Methodological and analytical considerations in behavioral intervention research

Paolo Ghisletta

Dept. of Psychology, Faculty of Psychology and Educational Sciences, University of Geneva
Swiss Distance Learning University
Swiss National Centre of Competence in Research LIVES - Overcoming vulnerability: Life course perspectives, Universities of Lausanne and Geneva

March 26, 2017

This workshop

Behavioral interventions to enhance cognition: Toward a consensus on methods

- ambitious
- controversial
- much needed!

Introduction (con't)

Different perspectives on "brain training"

A consensus on the brain training industry from the scientific community (October 20, 2014). Stanford Center on Longevity, CA., USA, and Max Planck Institute for Human Development, Berlin, Germany.

. . . The Stanford Center on Longevity and the Berlin Max Planck Institute for Human Development gathered many of the world's leading cognitive psychologists and neuroscientists to share their views about brain games and offer a consensus report to the public.. . .
. . . We object to the claim that brain games offer consumers a scientifically grounded avenue to reduce or reverse cognitive decline when there is no compelling scientific evidence to date that they do.

An open letter to the Stanford Center on Longevity (December, 2014). Cognitive Training Data (http://www.cognitivetrainingdata.org).

. . . A substantial and growing body of evidence shows that certain cognitive training regimens can significantly improve cognitive function, including in ways that generalize to everyday life. This includes some exercises now available commercially.

A comprehensive review on "brain training"

Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., & Stine-Morrow, E. A. L. (2016). Do "brain-training" programs work? *Psychological Science in the Public Interest, 17(3)*, 103-186. https://doi.org/10.1177/1529100616661983

To date, the field has lacked a comprehensive review of the brain-training literature, one that examines both the quantity and the quality of the evidence according to a well-defined set of best practices. This article provides such a review, focusing exclusively on the use of cognitive tasks or games as a means to enhance performance on other tasks. We specify and justify a set of best practices for such brain-training interventions and then use those standards to evaluate all of the published peer-reviewed intervention studies cited on the websites of leading brain-training companies listed on Cognitive Training Data.

Problems with intervention studies

Simons et al., 2014, p. 171

- Severe problems (that preclude any conclusions about the causal efficacy)
  - No pretest baseline
  - No control group
  - Lack of random assignment to conditions
- Substantial problems (ambiguous or inconclusive evidence)
  - passive control group
  - lack of preregistration
  - scattershot publishing without full documentation
  - small N
  - contingent analyses
  - subgroup analyses
- Potential problems (that were not completely addressed)
  - active but unmatched control group
  - inadequate preregistration
  - departures from preregistration
  - lack of blinding when using subjective measures

Preregistration
Simons et al., 2014, p. 171

Mandatory in US-funded clinical trials in medicine since 2000.

Preregistration requires:

- nature of the experimental intervention
- all of the conditions
- all outcome measures
- complete analysis plan

A study by the National Heart, Lung, and Blood Institute found that between 1970 and 2012, 57% of studies from before 2000 reported positive effects, vs. only 8% thereafter (Kaplan & Irvin, 2015).

Recommendations for analysis

In Simons et al. (2014, pp. 165–168)

- directly compare improvements (*a difference in statistical significance is not the same as a significant difference*)
- control for multiple comparisons
- provide data and statistics
- focus on measuring and estimating evidence (e.g., Bayes factors vs. NHST)
- avoid duplicate or scattershot publications

Not in Simons et al. (2014)

- consider trial-by-trial analyses (vs. aggregate)
- task-specific vs. ability-general analyses and inferences
- beware of outliers

Trial-by-trial analysis

Sajda, P., Rousselet, G. A., & Pernet, C. R. (2011). Single-trial analyses of behavioural and neuroimaging data in perception and decision-making. *Frontiers in Psychology, 322, (2).* 10.3389/fpsyg.2011.00322

Methods that consider the variance within subjects.

