIGH SCORE MAY NOT MEAN WHAT THE
RESEARCHER THINKS IT MEANS).**

RELIABILITY

Source: Canva



CAN THE INSTRUMENT PRODUCE CONSISTENT RESULTS?

01

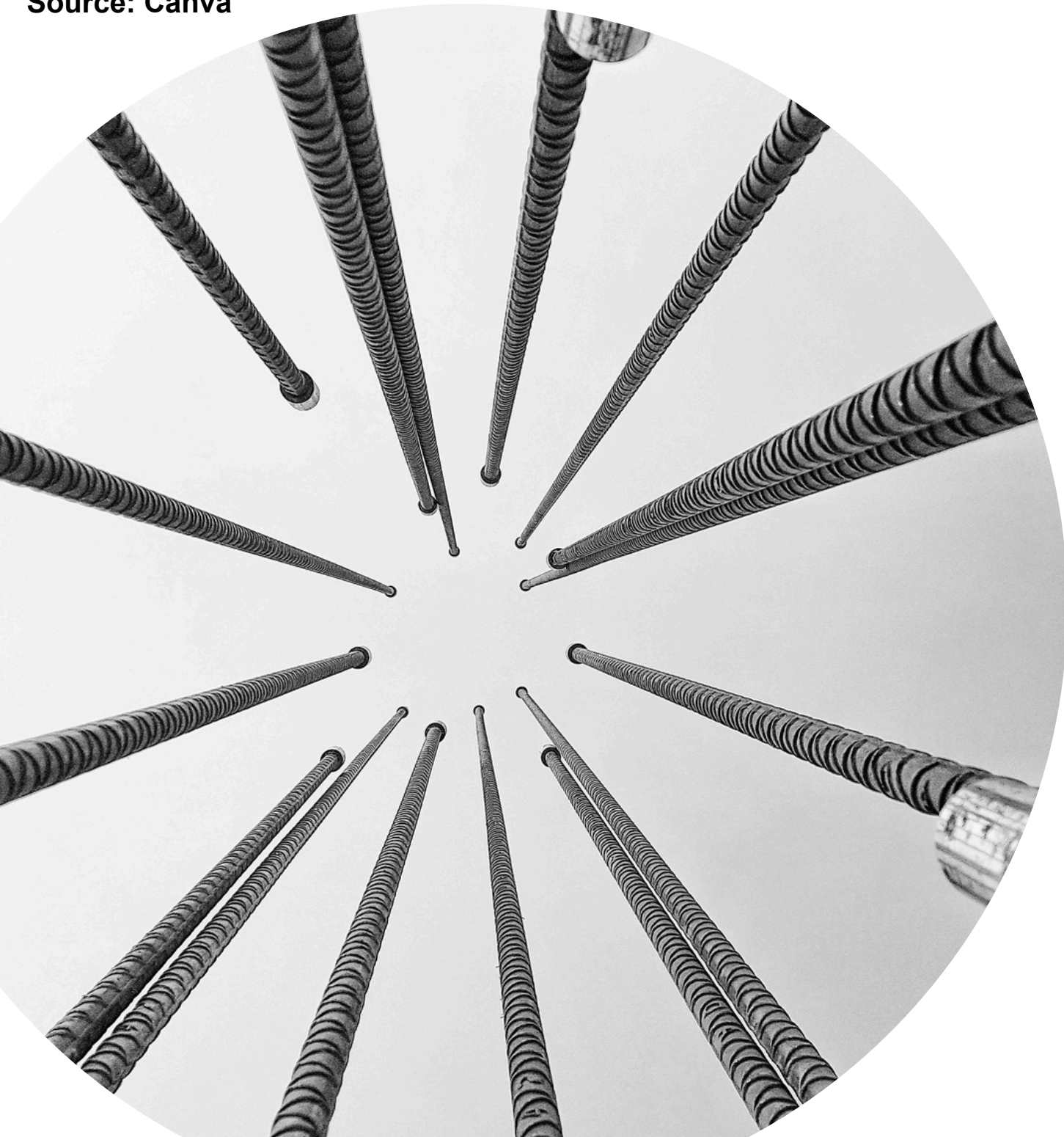
Reliability concerns the consistency of measurement (e.g., items measuring work-life fulfilment should behave coherently as one scale system).

