

A person wearing a white lab coat is seated at a wooden desk. They are holding a silver pen in their right hand and a clipboard with a white sheet of paper in their left hand. A smartphone is visible on the desk to the left of the person. The background is blurred, showing what appears to be a laboratory or office setting.

PSY443- PSYCHOLOGICAL TESTING AND ASSESSMENT STEPS IN SCALE DEVELOPMENT-2

STEPS IN SCALE DEVELOPMENT

1. **Determine the content of the scale**
2. **Generate an item pool- item writing**
3. **Determine a response format**
4. Have experts review the item pool
5. Pilot testing and item analysis
6. Applying the scale to the target population
7. Reliability and validity analysis
8. Optimize scale length (final version of the scale)

Last week...

- 1. Decide what you want to measure:** Define the main purpose of your research and identify the concept or construct you want to measure.
 - Critical for clarifying the direction and content of the scale.
- 2. Generate an item pool:** Write a set of items that represent the construct you aim to measure.
 - Items should cover all relevant dimensions or subcomponents.
- 3. Determine the response format:** Decide how participants will respond to the items in the scale.
 - This could include options like Likert-type scales, true/false answers, or open-ended formats...

Step 4: Have Experts Review the Item Pool

- The primary question to address is: *“To what extent are the items included in the survey sufficient for covering and collecting the required factual and/or judgmental data?”*
 - To answer this question, which is related to the content validity of the survey, experts are consulted.
- Ask people who are knowledgeable in the content area to review your initial item pool
 - Maximizes your **content validity! THAT'S IMPORTANT!!**
 - Confirms or invalidates your definition of the phenomenon
- Have them rate *how relevant they think each item is to what you intend to measure*
 - Especially important if you are creating a measure that will consist of separate scales to measure multiple constructs

Step 4: Have Experts Review the Item Pool

- This step parallels hypothesis testing
 - Hypothesis = Your thoughts about what each item measure
 - Data (confirming or disconfirming) = Your experts responses
- How to do it:
 - Give them a working definition of the construct
 - Ask them to rate the relevance of each item to the construct as you have defined it
 - Ask for comments on individual items (e.g., clarity, conciseness, alternative wordings)

Step 4: Have Experts Review the Item Pool

- Experts can also offer alternative ways to measure the construct of interest
- **Final decision to include or exclude items is your responsibility!!**
 - Experts may not understand principles of scale construction
 - Attend to their suggestions, but make your own informed decisions about how to actually use their advice

Step 4: Have Experts Review the Item Pool

- Expert form is prepared to evaluate the items in terms of **relevance** and **clarity**. This form is used to gather expert opinions to assess the content validity of the scale items. The expert form typically includes the following components:
 1. **For the experts**, a brief explanation of the general purpose of the scale and the construct intended to be measured.
 2. **Item List**: A list of all scale items presented for experts to evaluate.
 3. **Evaluation Criteria**: Experts are asked to rate each item based on the following criteria:
 1. **Relevance (İlgililik)**: How appropriate is the item in representing the construct being measured?
 2. **Clarity (Açıklık)**: How clear and understandable is the item to the respondents?

ROMANTİK İLİŞKİLERDE ADALETE DUYARLILIK ÖLÇEĞİ (MAĞDUR BOYUTU)

UZMAN DEĞERLENDİRME FORMU

Sayın Uzman,

Adalet duyarlılık, kişilerin kendilerine adil olmayan bir biçimde davranıldığında ya da adaletsiz durumlarla karşılaştıklarında gösterdikleri tepkilere ve öfke yoğunluklarına odaklanan bir kavram olarak tanımlanmaktadır. Gollwitzer ve Rothmund (2011)'e göre, bazı insanlar adaletsiz durumlarla karşılaştıklarında güçlü olumsuz duygular yaşarken, bazıları ise adaletsiz sonuçlara veya adil olmayan durumlara daha az önem vermektedir.

