

روایی درونی

این موضوع که آیا تغییر حاصل در ب (متغیر وابسته) منحصرأ ناشی از تغییر در الف (متغیر مستقل) است روایی درونی نامیده می‌شود. "روایی درونی حداقلی اساسی است که بدون آن هیچ آزمایشی تفسیر پذیر نیست. آیا مداخلات آزمایشی در این مورد خاص واقعاً تفاوتی ایجاد کرده‌اند؟ (کمپبل و استانلی² 1966). عوامل تهدید کننده‌ای که اگر کنترل نشوند روایی درونی را به مخاطره می‌اندازند عبارتند از:

1- تاریخ³. رویدادهایی بجز مداخله آزمایشی که در فاصله زمانی بین پیش آزمون⁴ و پس آزمون⁵

به‌وقوع می‌پیوندند. بین پیش آزمون و پس آزمون ممکن است رویدادهای تغییرزای بسیار دیگری علاوه بر مداخله آزمایشی رخ دهند. اگر پیش آزمون (O1) و پس آزمون (O2) در روزهایی متفاوت انجام گرفته باشند، آنگاه رویدادهای بین آنها ممکن است موجب تفاوت شده باشند. برای اینکه تاریخ به صورت فرضیه رقیب موجهی درآید، چنین رویدادی باید در مورد اکثر آزمودنی‌های مورد مطالعه رخ بدهد، مثلاً در کلاس درس دیگر یا از طریق برنامه ای خبری که پخش گسترده ای داشته است. هرچه مدت زمان بین پیش آزمون و پس آزمون بیشتر باشد تاریخ توضیح رقیب موجه تری برای تغییر می‌گردد. در آزمایشی که در مدت یکی دو ساعت انجام می‌شود تاریخ ممکن است مسئله ای جزئی قلمداد شود، گرچه حتی در اینجا نیز باید مراقب منابع خارجی از قبیل خنده، رویدادهای مخل و غیره بود. دسته ای از اثرهای ممکن مربوط به فصل سال یا مربوط به برنامه رویدادهای موسسه ای را با متغیر تاریخ طبقه بندی می‌کنیم. بدین ترتیب، خوشبینی ممکن است با فصول سال و اضطراب با برنامه امتحانات نیمسال تغییر کنند. چنین اثرهایی ممکن است بین پیش آزمون و پس آزمون تغییری ایجاد کنند که با اثر مداخله آزمایشی اشتباه گردند.

2- رشد و نمو¹. تغییراتی که با گذشت زمان در آزمودنی‌ها روی داده و از مداخله آزمایشی مستقل‌اند.

اصطلاح رشد و نمو در اینجا دلالت بر همه فرآیندهای بیولوژیکی یا روانشناختی دارد که مستقل از رویدادهای بیرونی ویژه، با گذشت زمان به‌صورتی منظم تغییر می‌کنند. بدین ترتیب، بین پیش آزمون و پس آزمون، دانش آموزان ممکن است مسن تر، گرسنه تر، خسته تر، کسل تر و غیره شده باشند و تفاوت حاصل ممکن است بازتابنده این متغیرها باشد تا مداخله آزمایشی.

3- آزمودن². اثرهای گرفتن یک آزمون بر متغیر به‌دست آمده در آزمون دیگر، یا تغییراتی که در

عملکرد آزمودنی در نتیجه آشنا شدن وی به آن زمینه به علت سنجش پیشین عملکرد روی داده است. در آزمون‌های پیشرفت تحصیلی و هوش، دانش آموزانی که برای بار دوم آزمونی را می‌گیرند، یا صورت دیگری از آزمون را می‌گیرند، معمولاً از آنهاپی که برای نخستین بار آزمون را می‌گیرند

بهبتر عمل می کنند (انستازی³ 1958 صفحات 190-191، کین وهایم⁴، 1950). این اثرهای به‌طور متوسط به‌اندازه سه تا پنج نمره بهره هوشی نسبت به آزمون‌گیرندگان بی تجربه افزایش ایجاد می‌کنند. در رابطه با آزمون‌های شخصیتی، اثر مشابهی مشاهده شده است، به‌طور کلی با آزمون‌های دوم سازگاری بهتری نشان داده می‌شود. در یک مجموعه آزمون شخصیتی یا سازگاری که در بالای آنها نام افراد نوشته می‌شود، بار اول برگزاری آزمون در حکم موقعیتی از مشکل گشائی⁵ است که در آن دانش آموز تلاش می‌کند تا هدف تغییر قیافه یافته آزمون را کشف کند. با انجام این کار یا با گفتگو با دوستانش در مورد برخی سئوالات عجیب و غریب، وی یاد می‌گیرد که بار دوم چگونه باید خود را بنحو قابل قبولی عرضه کند.