02

Internal consistency is important, but it is not the whole story (e.g., a scale may hang together internally but still shift when the construct itself is unstable over time).

RELIABILITY

Source: Canva



CAN THE INSTRUMENT PRODUCE CONSISTENT RESULTS?

03

A reliable instrument is not automatically valid (e.g., consistent items can still measure the wrong thing).

04

Reliability gives confidence in the score, but validity gives meaning to the score (e.g., reliability makes measurement dependable; validity makes it interpretable).

PILOT TESTING AND SCALE PURIFICATION

WHY GOOD INSTRUMENTS ARE REFINED, NOT RUSHED

- 01** Pilot testing helps detect ambiguity, confusion, and redundancy (e.g., respondents may interpret “firm engagement” very differently if the item wording is vague).
- 02** Scale purification removes weak, overlapping, or misleading items (e.g., items that blur prestige and loyalty should not remain in a prestige scale).



source: <https://share.google/OgNp52sR38PNBB0Ev>

PILOT TESTING AND SCALE PURIFICATION

WHY GOOD INSTRUMENTS ARE REFINED, NOT RUSHED

03 Refinement improves both validity and reliability (e.g., the final instrument becomes cleaner, sharper, and easier to interpret).

04 A polished scale usually emerges through multiple rounds, not one draft (e.g., item generation, expert review, testing, purification, and validation).



source: <https://share.google/OgNp52sR38PNBB0Ev>

ENTREPRENEURSHIP EXAMPLE



MEASURING THE SOCIAL IDENTITY OF ENTREPRENEURS

01

Entrepreneurial identity is not a vague personality trait (e.g., founders differ in how central entrepreneurial roles are to their self-definition).

ENTREPRENEURSHIP EXAMPLE



MEASURING THE SOCIAL IDENTITY OF ENTREPRENEURS

02

A scale can deepen theory by clarifying what is being measured (e.g., social identity helps explain why founders pursue different venture paths).

ENTREPRENEURSHIP EXAMPLE



MEASURING THE SOCIAL IDENTITY OF ENTREPRENEURS

03

International validation strengthens confidence in the construct (e.g., testing across countries reduces the risk that the scale reflects only one local context).

ENTREPRENEURSHIP EXAMPLE



MEASURING THE SOCIAL IDENTITY OF ENTREPRENEURS

04

Good measurement makes entrepreneurial heterogeneity visible (e.g., founders are not all driven by the same self-concept).