Adalet duyarlılığın altboyutlarından biri olan **mağdur duyarlılığı** ise bireyin adaletsiz durumlarda benliği ile ilgili endişeleri, anti-sosyal eğilimleri, tekrarlayıcı düşünceleri ve yoğun duygusal tepkileri şeklinde işlev görmektedir (Gollwitzer, Schmitt, Schalke, Maes ve Baer, 2005; Schmitt, Baumert, Gollwitzer ve Maes, 2010). Mağdur duyarlılığı yüksek olan kişiler başkalarıyla ilgili olan adaletten ziyade **kendileri ile ilgili olan adaletle ilgilidirler** (Gollwitzer ve ark., 2005; Schmitt, Gollwitzer, Maes ve Arbach, 2005). Adalet duyarlılığın mağdur boyutu, **sömürülme korkusu içerir ve kişinin kendisine karşı adil davranılma konusundaki kaygısını** yansıtır (Schmitt ve ark., 2005). Schmitt ve ark. (2005) yürüttükleri bir araştırmada, kişilerin **paranoya, nevroitiklik, şüphencilik, intikam ve kıskançlık gibi kendileriyle ilgili olan endişeleriyle mağdur duyarlılığı arasında yüksek düzeyde olumlu bir ilişki** olduğunu rapor etmişlerdir.

Bu çalışmada, romantik ilişkilerde adalet duyarlılığın altboyutlarından biri olan mağdur duyarlılığını değerlendirmeye yönelik geçerli ve güvenilir bir ölçme aracı geliştirilmesi amaçlanmaktadır. Ölçeğin hedef kitlesini **en az 1 yıldır evli olan yetişkinlerin** oluşturması planlanmaktadır.

1. Elinizdeki bu formda “romantik ilişkilerde adalet duyarlılığın altboyutlarından biri olan mağdur duyarlılığını” ölçmek amacıyla 5’li likert tipi şeklinde geliştirilmek istenen ölçek için düşünülen olası maddeler bulunmaktadır. Her bir maddenin yapı ile ne derece ilgili (**4: çok ilgili, 1: hiç ilgili değil**) olduğunu ve söz konusu maddenin uygulanmak istenen hedef kitle için ne derece anlaşılır/açık (**4: çok açık, 1: hiç açık değil**) olduğunu ilgili rakamı işaretleyerek değerlendirmeniz beklenmektedir.
2. Maddeler için kullanılması planlanan Likert tipi ölçeğin “**5:kesinlikle katılıyorum, 4:katılıyorum, 3:biraz katılıyorum, 2:katılmıyorum, 1:kesinlikle katılmıyorum**” olarak derecelendirilmesi planlanmaktadır. **Derecelendirme ile ilgili farklı önerileriniz var ise** sayfanın son kısmında bulunan “NOT” kısmında belirtebilirsiniz.
3. Yazılan maddeler ile ilgili varsa öneri ve düzenlemelerinizi her bir maddenin altında yer alan boşluklara belirtebilirsiniz.

Değerli zamanınızı ayırarak ölçeğin geliştirilmesine katkıda bulunduğunuz için teşekkür ederiz.

Dr. Öğr. Üyesi Sena DOĞRUYOL

MADDELER	İLGİLİLİK					AÇIKLIK			
	4	3	2	1		4	3	2	1
1. Partnerim/eşim tarafından sömürüldüğümü (<i>duygusal, maddi, manevi gibi</i>) hissettiğimde bunu kabul edemem.	4	3	2	1		4	3	2	1
2. Partnerimle/eşimle bir konuda tartıştığımızda haklı olmama rağmen suçlanırsam günlerce tartıştığımız konuyu düşünürüm.	4	3	2	1		4	3	2	1
3. Katıldığımız bir etkinlikte benim hak ettiğim ödülü partnerim/eşim aldığında ona öfkelenirim.	4	3	2	1		4	3	2	1
4. Kendimi geliştirmek için partnerimden/eşimden az imkana sahip olmak moralimi bozar.	4	3	2	1		4	3	2	1
5. Haklı olduğum halde ailelerimizin partnerimi/eşimi kayırmasını uzun süre sindiremem.	4	3	2	1		4	3	2	1
6. Partnerimin/eşimin dikkatsizliği yüzünden zarar gördüğümde ona öfkelenirim.	4	3	2	1		4	3	2	1
7. Partnerim/eşim yaptığım yanlışları yüzüme vurduğunda haksızlığa uğramış hissederim.	4	3	2	1		4	3	2	1
8. Partnerim/eşim tarafından baskı altında bırakılmaya katlanamam.	4	3	2	1		4	3	2	1
9. Partnerim/eşim tarafından sert bir şekilde eleştirilmeye tahammül edemem.	4	3	2	1		4	3	2	1
10. Partnerimle/eşimle fikir ayrılığına düştüğümde düşüncelerimi önemsememesine tahammül edemem.	4	3	2	1		4	3	2	1
11. Partnerimin/eşimin ilişkimiz için aldığı bir karar yalnızca ona yarar sağlıyorsa kötüye kullanılmış hissederim.	4	3	2	1		4	3	2	1