با معرفی مشکل اثرهای آزمودن تمایزی می‌ان‌سنجش‌های بالقوه در مورد واکنش دار بودن آنها پیش می‌آید. این موضوع مدت‌های مدید در علوم اجتماعی یک امر بدیهی بوده است که فرآیند سنجش ممکن است آنچه را که سنجیده می‌شود تغییر دهد. اثر واکنشی را هر وقت که فرآیند آزمودن فی‌نفسه محرکی برای تغییر است تا ضبط فعل پذیر رفتار، می‌توان انتظار داشت. بدین ترتیب، قرار دادن مشاهده‌کنندگانی در کلاس درس برای مشاهده مهارت‌های روابط انسانی پیش آموزش معلم ممکن است بخودی خود نحوه عمل معلم را تغییر دهد. گذاشتن می‌کروفونی روی میز ممکن است انگاره تعامل گروهی را دگرگون سازد و غیره. به‌طور کلی انتظار می‌رود که هرچه ابزار آزمون بدیع‌تر و انگیزنده‌تر باشد، واکنشی‌تر باشد.

4- به‌کارگیری یا فرسودگی ابزار⁶. تغییرات در درجه بندی یک ابزار سنجش یا تغییرات در مشاهده‌کنندگان یا نمره‌های به‌کارگرفته‌شده ممکن است در سنجش‌های حاصل تغییراتی ایجاد کند. بنابراین، تغییرات در سنجش‌های عملکرد آزمودنی ممکن است ناشی از تغییرات در وسائل یا شرایط سنجش باشد، از قبیل فرسودگی قطعات در مورد ابزار یادگیری، یکنواختی یا خستگی برای ناظران انسانی. پس این اصطلاح به تغییرات مستقل در ابزار سنجش اشاره دارد که ممکن است تفاوت بین پیش‌آزمون و پس‌آزمون را تبیین سازند. جایی که مشاهده‌کنندگان انسانی برای به‌دست دادن پیش‌آزمون و پس‌آزمون بکار می‌روند، فرآیندهای یادگیری، خستگی و غیره در مشاهده‌کنندگان تفاوت‌های بین پیش‌آزمون و پس‌آزمون را ایجاد خواهد کرد. اگر به مقالاتی نمره داده می‌شود، معیارهای نمره‌گذاری ممکن است بین O1 و O2 تغییر کند (این امر موجب این پیشنهاد شد که به عنوان یک وسیله‌مهار، مقالات مربوط به O1 و O2 در هم شده و بدون اطلاع از اینکه کدام مربوط به پیش‌آزمون و کدام مربوط به پس‌آزمون است نمره داده شوند). اگر مشارکت در کلاس درس مورد مشاهده است، آنگاه مشاهده‌کنندگان ممکن است در نوبت دوم ماهرتر، یا لاقیدتر شوند. اگر والدین مصاحبه می‌شوند، آشنایی مصاحبه‌گر با برنامه مصاحبه و با

آن والدین خاص ممکن است تغییراتی را موجب شود. یک تغییر در مشاهده کنندگان در فاصله بین 01 و 02 می تواند علت تفاوت باشد. .

5- بازگشت آماری¹. این عامل هنگامی عمل می کند که گروهها براساس نمره‌های انتهای آنها انتخاب شده باشند، بدین معنی که افراد بر حسب نمره‌های بالا یا پایین جهت گماردن به گروهها انتخاب شوند. مثلاً در یک آزمایش بهبودی بخش، دانش آموزان بعلت اینکه در یک آزمون پیشرفت تحصیلی (که برای آنها 01 محسوب می شود) ضعیف عمل کرده‌اند برای یک مداخله آزمایشی ویژه انتخاب می شوند. پس از اعمال مداخله بعنوان پس آزمون 02 یک فرم موازی را می گیرند یا همان آزمون پیشرفت تحصیلی را تکرار می کنند. میانگین 02 برای این گروه تقریباً به‌طور مطمئنی بیشتر از 01 خواهد بود. این نتیجه ناشی از هیچ‌گونه اثر واقعی مربوط به مداخله آزمایشی (X) ، یا اثر تمرین آزمون - با آزمون و غیره نیست، بلکه جنبه ای از همبستگی غیر کامل بین 01 و 02 است.