MARKETING EXAMPLE

Source: Canva



01

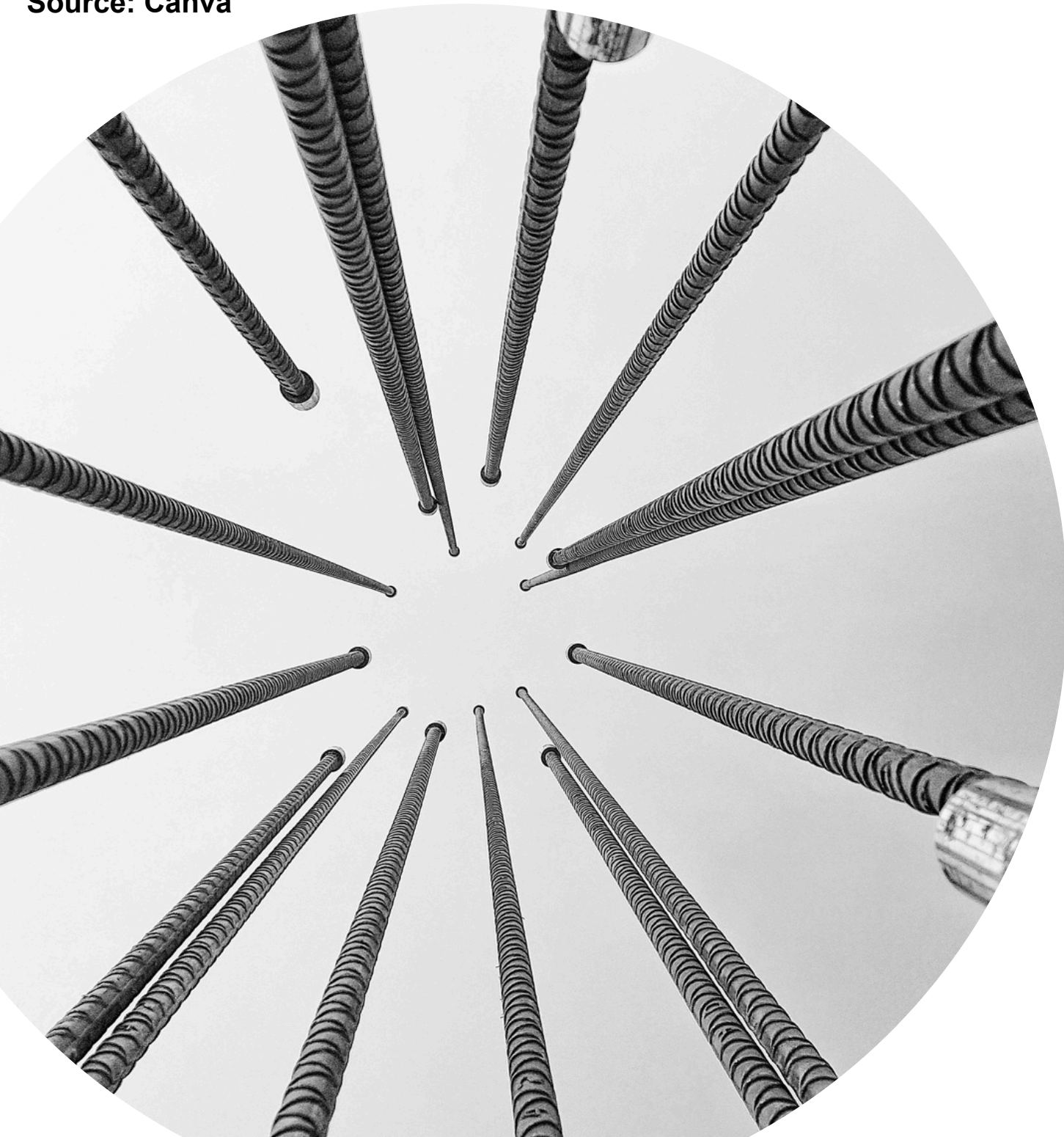
Popular concepts often need sharper measurement, not more repetition (e.g., masstige marketing is widely discussed, but cannot remain only a fashionable label).

02

A strong scale forces a concept to become analytically precise (e.g., prestige, self-congruence, and willingness to pay premium must be distinguished rather than blended together).

MARKETING EXAMPLE

Source: Canva



03

Validity matters because managers may act on the results (e.g., a brand strategy based on a flawed scale may misread consumers' masstige perceptions).

04

Generalizability is stronger when the instrument is tested across industries (e.g., restaurant, airline, and automobile contexts reveal whether the scale travels well).