Step 4: Have Experts Review the Item Pool- Pilot Test Form

- Consider running a focus group
 - Meet with a small group of individuals to get detailed feedback on their opinions (~5-10 people)
 - Gives you feedback from a sample that is similar to the sample you will eventually give the scale to
 - Especially important if working with “special” population (e.g., children, elderly people)

Step 4: Have Experts Review the Item Pool- Pilot Test Form

- Things you might do in a focus group
 - Identify difficult to read items and ask if the items are confusing or difficult to read (checks reading level)
 - Identify items that you are unsure of whether they measure what you think they measure and ask participants what the items mean to them (checks your construct validity)
 - Ask “How would you answer this statement?” and “Why would you answer it that way?”
 - For each item ask the individuals “Is this something they sample you will eventually use) would say?”

Step 5: Pilot Testing and Item Analysis

- Sometimes you may want to include items that will determine the validity of the final scale
 - Items that might **detect flaws or problems**
- After the expert review, the finalized items are selected for inclusion in the scale.
- This step involves revising, **removing, or rewording items** based on the **feedback** to improve clarity and relevance

Step 6: Applying the Scale to the Target Population

- Administer your initial pool of items along with construct-related and validity items
- How many participants should you collect?
 - Depends on the length of the scale
 - Fewer items requires fewer participants
 - When the ratio of participants to items is low correlations among items can be substantially influenced by chance factors
 - Depends on how representative the development sample is...

A Question😊

***WHAT COULD BE THE POSSIBLE
NONREPRESENTATIVENESS OF THE
DEVELOPMENTAL SAMPLE?***



Step 6: Applying the Scale to the Target Population

○ Possible nonrepresentativeness of the developmental sample

1. Level of the attribute may be different than the population for which the scale is intended
 - Special population??
2. Sample is qualitatively different from the target population

Step 7: Reliability and Validity Analysis

- High correlations among items
- We can make inferences about this relation by examining the correlations among items
 - Higher correlations among items → higher individual item reliabilities
 - More reliable individual items → More reliable scale
 - We therefore want items to be *highly intercorrelated* in a correlation matrix

Step 7: Reliability and Validity Analysis

Reverse scored items

1. To check respondents' attention levels

- Some participants respond quickly or carelessly, marking the same option for all items.

2. To better represent the construct being measured

- Reverse items capture the **opposite pole** of the construct. For example, a self-esteem scale should not include only items like "I feel valuable," but also:
- "I often feel inadequate," which is reverse-coded. This increases **content validity** by representing the full spectrum of the construct.

3. To increase item variety and **reduce measurement error**

- If all items are phrased similarly, the scale may become monotonous and more prone to error. Reverse items diversify item wording and structure, thereby reducing systematic response patterns.

Step 7: Reliability and Validity Analysis

- Cronbach's coefficient alpha
 - Indicator of scale's reliability – how successful you've been...
 - Ranges from 0.0 to 1.0
 - **.70 is acceptable**

Step 8: Optimize Scale Length

Final version of the scale

- Scale length affects reliability
 - Alpha is influenced by the degree of covariation among items and the number of items in the scale
 - Items with *average* inter-item correlations – adding items will increase alpha, removing items will decrease alpha
 - Shorter scales are less burdensome to participants

Step 8: Optimize Scale Length

Final version of the scale

- Dropping bad items
 - Dropping items with sufficiently lower-than-average item correlations will raise alpha
 - Retaining items with slightly below average item correlations will actually increase alpha
- Adjusting scale length
 - Use reliability analyses in SPSS to decide
 - Items with lowest item-scale correlations should be dropped first

Step 8: Optimize Scale Length

Final version of the scale

- Adjusting scale length (continued)
 - **Communality** (i.e., squared multiple correlation) = extent to which the item shares variance with the other items – Açıklanan (paylaşılan) Ortak Varyans- **THAT'S IMPORTANT!!**
 - Items with low communality estimates should be dropped
 - Should see convergence across these methods
 - Also consider that the reliability of alpha as an estimate of reliability increases with the number of items

Any questions??😊