6- به دلیل اهمیت آگاه شدن بر این مطلب شرحی مقدماتی درباره آن عرضه می گردد. شکل 1 داده‌هایی مصنوعی ارائه می‌دهد که در آنها پیش آزمون و پس آزمون برای کل جمعیت همبستگی 0/50 دارند بدون اینکه در میانگین یا واریانس² گروهی تفاوتی وجود داشته باشد. خط بازگشت ب نشان دهنده بهترین پیش بینی از پیش آزمون به پس آزمون است.

7	8	9	10	11	12	13	میانگین پس آزمون
			1	1	1	1	11/5
		1	1	2	1	1	11
	1	2	3	3	2	1	10/5
1	1	3	4	3	1	1	10
1	2	3	3	2	1		9/5
1	1	2	1	1			9
1	1	1	1				8/5

میانگین
پیش آزمون

11/5
11
10/5
10
9/5
9
8/5

شکل 1- الف. نمودار پراکندگی فراوانی نمره‌های پس آزمون برای هر طبقه از نمره‌های پیش آزمون و بالعکس.

خط بازگشت ج که نشان دهنده بهترین پیش بینی از پس آزمون به پیش آزمون است.

پیش بینی از گروه‌های همگن پس آزمونی	به میانگین پیش آزمون	به میانگین پس آزمون	پیش بینی از گروه‌های همگن پیش آزمونی
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0
0	0	0	0

ج-ا

شکل

شکل 1-ب

شکل 1- بازگشت هنگام پیش بینی نمره‌های پس آزمون از پیش آزمون و بالعکس

در این مورد فرضی، هیچ‌گونه تغییر حقیقی، به‌وقوع نپیوسته است، بلکه همان‌طور که معمول است نمره‌های خطا آمیز¹ آزمون یک همبستگی با بازآزمون نشان می‌دهد که به‌طور قابل توجهی کمتر از یک است. اگر همان‌طور که در مثال بالا پیشنهاد شد شخص تنها به افرادی بنگرد که نمره‌های خیلی پایین دارند، مثلاً نمره‌های 7 و تنها به نمره‌های پس آزمون این دانش آموزان نظر افکند، درمی‌یابد که نمره‌های پس آزمون آنها پراکنده‌اند. لیکن رویهمرفته بهتر بوده و به‌طور متوسط تا نیمه راه میانگین گروهی بازگشت کرده‌اند (بدین معنی که ضریب همبستگی 0/50 است) و منتج به میانگین 8/5 شده‌اند. لیکن بجای اینکه

1-Statistical regression

11

2-Variance

1-fallible

2-Mc Nemar

این بازگشت شاهدهی از پیشرفت باشد بیانی مجدد، هر چند ویژه، از واقعیت غیر کامل بودن همبستگی و درجه آنست.

چون زمان می گذرد و بین پیش آزمون و پس آزمون رویدادهائی رخ می دهند، شخص وسوسه می گردد تا این تغییر را بنحوی علی در جهت گذشت زمان مربوط کند. لیکن توجه کنید که یک تحلیل در جهت عکس گذشت زمان هم در اینجا ممکن است. اگر به کسانی بنگریم که نمره‌های پس آزمون آنها 7 است و به پراکندگی نمره‌های پیش آزمون آنها نظر افکنیم، مضمونی معکوس از آنها حاصل می شود، بدین معنی که نمره‌ها بدتر شده‌اند. اشتباه آمیزترین استنباطات علی هنگامی به عمل می‌آیند که داده‌ها به صورت شکل 1.1. ب عرضه شده باشند (یا بخش فوقانی یا تحتانی شکل ا. ب). در اینجا بنظر می‌رسد که باهوش‌ها کودنتر و کودن‌ها باهوشتر شده‌اند و گوئی این امر در نتیجه اثر خنثی کننده و همگن‌ساز محیط به عمل آمده است. در حالی که این تفسیر غلط حاکی از این است که پراکندگی جمعیت در پس آزمون باید کمتر از پیش آزمون باشد، هر دو پراکندگی در واقع برابرند. وانگهی، با تحلیل گروه‌هائی ناب از نمره‌های پس آزمون (مثل خط بازگشت ج و شکل 1.1 ج) می‌توانیم استنباط متضادی به عمل آوریم. همان‌طور که مک نمار² (1940) خاطر نشان ساخته است استعمال تحلیل‌های مهار کننده با زمان معکوس و آزمایش مستقیم برای کشف تغییرات در پراکندگی‌های جمعیت احتیاط‌های مفیدی علیه چنین تفاسیر اشتباه آمیزند.