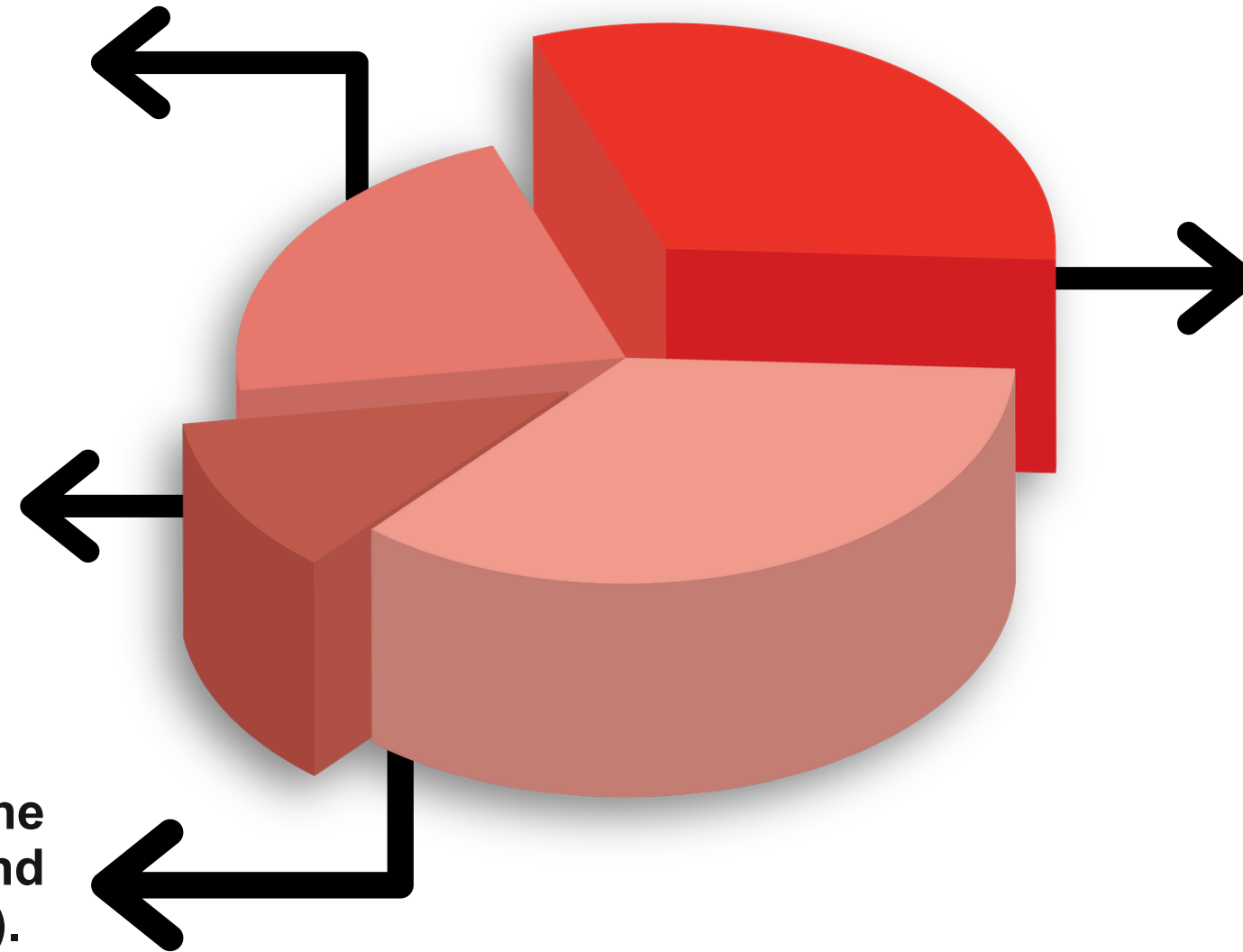
DIGITAL / SERVICE / HRM EXAMPLE

Measuring Frontline Employee Empowerment in Analytics-Driven Services

Old constructs may need new instruments in new settings (e.g., empowerment in analytics-driven services is not identical to empowerment in traditional service environments).

Scale development must follow contextual change (e.g., digital tools, information access, and analytics reshape what empowerment means for frontline employees).

Mixed evidence-building can strengthen the scale (e.g., literature review, interviews, and survey-based validation may work together).



Measurement quality matters because service performance may depend on it (e.g., mismeasuring empowerment may lead firms to redesign jobs in the wrong direction).

