

روایی و پایایی پرسشنامه

دکتر ریحانه احسانی
متخصص پزشکی اجتماعی

▶ شیوه های جمع آوری اطلاعات در پژوهش ها مختلف شامل: مشاهده

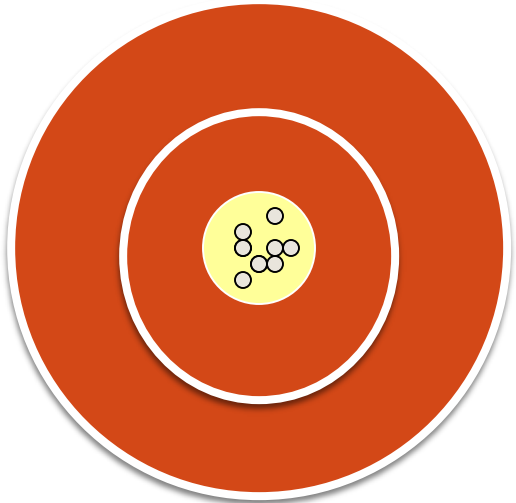
(معاینه) ، مصاحبه و پرسشنامه است. اما قبل از جمع آوری اطلاعات هر یک از روش های فوق نیاز است این سوال مطرح گردد که آیا ابزار مورد استفاده روایی و پایایی لازم را دارد یا خیر؟

▶ چه در طراحی ابزار (پرسشنامه) و چه در بومی سازی

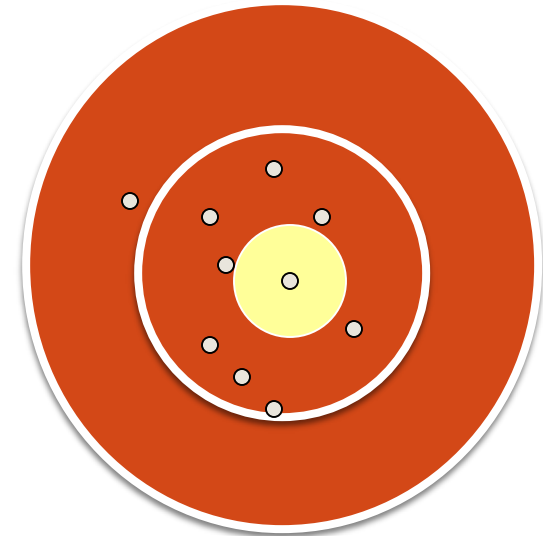
اعتبار (validity)

- ▶ مقدار سنجیده شده چه میزان به واقعیت نزدیک است؟
- ▶ اندازه گیری فشارخون با يك كاف كوچك هر چند ممكن است پایا باشد ولي معتبر نیست.
- ▶ سوال در خصوص بعضي موضوعات خاص ممكن است پایا باشد ولي معتبر نباشد.
- ▶ میزان درآمد شما در ماه چقدر است؟
- ▶ در هفته چند ساعت کتاب می خوانید؟
- ▶ چه میزان سیگار کشیدن را مضر می دانید؟ (به خصوص اگر از افراد سیگاری پرسیده شود)

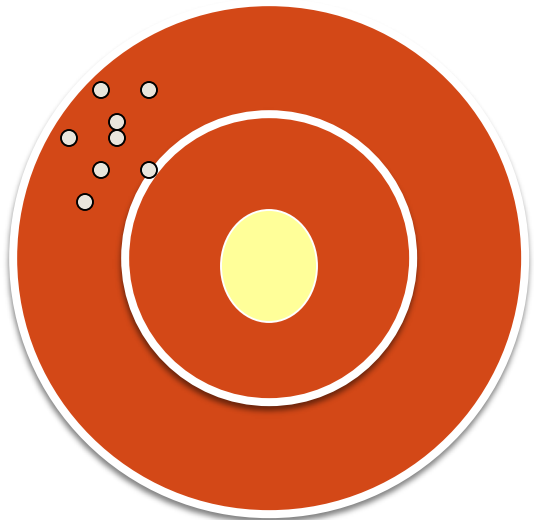
Validity & Reliability:



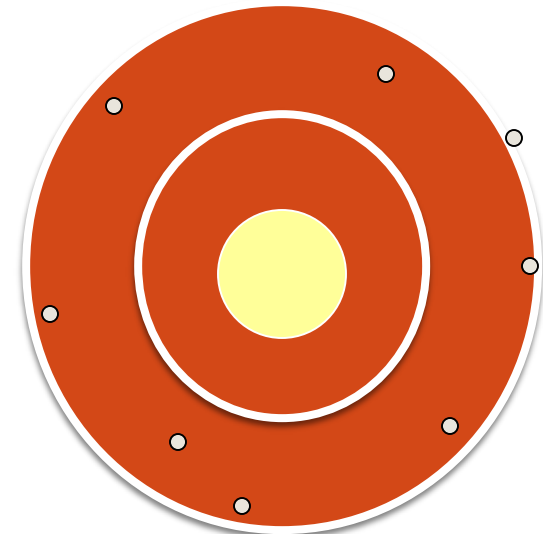
Valid & Reliable



Valid BUT Not Reliable



Reliable BUT Not Valid



Not Valid & Not Reliable

انواع روایی:

1-روایی درونی

قابلیت ابزار مورد نظر در اندازه گیری هدف

-روایی محتوایی

-روایی صوری

-روایی ملاکی (روایی همزمان- روایی پیش بینی)

-روایی سازه (روایی همگرایی- روایی واگرایی)

2-روایی بیرونی

قابلیت تعمیم نتایج بدست آمده به گروهی مشابه یا گروه بزرگتر

1) روایی محتوا content validity

- ▶ اعتبار محتوا نوعی اعتبار است که معمولاً برای بررسی اجزای تشکیل دهنده يك ابزار اندازه گیری به کار برده می‌شود.
- ▶ اعتبار محتوای يك ابزار اندازه گیری به سوال های تشکیل دهنده آن بستگی دارد. اگر سوال های ابزار معرف ویژگی ها و مهارت های ویژه ای باشد که محقق قصد اندازه گیری آنها را داشته باشد، آزمون دارای اعتبار محتوا است.
- ▶ برای اطمینان از اعتبار محتوا، باید در موقع ساختن ابزار (مانند طراحی پرسشنامه) چنان عمل کرد که سوال های تشکیل دهنده ابزار معرف قسمت های محتوای انتخاب شده باشد.
- ▶ اعتبار محتوای يك آزمون معمولاً توسط افرادی متخصص در موضوع مورد مطالعه تعیین می‌شود. از این رو اعتبار محتوا به قضاوت داوران بستگی دارد.

روایی محتوایی :

- ▶ هدف از روایی محتوایی ، ارزیابی محتوای ابزار در اندازه گیری هدف است.
- ▶ برای تعیین روایی محتوایی از قضاوت افراد خبره (expert) ، تحت عنوان پانل خبرگان ، در زمینه تخصصی مورد نظر استفاده می شود .

What is content validity?

- Are we measuring what we think we are measuring?" and refers to the extent to which an instrument actually measures what it alleges to measure.



-
- “Content validity is established by showing that the test items are a sample of a universe in which the investigator is interested.”
 - “Content validity is ordinarily to be established **deductively, by defining a universe of items** and sampling systematically within this universe to establish the test.”



It permit us to answer two question:

- Does the test **cover** a representative sample of the specified construct?
- Is test performance reasonably **free from** the influence of irrelevant variables?



Specific procedure

- **Thorough and systematic** examination of relevant domains and subjects in textbooks and internet
- Consultation with experts

What is Quality of Life?

What are the QoL domains?

What are the objectives of QoL assesment?



Select a panel of experts



نحوه ارزیابی روایی محتوایی :

از دوروش

- ▶ الف-) **کیفی**: در بررسی کیفی اعتبار محتوا، از افراد متخصص خواسته می شود تا ابزار را از نظر این که گویه ها تا چه اندازه ای مفهوم و حیطة هاي آن را پوشش مي دهند بررسی نمایند.
- ▶ از شرکت کنندگان هم می شود استفاده کرد.
- ▶ نظرات خود را نسبت به حذف، تغییر عبارات و یا اضافه نمودن بعضی از عبارات ها به ابزار طراحی شده بیان نمایند.
- ▶ باز خورد خود را ارائه می دهند که بر اساس آن موارد اصلاح می شود.

ب) بررسی روایی محتوایی به روش کمی :

$$CVR = \frac{\left(N - \frac{N}{2} \right)}{\frac{N}{2}}$$

نسبت روایی محتوا^۱ (CVR)

که در آن، N_e ، تعداد کارشناسانی است که سؤال را ضروری می‌دانند و N ، تعداد کل کارشناسان می‌باشد. بدیهی است بر اساس این فرمول اندازه CVR می‌تواند بین -1 و $+1$ تغییر نماید.