می‌توانیم به موضوع بازگشت به سوی میانگین به طریق مربوط دیگری نیز نظر افکنیم. هرچه نمره‌ای منحرف‌تر باشد، مقدار خطای اندازه‌گیری که در بردارد احتمالاً بیشتر است. بدین ترتیب که در یک معنا کسی که معمولاً نمره خیلی بالائی گرفته است به‌طور غیر معمولی از بخت و اقبال نیکوئی بهره مند شده است (خطای زیاد مثبت) و کسی که نمره خیلی پایینی کسب کرده است بد شانس بوده است (خطای زیاد منفی). از آنجا که بخت و اقبال بی ثبات است، در یک پس آزمون انتظار می‌رود که کسانی که نمره بالا گرفته‌اند به‌طور متوسط افت کنند و آنهائی که نمره پایین گرفته‌اند موضع نسبی خود را بهبود بخشند. (همین منطق هم وقتی که شخص از نمره‌های پس آزمون آغاز کرده و به نمره‌های پیش آزمون برگشت نماید حکم فرماست).

بازگشت بسوی میانگین پدیده‌ای همه جا حاضر است و محدود به پیش آزمودن و پس آزمودن با یک آزمون واحد یا صورت‌های مقایسه پذیر یک آزمون نیست. مدیر مدرسه‌ای که مشاهده می‌کند باهوش‌ترین دانش آموزان بالاترین نمره را در آزمون پیشرفت تحصیلی ندارند و کودن‌ترین دانش آموزان پایین‌ترین نمره را در آزمون پیشرفت تحصیلی نمی‌گیرند، اگر اعلام کند که مدرسه اش باهوش‌ترین دانش آموزان را به‌اندازه کافی بر نمی‌انگیزد و کودن‌ترین دانش آموزان را سخت بکار می‌کشد مرتکب سفسطه بازگشت شده است. انتخاب دانش آموزانی که بالاترین و پایین‌ترین نمره را در آزمون پیشرفت تحصیلی دارند و

نگریستن به بهره‌های هوشی آنها وی را مجبور می‌سازد که بهمان طریق غیر منطقی نتیجه‌گیری متضادی به‌عمل آورد.

با اینکه بازگشت در اینجا بر حسب خطاهای اندازه‌گیری مورد بحث قرار گرفته است، بنحوی کلی‌تر تابعی از درجه همبستگی است. هرچه همبستگی کمتر باشد، بازگشت به سوی میانگین بیشتر است. عدم همبستگی کامل ممکن است ناشی از خطا باشد و یا به‌علت منابع منظم پراکندگی خاص یک سنجش یا دیگری. بدین ترتیب، اثرهای بازگشتی ملازم‌های اجتناب‌ناپذیری از همبستگی غیر کامل آزمون - بازآزمون برای گروه‌هایی است که بخاطر انتهائی بودنشان انتخاب شده‌اند، هرچند که این اثرهای همه‌جا ملازم‌های ضروری نمره‌های انتهائی نیستند.

اگر گروهی که بدلائل مستقل انتخاب شده است میانگینی انتهائی داشته باشد، انتظار کمتری می‌رود که میانگین گروهی در یک آزمون مجدد بازگشت کند. چه منابع تصادفی یا خارجی واریانس آزاد بوده‌اند تا نمره‌های اولیه را در هر دو جهت تحت تاثیر قرار دهند. لیکن برای گروهی که بعلت انتهائی بودن آن در یک متغیر خطا آمی‌ز انتخاب شده است، این موضوع صحت ندارد. انتهائی بودن آن مصنوعی است و بسوی میانگین جمعیتی که از آن برگزیده شده است بازگشت خواهد کرد.