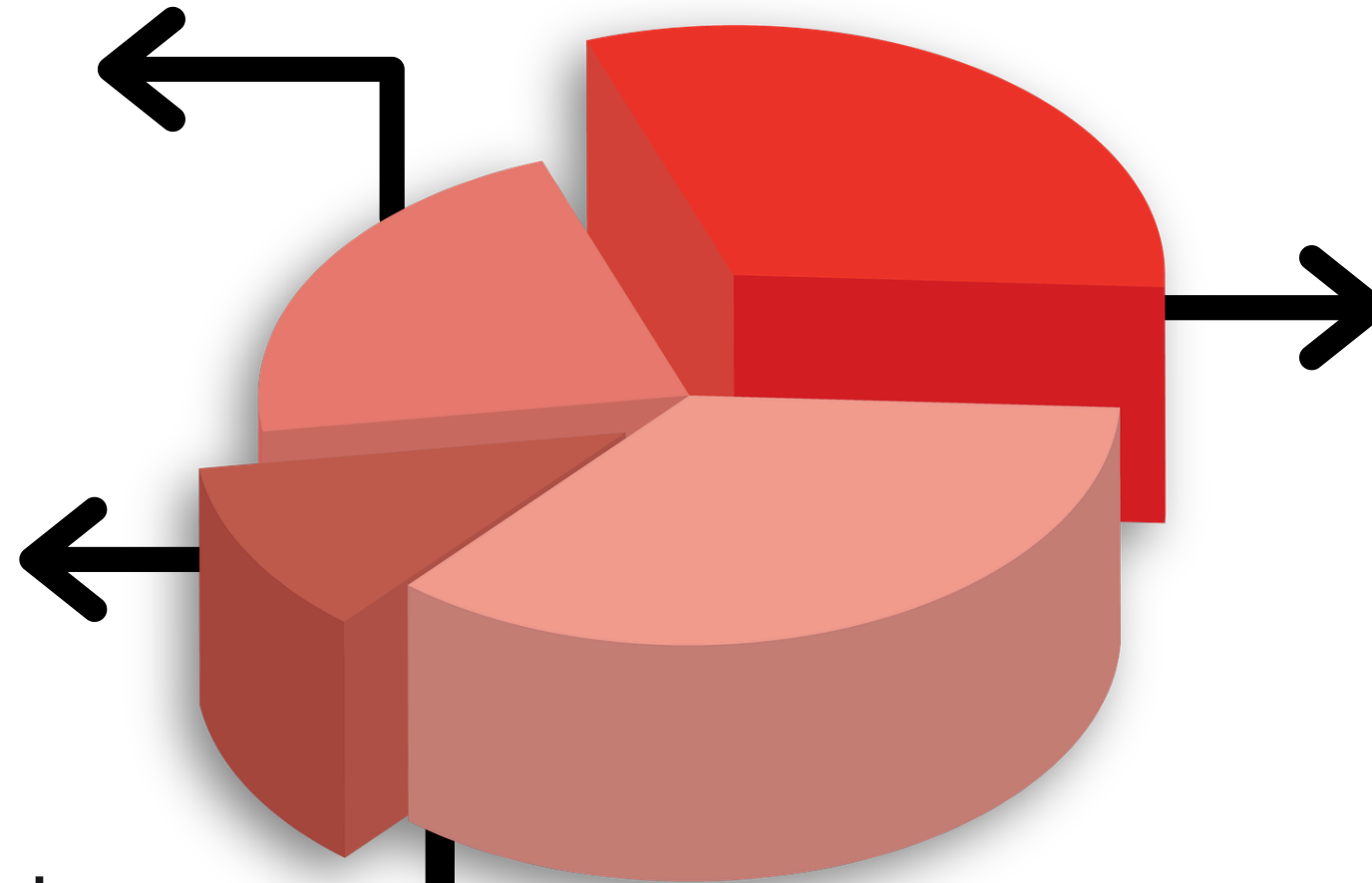
COMMON MISTAKES IN INSTRUMENT DEVELOPMENT

Why Many Studies Look Strong but Measure Weakly

Many researchers jump too quickly from variable names to questionnaire items (e.g., writing “digital mindset” items before defining what the construct actually means).

Popularity of a construct does not guarantee quality of its measurement (e.g., engagement, resilience, and empowerment can all be overused but poorly operationalized).