▶ از پانل درخواست می‌شود تا هر آیتم را براساس **طیف سه قسمتی** (ضروری است ، مناسب است ، ضرورتی ندارد) بررسی نماید.

▶ **مقدار CVR محاسبه شده** از فرمول با مقدار متناظر آن در جدول مقایسه و اگر از آن بزرگتر باشد اعتبار محتوای آن آیتم تایید می‌گردد.

حداقل مقدار نسبت روایی محتوا (CVR)، آزمون یک دامنه؛ $p = +/0.5$

حداقل مقدار	تعداد پانلیست‌ها
-/99	5
-/99	6
-/99	7
-/75	8
-/78	9
-/62	10
-/59	11
-/56	12
-/54	13
-/51	14
-/49	15
-/42	20
-/37	25
-/33	30
-/31	35
-/29	40

جدول لاوشه

فرض کنید پرسشنامه ای به 20 نفر از پانل متخصصان داده می شود تا در مورد آن نظر دهند و همچنین فرض کنید برای آیتم خاص پاسخ ها به صورت زیر بوده است :

آیتم 1: ضروری	مناسب	غیر ضروری
تعداد نفر 16	2	2

$CVR = 0.6$ که بزرگتر از 0.42 (مقدار جدول) پس تایید می گردد.

شاخص روایی محتوا (CVI)

- ▶ **برای بررسی CVI** ، 3 معیار مرتبط بودن، سادگی و وضوح به صورت مجزا در یک طیف 4 قسمتی برای هر یک از آیتم ها توسط متخصصان مورد بررسی قرار می گیرد. (مثلا 1: غیرمرتبط ، 2: تا حدودی مرتبط ، 3: مرتبط و 4: کاملا مرتبط).
- ▶ امتیاز CVI بوسیله جمع امتیازات موافق برای هر آیتم که رتبه 3 و 4 را کسب کردند بر تعداد کل متخصصان محاسبه خواهد شد .
- ▶ پذیرش آیتم ها براساس CVI بالاتر از 0.79 خواهد بود.

مثال: CVI

فرض کنید هدف ارزیابی محتوایی براساس شاخص CVI است پاسخ ها را به صورت زیر در نظر بگیرید:

آیتم 1 : 1: غیر مرتبط 2: تا حدودی مرتبط 3: مرتبط 4: کاملا مرتبط

تعداد نفر 2 2 8 8

16 با همین روش برای تمام آیتم ها محاسبه

$$CVI = \frac{16}{20} = 0.8$$

می شود و میانگین آن ها به عنوان CVI

مقیاس تعریف می شود در صورتی که مقیاس بالاتر از 0.79 باشد روایی مورد تایید

است.



شاخص روایی محتوا (CVI)

		ساده بودن				واضح بودن				مربوط بودن						
نیمه		نیمه				نیمه				نیمه				نیمه		
		کاملاً ساده است	نسبتاً ساده است	ساده نیست		کاملاً واضح است	واضح است	نسبتاً واضح است	واضح نیست		کاملاً مربوط است	مربوط است	نسبتاً مربوط است	مربوط نیست		
		4	3	2	1	4	3	2	1		4	3	2	1		

روایی صورتی (Face)

- ▶ آیا **ظاهر ابزار** برای آزمودنی ها ، عوامل اجرایی و سایر مشاهده گران مناسب طراحی شده است یا خیر ؟
- ▶ محاسبه به همان روش تعیین روایی محتوا با اندکی تفاوت
- ▶ برای تعیین روایی صورتی از دو روش **کیفی و کمی** استفاده می شود .
- ▶ در تعیین **کیفی روایی صورتی** ، موارد سطح دشواری (**Difficulty**) ، میزان عدم تناسب (**Irrevalancy**) و ابهام (**Ambiguity**) مورد بررسی و اصلاح می شود.
- ▶ از نگارش ، جمله بندی و ظاهر منطقی و جالب برای ابزار استفاده شده ؟

▶ در گام بعدی جهت کاهش و حذف آیتم های نامناسب و تعیین اهمیت هر یک از آیتم ها ، از روش کمی تاثیر آیتم (**Item Impact Method**) استفاده خواهد شد .

▶ $\text{Impact score} = \text{importance} \times \text{frequency}(\%)$

▶ مثال :

▶ آیتم 1:1:خیلی ضعیف 2:ضعیف 3:متوسط 4:قوی 5:خیلی قوی

▶ تعداد نفر 5 10 20 10 5

▶ فراوانی % 10 %20 %40 %20 %10

▶ $1*0.1+2*0.2+3*0.4+4*0.2+5*0.1=3.3$

▶ با توجه به اینکه امتیاز میزان تاثیر این آیتم بیشتر از 1.5 می باشد آیتم برای تحلیل بعدی مناسب و حفظ خواهد شد .