اثرهای بازگشتی از گونه‌ای غیر مستقیم‌تر می‌توانند ناشی از انتخاب دارندگان نمره‌های انتهائی در سنجش‌هایی بجز پیش‌آزمون باشند. موردی را در نظر بگیرید که در آن دانش‌آموزانی که در امتحان یک درس رد شده‌اند برای آموزش انفرادی انتخاب شوند بعنوان یک پیش‌آزمون فرم الف بکار می‌رود. احتمال می‌رود که آزمون کلاس درس با فرم الف همبستگی بیشتری داشته باشد تا با فرم ب که سه ماه دیرتر بکار رفته است (اگر این آزمون در هر دو نوبت به همه کلاس داده شده باشد). هرچه همبستگی بیشتر باشد، بازگشت بسوی میانگین کمتر است. بدین ترتیب مردودان کلاس در رابطه با پیش‌آزمون بازگشت صعودی کمتری خواهند داشت تا نسبت به پس‌آزمون. این امر پیشرفتی کاذب فراهم می‌کند که ممکن است با اثر تلاش آموزش انفرادی اشتباه گرفته شود.

6-گزینش¹. سوگیری‌های² آگاهانه و ناآگاهانه می‌توانند منجر به گزینش اختلافی در آزمودنی‌های گروه‌های مقایسه شوند. بنابراین، اگر گماردن آزمودنی‌ها به گروه‌های متفاوت (گروه‌های آزمایشی و گواه) در طرح بر هر مبنائی بجز گماردن تصادفی از یک جمعیت مشترک باشد، آن دیگر مبانی گزینش سوگیری‌هائی اصولی ببار خواهند آورد. پس اگر 01 و 02 با هم متفاوت باشند، این تفاوت می‌تواند ناشی از انتخاب اختلافی افرادی باشد که گروه‌ها را تشکیل می‌دهند. بدین معنی که گروه‌ها ممکن بود به‌رحال و بدون وقوع مداخله آزمایشی (X) با هم متفاوت باشند.

7- افت آزمایشی³. افت آزمایشی عبارتست از ازدست‌دادن آزمودنی‌های گروه‌های مقایسه بنحو متفاوت. این عامل تهدیدکننده مبین این است که تفاوت‌های بین O1 و O2 ممکن است ناشی از افت اختلافی در آزمودنی‌های گروه‌های آزمایشی و گواه باشد. بدین ترتیب، حتی اگر دو گروه در آغاز با هم یکسان باشند، ممکن است هنگام پس آزمون با هم متفاوت باشند. این امر به علت هیچ‌گونه تغییری از ناحیه اعضای گروه‌ها نیست، بلکه به دلیل افت انتخابی افراد از یکی از گروه‌هاست. در پژوهش‌های آموزشی این مشکل اغلب در مطالعاتی خودنمایی می‌کند که هدف آنها تعیین اثرهای آموزش دانشگاهی است. براساس مقایسه سنجش‌هائی از دانشجویان سال اول (که مداخله آزمایشی را دریافت نکرده‌اند) با دانش آموزان سال چهارم (که مداخله آزمایشی را دریافت کرده‌اند).

8- تعامل‌ها با گزینش⁴. بسیاری از تهدیدهای فوق‌الذکر نسبت به روایی درونی می‌توانند با گزینش تعامل کرده و نیروهایی پدید آورند که ممکن است مصنوعاً به‌عنوان اثرهای مداخله آزمایشی قلمداد شوند. شاید معمول‌ترین این نیروها تعامل گزینش و رشد و نمو باشد. این تعامل وقتی حاصل می‌شود که گروه‌های آزمایشی مرکب از انواع مختلفی از افراد باشند که با سرعت‌های متفاوت رشد می‌کنند. اگر گروه آزمایشی مرکب از بیماران روان‌درمانی باشد و گروه گواه متشکل از افراد دیگری که آزمون و بازآزمون شده‌اند. یک دستاورد ویژه گروه آزمایشی ممکن است به‌عنوان فرایند بهبودی خود بخود که خاص چنین گروهی افراطی است تفسیر شود، دستاوردی که حتی بدون مداخله آزمایشی هم به‌وقوع می‌پیوندد. چنین تعاملی بین گزینش و رشد و نمو (یا تعامل بین گزینش و تاریخ، یا تعامل بین گزینش و آزمون) می‌تواند با اثر مداخله آزمایشی اشتباه گرفته شود و بدین ترتیب تهدیدی برای روایی درونی آزمایش به‌شمار می‌آید. فرضیه تعامل بین گزینش و رشد و نمو گاهی حتی در جایی هم که گروه‌ها در نمره‌های پیش آزمون یکسانند قابل پذیرش خواهد بود. معمول‌ترین این موارد در جایی است که یک گروه میزان بیشتری از رشد و نمو یا تغییر خودبخود از گروه دیگر داشته باشد.