Researchers often overtrust statistics and underinvest in conceptual clarity (e.g., a high loading does not automatically prove that an item is substantively correct).



Weak measurement undermines the credibility of the whole model (e.g., elegant hypotheses collapse when the constructs are poorly captured).

CONCLUSION!

01

Instrument development is where conceptual rigor becomes measurement rigor (e.g., abstract constructs must be turned into defensible empirical indicators).

02

Validity asks whether the instrument means what it claims to mean (e.g., whether a scale really captures identity, engagement, or empowerment rather than something adjacent).

CONCLUSION!

03

Reliability asks whether the instrument is consistent enough to be trusted (e.g., whether the score can be treated as dependable rather than erratic).

04

Strong research requires instruments that are both theoretically grounded and empirically defensible (e.g., a weak instrument can damage even a strong model).

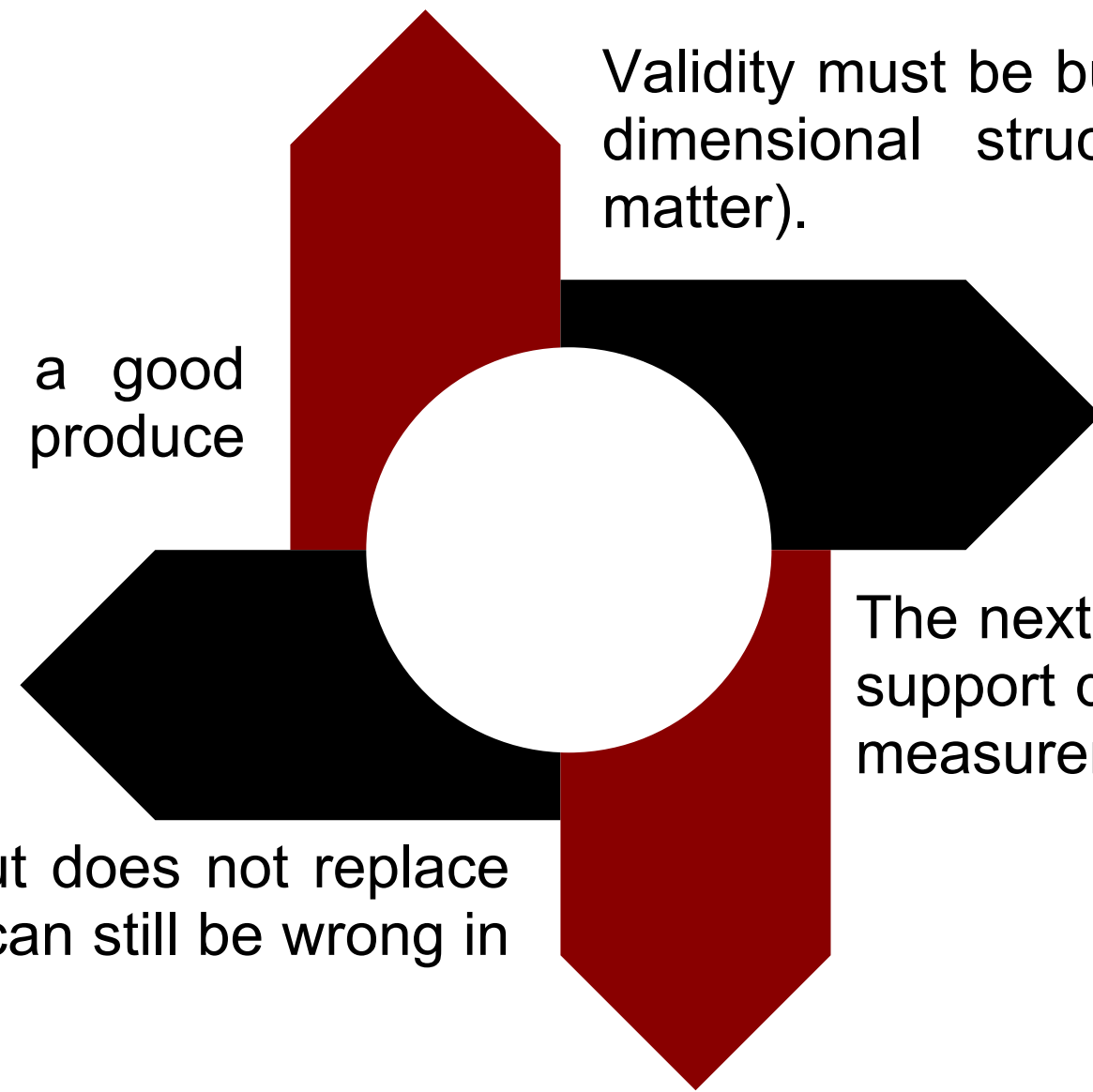
KEY TAKEAWAYS AND CLOSING REMARKS

A good instrument begins with a good concept (e.g., unclear constructs produce unclear items).

Reliability supports confidence, but does not replace meaning (e.g., a consistent scale can still be wrong in substance).

Validity must be built, not assumed (e.g., item relevance, dimensional structure, and construct boundaries all matter).

The next step is to learn how trustworthy instruments support quantitative analysis (e.g., moving from good measurement to sound statistical interpretation).





See you in Meeting
10, with the topic:

**Descriptive Statistics
and Interpretation**



REFERENCE

Research Methods for Business Students, Mark N. K. Saunders, Philip Lewis, Adrian Thornhill, Pearson, 2024, 888 pages.

Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, John W. Creswell, J. David Creswell, SAGE Publications, 2018, 304 pages.

Research Methods For Business: A Skill Building Approach, Roger Bougie, Uma Sekaran, John Wiley & Sons, 2020, 432 pages.