اعتبار ملاکی

- ▶ **روایی ملاکی (Criterion – related)**: میزان همبستگی بین نمرات حاصل از یک ابزار با نمرات حاصل از ابزار اندازه گیری دیگر (ملاک یا GS) است. که دو نوع است:
(در عمل سخت است)
- ▶ **الف) روایی همزمان**: میزان توافق دو فرم هم ارز در اندازه گیری هدف واحد
- ▶ **ب) روایی پیش بینی**: میزان همبستگی دو ارزیابی روی هدف واحد در دو زمان متفاوت

الف) اعتبار همزمان (Concurrent) :

- ▶ میزان توافق دو فرم هم ارز در اندازه گیری هدف واحد
- ▶ اعتبار همزمان در مواردی به کار می‌رود که داده‌های حاصله از دو اندازه گیری در يك زمان در دسترس باشد.
- ▶ در این گونه موارد عملکرد در يك آزمون به عملکرد در آزمون دیگر مرتبط می‌گردد. این فرایند اعتبار همزمان نامیده می‌شود.
- ▶ اعتبار همزمان در مواردی محاسبه می‌شود که هدف جانشین کردن يك ابزار اندازه گیری به جای ابزار دیگری باشد. این امر بیشتر به علت ملاحظات مربوط به سهولت اجرای آزمون یا جانشین کردن يك آزمون کوتاه‌تر به جای يك آزمون طویل‌تر است. در این مورد نیز ضریب همبستگی بین نمره‌های حاصل از اجرای آزمون با نمره‌های بدست آمده از اجرای آزمون دیگر یا اندازه‌های حاصل از اندازه گیری دیگری که همزمان به عمل آمده است به عنوان میزان اعتبار اندازه گیری به کار می‌رود.

ب) اعتبار پیش بین (Predictive):

- ▶ اعتبار پیش بین عبارتست از بررسی رابطه نمره های آزمونی که برخی ویژگی ها را می سنجد و آن چه ادعای پیش بینی آن را دارد.
- ▶ اعتبار پیش بین می تواند به وسیله رابطه عملکرد در يك آزمون با آزمون ملاك رفتاری به دست آید. به عبارت دیگر ضریب همبستگی نمره های حاصل از اجرای آزمون با نمره های متغیر ملاك، نمایانگر اعتبار پیش بین است.
- ▶ برای مثال در يك تحقیق، 882 مرد و زن به يك آزمون رغبت شغلی پاسخ گفته اند. آزمونی ها 12 ساله بودند و در 19 سالگی مجدد مورد بررسی قرار گرفتند. 51 درصد به شغلی که در آزمون انتخاب کرده بودند اشتغال داشتند. این داده ها اعتبار پیش بین این ابزار را نشان می دهد.

روایی سازه ای (Construct validity)

- ▶ به این موضوع می پردازد که آیا آیتم های سازنده یک سازه به **صورت معقولی** در کنار همدیگر سازه مورد نظر را اندازه گیری می کنند یا خیر؟ یا به عبارت دیگر آیا اجزای مقیاس مورد نظر توانایی تشکیل مقیاس را دارند یا برخی از آنها نا مرتبط اند
- ▶ این روایی به **روایی عاملی** نیز معروف است .
- ▶ روایی سازه ای به دو دسته کلی تقسیم می شود:
- ▶ **1-روایی همگرایی:** به همگرایی (همسوبودن یا همبستگی) آیتم ها اشاره می کند
- ▶ **2-روایی افتراقی:** آیتم های تشکیل دهنده دو مقیاس نباید به گونه ای با یکدیگر همبستگی داشته باشد که بتوان گفت موضوع واحدی را اندازه می گیرند

انواع روایی سازه ای :