9- ابهام در مورد جهت نفوذ علی¹. این تهدید موقعیتی است که در آن کلیه توضیحات مربوط به متغیرهای موجه ثالث در ارتباط با رابطه الف با ب رد شده‌اند، ولی روشن نیست که آیا الف علت ب است یا ب علت الف. این تهدید نسبت به روایی درونی در مطالعات همبستگی ساده که اغلب روشن نیست علت کدام و معلول کدامست از برجستگی برخوردار است. مثلاً، معلوم نیست که آیا سرپرستی کمتر موجب بهره‌وری بیشتر می‌شود یا بهره‌وری بیشتر سبب سرپرستی کمتر می‌گردد. این تهدید خاص در اکثر آزمایش‌ها از برجستگی برخوردار نیست، چه ترتیب زمانی روشن است. همچنین در آن مطالعات همبستگی که یک جهت از نفوذ علی نسبتاً غیر موجه است این

تهدید مشکلی محسوب نمی شود (مثلاً موجه‌تر این است که استنباط کنیم کاهش در درجه حرارت هوای محیط موجب افزایش در مصرف مواد سوختی است تا اینکه افزایش در مصرف مواد سوختی باعث کاهش درجه حرارت محیط است) لیکن در تمامی دیگر مطالعات همبستگی مشکلی به حساب می آید.

به کاربردن آزمایش‌های تصادفی حفاظ مناسبی علیه تمام تهدیدهای فوق‌الذکر نسبت به روایی درونی فراهم می‌کند. لیکن پنج تهدید زیر ممکن است روایی درونی آزمایش‌های حقیقی و هم روایی درونی شبه آزمایش‌ها را به مخاطره افکنند. این‌ها نمایشگر تهدیدهای اضافی هستند که در فهرست تهدیدهای کمپبل و استانلی (1963, 1966) نیامده‌اند. سه تهدید اول اثر معادل ساختن گروه‌های آزمایشی و گواه را دارند درحالی‌که دو تهدید آخر ممکن است تفاوتی تصنعی ایجاد کنند.

10- پخش شدن مداخله یا تقلید کردن از آن². اگر مداخله‌های آزمایشی متضمن برنامه‌هایی اطلاعاتی باشند که کاملاً پخش شده باشند یا اگر گروه‌های آزمایشی و گواه بتوانند با یکدیگر ارتباط برقرار سازند، آزمودنی‌های گروه گواه ممکن است به یادگیری آن اطلاعات پرداخته و به دریافت مداخله نائل آیند. بدین ترتیب، آزمایش بی‌روایی می‌گردد چه عملاً تفاوتی بین گروه‌های آزمایشی و گواه در مداخله‌هایی که دریافت کرده‌اند وجود ندارد. در بسیاری از شبه آزمایش‌ها مطلوب این است که واحدهایی داشته باشیم که حتی المقدور در همه جهات بجز مداخله آزمایشی مشابه باشند. یک طریقه برای انجام این کار این است که از واحدهایی که مجاورت فیزیکی دارند نمونه‌گیری کنیم. لیکن خود این قرابت ممکن است منجر به این گردد که مداخله توسط کلیه واحدها تجربه شود.

11- معادل ساختن جبرانی مداخله³. هنگامی که مداخله آزمایشی نتایجی ببار آورد که عموماً مطلوب پنداشته شوند ممکن است بعلت احساس نابرابری در گروه گواه اکراه و بی‌میلی ظاهر شود و برای اجتناب از این امر نتایج مشابهی در اختیار گروه گواه نیز قرار بگیرد. این امر ممکن است گروه گواه را بر آن دارد تا واکنشی نظیر گروه آزمایشی به‌عمل آورد و در نتیجه اثرهای مداخله آزمایشی را پوشیده نگه دارد.