By using the information contained in the variance of individual trials, the single-trial approach goes beyond the activity of the average brain: it reveals the specificity of information processing in individual subjects, across tasks and stimulus space, revealing both inter-individual commonalties and differences.

Advantages of TBT analysis

- avoid ecological fallacy (Robinson, 1950)
- test rather than assume parallelism of learning trajectories across individuals (Macdonald, Stigsdotter-Neely, Derwinger, & Bäckman, 2006)
- test how within-person change parameters vary between individuals as a function of specific characteristics
- provide more sensitive statistical tests across units of analysis, with fewer assumptions
- may more adequately test for possible transfer effects (Green, Strobach, & Schubert, 2014)
- modern methods accommodate for missing data and incomplete training schedules

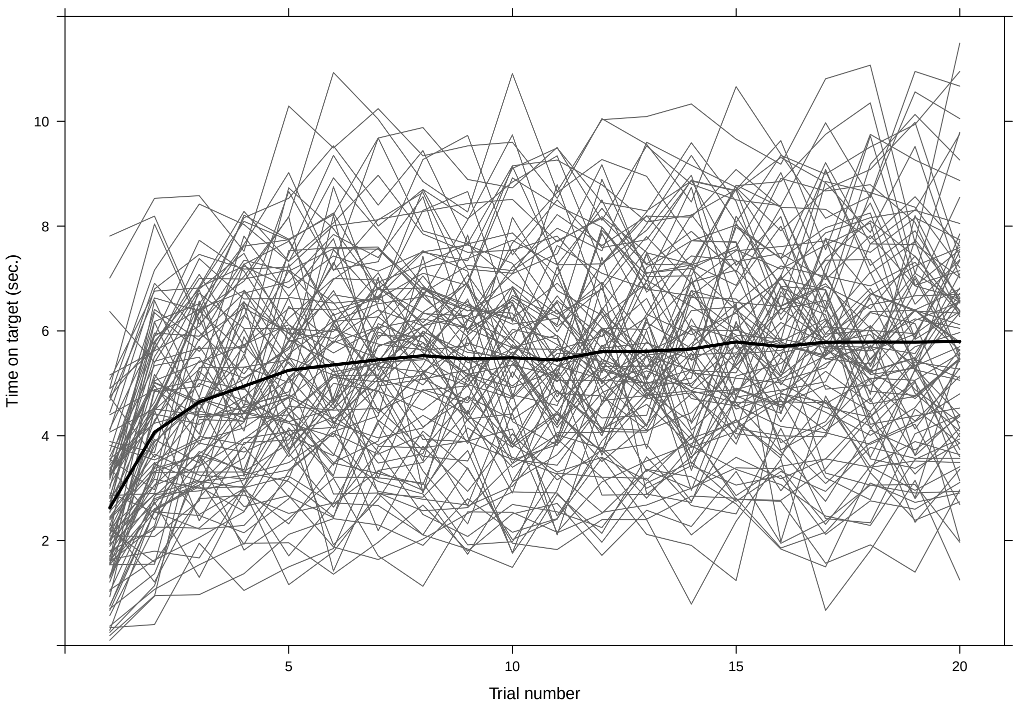
Example of TBT analysis

Ghisletta, P., Kennedy, K. M., Rodrigue, K. M., Lindenberger, U., & Raz, N. (2010). Adult age differences and the role of cognitive resources in perceptual-motor skill acquisition: Application of a multilevel negative exponential model. *Journal of Gerontology: Psychological Sciences, 65*, 163-173.

Ghisletta, P., Cantoni, E., & Jacot, N. (2015). Nonlinear Growth Curve Models. In: Stemmer, M., von Eye, A., & Wiedermann, W. (Eds.), *Dependent data in social sciences research: Forms, issues, and methods of analysis*. Berlin, Germany: Springer.

Learning perceptual-motor skill in $N = 102$ adults (19–80 years).