Business Research Methods, Emma Bell, Bill Harley, Alan Bryman, Oxford University Press, 2022, 647 pages.

Measuring the Social Identity of Entrepreneurs: Scale Development and International Validation, Melissa S. Cardon, Joakim Wincent, Jagdip Singh, Mateja Drnovšek, Journal of Business Venturing, Elsevier, 2016, 542-572.

Entrepreneurial Fear of Failure: Scale Development and Validation, Gabriella Cacciotti, James C. Hayton, J. Robert Mitchell, David G. Allen, Journal of Business Venturing, Elsevier, 2020, 106041.

Entrepreneurial Hustle: Scale Development and Validation, Devin Burnell, Emily Neubert, Greg Fisher, Matthew R. Marvel, Regan Stevenson, Donald F. Kuratko, Journal of Business Venturing, Elsevier, 2024, 106407.

Masstige Marketing: A Scale Development and Validity Study, Muhammad Ishtiaq Ishaq, Ali Raza, Boris Bartikowski, Huma Sarwar, Journal of Business Research, Elsevier, 2023, 114112.

Frontline Employee Empowerment: Scale Development and Validation Using Confirmatory Composite Analysis, Saradhi Motamarri, Shahriar Akter, Venkata Yanamandram, International Journal of Information Management, Elsevier, 2020, 102177.

Firm Engagement: Scale Development and Verification, Xiaoyun Han, Yundan Zheng, Catherine Prentice, Shuping Chen, Journal of Retailing and Consumer Services, Elsevier, 2024, 103902.

Work-Based Learning: Development and Validation of a Scale Measuring the Learning Potential of the Workplace (LPW), Irina Nikolova, Joris van Ruysseveldt, Hans De Witte, Jef Syroit, Journal of Vocational Behavior, Elsevier, 2014, 1-10.

Towards Work Life Fulfilment: Scale Development and Validation, Puja Khatri, Shalu Shukla, Asha Thomas, Atul Shiva, Abhishek Behl, Journal of Business Research, Elsevier, 2025, 115006.

Expanding the Boundaries of Digital Orientation Research: Scale Development and Validation, Bastian Kindermann, Corinna Vera Hedwig Schmidt, Florian Fengel, Steffen Strese, Journal of Business Research, Elsevier, 2024, 114895.

**THANK
YOU**