12- رقابت جبرانی⁴. جائیکه گماردن اشخاص یا واحدهای سازمانی به شرایط آزمایشی و گواه علنی باشد، شرایط رقابت اجتماعی پیش می‌آید. گروه گواه به تلاشی ناشی از همچشمی برانگیخته می‌گردد تا تفاوت مورد انتظار را کاهش داده یا معکوس سازد. این نتیجه به ویژه درجائی احتمال می‌رود که واحدهایی کامل (از قبیل بخش‌ها، کارخانه‌ها، گروه‌های کار و غیره) به مداخله‌ها گمارده شوند یا اینکه اگر مداخله موفقیت آمیز باشد موجب زیان و خسارت برای گروه گواه گردد. مثلاً سارتسکی⁵ (1972) خاطر نشان ساخته است که عملکرد قراردادی، امنیت شغلی معلمان را تهدید

می کند. وی شواهدی را بازنگری کرده است حاکی از اینکه عملکرد آموزشی کودکانی که توسط معلمان گروه گواه تدریس شده بودند در اثنای دوره آموزشی بهتر از سال‌های گذشته بود. دستاوردهای یادگیری بهتر توسط گروه‌های گواه منتج به این خواهد شد که تفاوت یادگیری دانش آموزان گروه گواه که توسط معلمان معمولی خود آموزش می دیدند و کودکان آزمایشی را که توسط معلمان پیمانی که برحسب پیشرفت‌های حاصله توسط کودکان به آنها حقوق پرداخت می شد، کاهش دهد رقابت جبرانی از بسیاری جهات شبیه معادل سازی جبرانی است. هرچند، معادل سازی جبرانی ناشی از پیش بینی مشکلاتی است که از ناحیه گروه‌هایی که مداخله‌های کمتر مطلوبی دریافت کرده‌اند پیش می آید در حالیکه رقابت جبرانی ناشی از طریقی است که اعضای این گروه‌ها نسبت به مداخله‌ها یا عدم مداخله‌ها واکنش نشان می دهند.

13- از دست دادن روحیه ناشی از آزرده‌گی توسط پاسخ دهندگانی که مداخله‌هایی کمتر مطلوب دریافت کرده‌اند¹. هنگامی که یک آزمایش مزاحمتی² باشد، واکنش گروه گواه علاوه بر همراه بودن با رقابت جبرانی می تواند توأم با آزرده‌گی و از دست دادن روحیه هم باشد. این امر بدان علت است که آزمودنی‌های گروه گواه⁽⁺⁾ ممکن است با پایین آوردن بهره وری و منافع شرکت، انتقام بکشد. این موقعیت احتمال دارد که منجر به یک تفاوت پس آزمون بین گروه‌های گواه و آزمایشی شود و کاملاً خطاست که این تفاوت را به مداخله نسبت داد. مناسبتر این است که گروه گواه را گروه مداخله آزرده‌گی نامی د. (البته این پدیده منحصر به گروه‌های گواه نیست، بلکه می تواند هر جا که مداخله‌ها از لحاظ مطلوبیت تفاوت داشته و پاسخ دهندگان از این تفاوت آگاه باشند، روی بدهد).

تاریخ موضعی³. گماردن تصادفی افراد به گروه‌های آزمایشی و گواه منابع نامربوط همزمان را که در تهدید "تاریخ" فهرست شدند مهار می کند. این امر در دو موقعیت صحت دارد: اول اینکه گروه‌های آزمایشی و گواه به نحوی یکسان در این رویدادهای نامربوط سهیم باشند. دوم اینکه این رویدادها در جلسات گردآوری داده‌ها که خاص گروه‌های گواه یا آزمایشی است روی ندهند. هنگامی که داده‌ها از گروه‌ها گردآوری می شود، گماردن تصادفی تمام آزمودنی‌های گروه آزمایشی به یک جلسه گروهی بدین معنی خواهد بود که رویدادهایی ویژه که در آن جلسه رخ می دهند با مداخله آزمایشی همراه شده و ممکن است موجب اثرهایی باشند علاج این کار، هر جا که میسر باشد، این است که مداخله نسبت به افراد یا نسبت به بسیاری از گروه‌های کوچک اعمال شده و جلسات مربوط به گروه‌های آزمایشی و گواه از لحاظ مکان و زمان تصادفی گردند. در این رویه اخیر، درجات آزادی مناسب مبتنی بر تعداد گروه‌ها خواهد بود نه افراد. احتمال یک تفسیر مربوط به تاریخ موضعی در شبه آزمایش‌ها اغلب بیشتر از آزمایش‌های حقیقی است چه در مورد قبلی احتمال می

رود که مداخله‌ها با تفاوت‌های موجود در جمعیت پاسخ دهندگان آمیخته گردند. بدین ترتیب، مداخله‌های متفاوت با تمام تجارب بی‌همتا که هر گروه واجد است همراه خواهد گشت.