Rotor pursuit data



Nonlinear mixed-effects models

The NMEM can be defined as (Davidian & Giltinan, 1995)

$$\mathbf{y}_i = \mathbf{f}_i(\boldsymbol{\beta}_i) + \mathbf{e}_i \tag{1}$$

$$\boldsymbol{\beta}_i = \mathbf{d}(\mathbf{a}_i, \boldsymbol{\beta}, \mathbf{b}_i) \tag{2}$$

- $\mathbf{y}_i$ is the $(n_i \times 1)$ data vector for the ith individual $(N = \sum_{i=1}^{m} n_i)$
- $\mathbf{f}_i = [f(x_{i1}, \beta_i), \ldots, f(x_{in_i}, \beta_i)]'$
- $\mathbf{x}_i$ is the vector of predictors
- $\mathbf{e}_i$ is vector of random errors
- $\mathbf{d}$ is a function of $\mathbf{a}_i$, $\boldsymbol{\beta}$, and $\mathbf{b}_i$
- $\mathbf{a}_i$ is a vector of individual characteristics
- $\boldsymbol{\beta}$ is a vector of fixed effects
- $\mathbf{b}_i$ is a vector of random effects

3-parameter exponential model

$$y_{ij} = \beta_i - (\beta_i - \alpha_i)\exp(-(t_{ij} - 1)\gamma_i) + e_{ij} \tag{3}$$

- α_i is the initial performance at time $t_{ij} = 1$
- β_i the final performance at $t_{ij} = 20$ (final asymptote)
- γ_i the rate of change (representing learning speed)
- $\alpha_i = \alpha + U_{1i}$, $U_{1i} \sim \mathcal{N}(0, \sigma_1^2)$
- $\beta_i = \beta + U_{2i}$, $U_{2i} \sim \mathcal{N}(0, \sigma_2^2)$
- $\gamma_i = \gamma + U_{3i}$, $U_{3i} \sim \mathcal{N}(0, \sigma_3^2)$
- $e_{ij} \sim \mathcal{N}(0, \sigma_e^2)$

Alternative models

Logistic

$$y_{ij} = \frac{\alpha_i \beta_i}{\alpha_i + (\beta_i - \alpha_i)\exp(-(t_{ij} - 1)\gamma_i)} + e_{ij}. \qquad (4)$$

Gompertz

$$y_{ij} = \beta_i \exp\left(\ln\left(\frac{\alpha_i}{\beta_i}\right)\exp(-(t_{ij} - 1)\gamma_i)\right) + e_{ij}. \qquad (5)$$

Chapman-Richard

$$y_{ij} = \beta_i(1 - \exp(-\gamma_i t_{ij}))^{\delta_i} + e_{ij}. \qquad (6)$$

von Bertalanffy

$$y_{ij} = (\beta_i^{\frac{1}{\delta_i}} - \exp(-\gamma_i t_{ij}))^{\delta_i} + e_{ij}. \qquad (7)$$

Schnute

$$y_{ij} = \left((\alpha_i^{\gamma_i} + (\beta_i^{\gamma_i} - \alpha_i^{\gamma_i}))\frac{1 - \exp(-\delta_i(t_{ij} - t_1))}{1 - \exp(-\delta_i(t_2 - t_1))}\right)^{\frac{1}{\gamma_i}} + e_{ij}. \qquad (8)$$

Model comparison

Table 1 Total number of parameters (p; fixed and random effects), -2 Log-Likelihood (-2LL) and Bayesian Information Criterion (BIC) values, and parameter estimates for each nonlinear function (only fixed effects and variances of random effects are shown for each parameter of the functions in Eqs. (6)–(11))