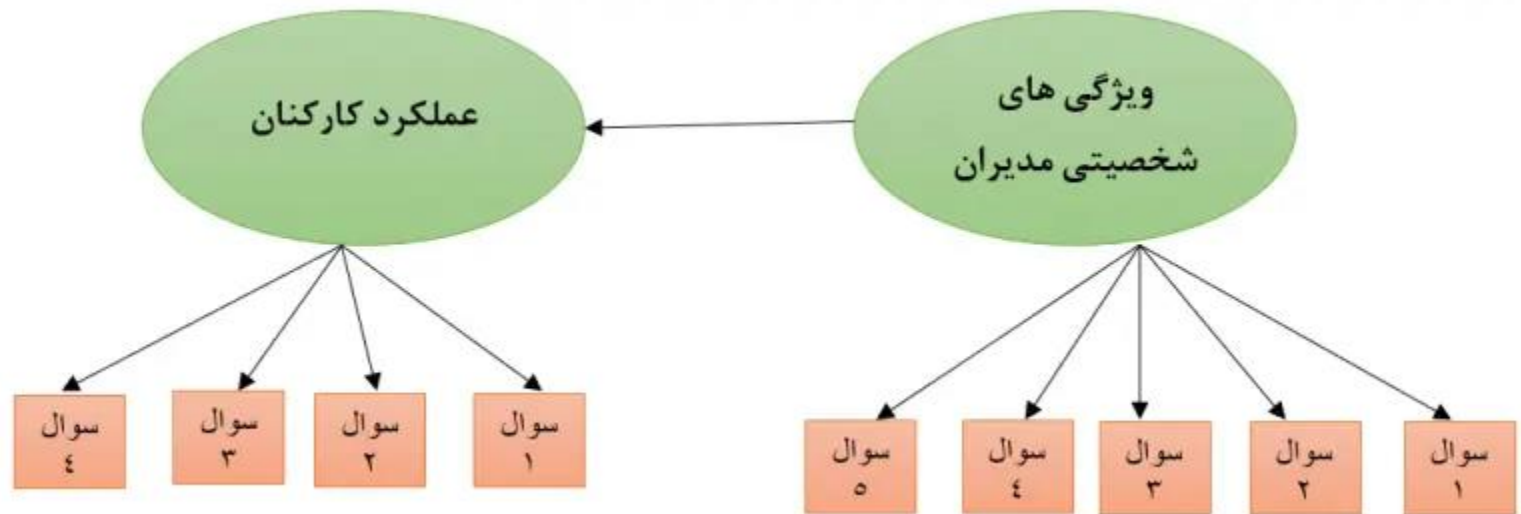
► الف) روایی همگرایی (Convergence):

External: یکی از روش های ارزیابی: همبستگی میان مقیاس مورد بررسی با ابزارهای مشابه ای که برای اندازه گیری ، سازه مورد نظر (توسط سایر مولفین) طراحی شده اند.

Internal: بیشترین همبستگی هر سوال باید با دامنه خودش باشد. همبستگی خود سوال با مجموع نمره دامنه خودش (با حذف سوال) بررسی می شود.

(Discriminant validity) (ب) روایی افتراقی یا Divergent واگرایی

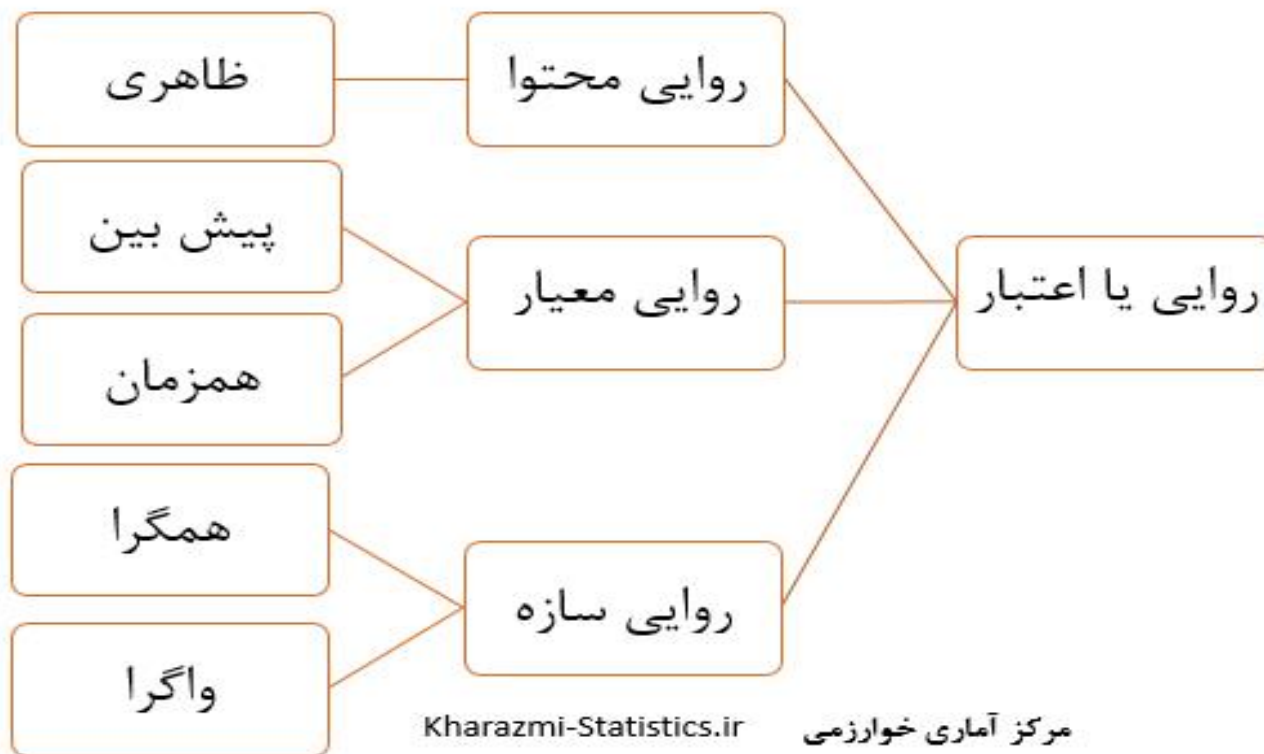
▶ روایی افتراقی بدین منظور انجام می گیرد که به **لحاظ آماری** آزمون شود آیا دو سازه ی مختلف با یکدیگر تفاوت دارند یا خیر؟ و این موضوع دقیقا **در نقطه مقابل** این است که آیا آیتم های مختلف یک سازه ، در داخل آن سازه همگرا می شوند و یا به عبارت دیگر ، همسانی درونی دارند یا خیر؟