تخمین روایی درونی یک رابطه، فرایندی استقرائی است که در آن محقق باید منتقد سختگیر خود باشد و باید بنحوی اصولی بیندیشد که چگونه هر کدام از عوامل فوق ممکن است بر داده‌های او نفوذ داشته باشد آنگاه باید داده‌ها را بررسی کند تا دریابد که کدام‌یک از تهدیدهای مربوط را می‌توان رد کرد. هنگامی که همه آنها را بتوان رد کرد، نتیجه‌گیری‌هایی مطمئن درباره اینکه رابطه‌ای علی وجود دارد ممکن است به‌عمل آورد. هنگامی که همه تهدیدها را نتوان رد کرد، شاید به‌دلیل اینکه داده‌های مناسب در اختیار نیست یا به‌علت اینکه داده‌ها حاکی از آنند که تهدیدی خاص ممکن است در واقع عمل کرده باشد، آنگاه محقق باید چنین نتیجه‌گیری کند که یک رابطه نشان داده شده بین دو متغیر ممکن است علی باشد یا نباشد.

گاهی محقق مجبور خواهد بود چنان عمل کند که گوئی رابطه علی است چه تصمیماتی عملی باید اتخاذ گردد و امکان دارد که رابطه علی باشد. در دیگر مواقع، تفسیرهای جایگزین ممکن است به‌اندازه‌ای غیر موجه به‌نظر آیند که مورد مسامحه قرارگیرند و محقق مایل به رد آنها باشد. هرگاه تفسیر جایگزین براساس یافته‌هایی از یک سنت پژوهشی با تعداد زیادی از یافته‌های مربوط و باز انجام‌گرفته غیر محتمل به‌نظر آید می‌توان با اطمینان آن را رد کرد. با وجود این، اغلب دشوار است که توافق میان داوری بالا درباره موجه بودن یک تفسیر جایگزین ویژه به‌دست آورد. وانگهی، نظریه آزمایشان تأکید زیادی بر آزمودن پیش‌بینی‌هایی نظری و آنچه‌ان غیر موجه می‌گذارند که نه عقل سلیم و نه دیگر نظریه‌ها چنین پیش‌بینی‌هایی به‌عمل نمی‌آورند. در این مطلب اعترافی ضمنی وجود دارد مبنی بر اینکه تفسیر غیرموجه گاهی صحیح است و اینکه تفسیر جایگزین غیرموجه باید شک و تردید ما را درباره اینکه آیا یک رابطه علی است یا نه کاهش دهند نه اینکه آنها را زائل سازند.

تفاوت عمده بین آزمایش‌های حقیقی و شبه آزمایش‌ها مربوط به روایی درونی است. هنگامی که پاسخ دهندگان به‌طور تصادفی به گروه‌های گواه و آزمایشی گمارده می‌شوند، هر دو گروه به‌گونه‌ای یکسان تشکیل می‌گردند (مشکلات گزینش، رشد و نمو، یا تعامل گزینش و رشد و نمو زائل می‌شوند) هر آزمودنی شرایط آزمایشی و وسائل پژوهشی یکسانی را تجربه می‌کند (مشکلات آزمودن یا به‌کارگیری ابزار زائل می‌شوند). گزینش عمدی آنهایی که نمره‌های بالا یا پایین در هر آزمون دارند بجز تحت شرایطی که پاسخ دهندگان نخست بر مبنای مثلاً، نمره‌های پیش‌آزمون به‌صورت زوج درآمده‌اند و آنگاه به شرایط آزمایشی گمارده شده‌اند، به‌عمل نمی‌آید (مشکل بازگشت آماری

وجود ندارد). بدین ترتیب تصادفی کردن اکثر لیکن نه همه تهدیدهای مربوط به روایی درونی را زائل می کند.

در رابطه با طرح‌های شبه آزمایشی موقعیت کاملاً متفاوت است. به‌عوض اتکاء به تصادفی کردن برای رد اکثر تهدیدهای مربوط به روایی درونی، محقق باید ابتدا همه آنها را آشکار گرداند و آنگاه یکی یکی رد کند. بنابراین، وظیفه او پزحمت‌تر است. با وجود این، استنباط علی نهایی به استحکام موقعی نخواهد بود که یک آزمایش واقعی انجام می گیرد.

دکتر حسین شکرکن

سر دبیر نشریه بین المللی روانشناسی

- استفاده از مطالب با ذکر منبع بلامانع است.