Function	p	-2LL	BIC	α	β	γ	δ	σ_1^2	σ_2^2	σ_3^2	σ_4^2	σ_e^2
Exponential	10	5752	5798	2.82	5.75	0.39	—	1.57	3.11	0.08	—	0.82
Logistic	10	5777	5823	2.99	5.73	0.53	—	1.57	3.07	0.10	—	0.83
Gompertz	10	5765	5810	2.91	5.74	0.46	—	1.57	3.09	0.09	—	0.83
Chapman-Richard	10	5738	5784	—	5.76	0.33	0.56	—	3.19	0.11	0.14	0.81
von Bertalanffy	10	5758	5804	—	5.74	0.42	2.08	—	3.10	0.09	0.16[a]	0.82
Schnute	15	5703	5772	2.62	5.75	4.24	0.23	1.40	3.27	35.44	0.26	0.79

[a] A non-significant parameter at the 5 % level

Effect size estimates

Table 2 Percentage of variance explained of each random effect by age and spatial abilities

Function	U_{1i}	U_{2i}	U_{3i}	U_{4i}
Exponential	18.53	20.71	4.26	–
Logistic	16.99	21.16	1.72	–
Gompertz	17.44	20.99	2.69	–
Chapman-Richard	–	19.71	0.88	−11.65
von Bertalanffy	–	20.74	−0.30	−23.17

- Spatial abilities predicted positively initial and final performance
- Age predicted negatively final performance
- Age predicted positively rate of learning (less than 5%)

Task-specific vs. ability-general analyses and inferences

The notion that performance on a single task cannot stand in for an entire ability is a cornerstone of scientific psychology. (Stanford and Berlin consensus, Oct. 2014).

- Tasks are highly specific
- We hope strong results to generalize across multiple tasks - and to ddo so beyond their specificities

Latent variable approach

Classical test theory (Gulliksen, 1950):

$$x_i = T_i + e_i \quad (9)$$

Thanks to Spearman (1904):

$$x_{1,i} = \lambda_1 f_{1,i} + u_{1,i} \quad (10)$$

$$x_{2,i} = \lambda_2 f_{1,i} + u_{2,i} \quad (11)$$

$$x_{3,i} = \lambda_3 f_{1,i} + u_{3,i} \quad (12)$$

Beware of the "nominalistic fallacy" (Cliff, 1983).

A task's characteristic may be due to u_k rather than to f. Intervention may influence f and/or u. Thus, studying intervention's effects only at the latent level may overshadow tasks' unique effects.

Exemple of task specific analyses

Ghisletta, P., Rabbitt, P., Lunn, M., & Lindenberger, U. (2012). Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. (2012) *Intelligence, 40*, 260-268.

Table 1
Longitudinal design of the Manchester Study for cognitive variables.

Test battery A					Test battery B				
Domain	λ	R^2	Task	Abbreviation	Domain	λ	R^2	Task	Abbreviation
Gf	.91	.83	Heim Intelligence Test 1	AH4-1	Gf	.82	.67	Culture Fair Test	CFT
Gf	.84	.71	Heim Intelligence Test 2	AH4-2	Gc	.81	.65	WAIS-R Vocabulary	WAISV
Gc	.88	.77	Raven Mill Hill Voc. A	MH-A	Speed	.69	.48	Visual Search	VS
Gc	.90	.81	Raven Mill Hill Voc. B	MH-B	Speed	.90	.81	Alphabet Coding Task	ACT
Memory	.72	.52	Verbal Free Recall	VFR	Speed	.72	.53	Semantic Reasoning	SR
Memory	.80	.65	Cumulative Verbal Recall	CVR	Memory	.74	.55	Imm. Verbal Free Recall	IVFR
Memory	.41	.17	Picture Recognition	PR	Memory	.79	.62	Delayed Verbal Recall	DVR
					VSMem	.69	.48	Shape Locations	ShL
					VSMem	.91	.82	Spatial Locations	SpL
					VSMem	.98	.95	Shape + Spatial Locations	ShSpL
					Memory	.65	.43	Propositions about People	PaP
					Memory	.67	.44	Memory Objects	MO
					Memory	.64	.41	Memory Objects + Position	MOP