Construct validity has two components:

- *Convergent validity* demonstrates association with measures that are or should be related.
- *Divergent validity* demonstrates a lack of association with measures that should not be related.





پایایی یا ثبات (reliability)

▶ در صورت تکرار يك اندازه گيري تا چه میزان مقادير سنجیده شده به یکدیگر نزدیک هستند.

▶ اگر دو مرتبه از يك فرد به فاصله چند دقیقه و در شرایط کاملاً یکسان فشارخون گرفته شود تا چه میزان مقادير به یکدیگر نزدیک هستند؟

▶ اگر به فاصله چند روز از يك فرد در خصوص تعداد ساعات مشاهده تلویزیون سوال شود آیا اعداد بیان شده به یکدیگر نزدیک هستند؟

Stability : ثبات (آزمون مجدد)

▶ در این روش پاسخ دهندگان 2 بار همان ابزار را تکمیل می کنند و یک ضریب همبستگی را برای رابطه بین نمرات ابزاری را که در دفعه اول و دفعه دوم برای همان گروه پاسخ دهنده حساب می کنند .

Reliability	Interpretation
.90 and above	Excellent reliability; at the level of the best standardized tests
.80 - .90	Very good for a classroom test
.70 - .80	Good for a classroom test; in the range of most. There are probably a few items which could be improved.
.60 - .70	Somewhat low. This test needs to be supplemented by other measures (e.g., more tests) to determine grades. There are probably some items which could be improved.
.50 - .60	Suggests need for revision of test, unless it is quite short (ten or fewer items). The test definitely needs to be supplemented by other measures (e.g., more tests) for grading.
.50 or below	Questionable reliability. This test should not contribute heavily to the course grade, and it needs revision.

ضریب آلفا در SPSS

To compute an alpha coefficient:

From the Data Editor Screen:

√Analyze. . . .Scale. . . .√Reliability analysis

Highlight and transfer scale items L→R column

Alpha is the default in the drop down box

√OK

**EXHIBIT 10.2. SPSS PRINTOUT FOR THE ALPHA COEFFICIENT
FOR THE MATERNAL INVOLVEMENT SCALE.**

Case Processing Summary

		N	%
Cases	Valid	560	98.9
	Excluded ^a	6	1.1
	Total	566	100.0

Reliability Statistics

Cronbach's Alpha	N of Items
.764	6

Alpha coefficient

^a Listwise deletion based on all variables in the procedure.

▶ روش دونیمه کردن آزمون:

در این روش آزمون به دو نیمه تقسیم شده و رابطه همبستگی بین دو نیمه آزمون سنجیده میشود. در این روش، محتوا و سختی سوالات باید با هم مشابه باشند و از آنجا که تعداد سوالات ابزار به دو قسمت تقسیم شده است بایستی روش ضریب همبستگی به کار رود.