Measurement occasion	1	2	3	4	1	2	3	4
N	5926	3771	2125	990	4258	2417	1169	507
Mean age (SD)	64.90 (7.45)	68.53 (6.87)	72.23 (6.41)	75.29 (5.91)	67.69 (7.05)	72.50 (6.38)	75.81 (6.33)	77.10 (5.61)
[min-max]	[43–93]	[49–92]	[54–93]	[62–97]	[42–96]	[47–95]	[51–96]	[54–95]

Note. Gf = fluid intelligence. Gc = crystallized intelligence. VSMem = visuo-spatial memory. Voc. = Vocabulary. Imm. = Immediate. λ = standardized factor loading from confirmatory factor analysis. R^2 = Task variance accounted for by domain-specific factor in confirmatory factor analysis.

Age, demographic, and retest effects

Table 2
Parameter estimates (and standard errors) of the multivariate multilevel model.

Task	Fixed effects															Random effects			
	I	lS	qS	lr	r_1	r_2	r_3	city	sex	soc_1	soc_2	soc_4	soc_5	soc_6	$soc_{M/U}$	I	lS	ε	lr
AH4-1	48.48	−.49	−.013	–	2.46	3.53	4.37	−.94	−.50	7.84	4.89	−.91	−6.38	−7.59	−4.46	61.51	.06	10.55	–
	(.35)	(.01)	(.001)		(.09)	(.15)	(.24)	(.23)	(.25)	(.58)	(.30)	(.33)	(.45)	(1.21)	(.50)	(1.33)	(.01)	(.21)	
AH4-2	49.46	−.59	−.013	–	3.14	5.52	7.97	.-98	−2.59	8.32	4.10	−.45	−5.03	−6.58	−3.69	56.79	.06	14.33	–
	(.35)	(.01)	(.001)		(.10)	(.16)	(.25)	(.23)	(.25)	(.57)	(.29)	(.32)	(.44)	(1.18)	(.50)	(1.29)	(.01)	(.29)	
MH-A	50.10	−.07	−.006	–	.68	−.24	.97	−2.04	−.60	8.95	6.18	−.28	−6.15	−8.36	−3.17	75.53	.02	17.14	–
	(.38)	(.02)	(.001)		(.12)	(.18)	(.28)	(.25)	(.27)	(.62)	(.32)	(.36)	(.49)	(1.30)	(.53)	(1.65)	(.01)	(.33)	
MH-B	49.82	−.08	−.005	–	.49	.04	1.94	−2.65	.12	8.65	5.99	−.46	−5.83	−7.65	−2.66	55.53	–	24.49	–
	(.35)	(.01)	(.001)		(.12)	(.18)	(.27)	(.23)	(.25)	(.57)	(.29)	(.33)	(.45)	(1.20)	(.51)	(1.29)		(.42)	
VFR	46.04	−.41	−.006	–	−.17	1.23	−.34	−.94	2.34	4.03	3.57	.33	−2.20	−3.54	−1.85	41.43	.05	38.99	–
	(.37)	(.02)	(.001)		(.17)	(.23)	(.32)	(.24)	(.26)	(.58)	(.30)	(.34)	(.47)	(1.33)	(.70)	(1.39)	(.01)	(.79)	
CVR	46.22	−.57	−.016	–	.32	3.30	−.23	−2.08	3.42	3.90	3.22	.15	−3.69	−6.91	−4.29	55.97	.13	28.24	–
	(.37)	(.02)	(.001)		(.14)	(.22)	(.33)	(.25)	(.26)	(.59)	(.30)	(.34)	(.48)	(1.33)	(.68)	(1.97)	(.01)	(.57)	
PR	47.16	−.36	−.015	–	.72	3.78	3.15	−.43	2.61	1.63	1.97	.19	−1.88	−6.19	−2.07	44.24	.12	39.41	–
	(.38)	(.02)	(.001)		(.16)	(.23)	(.34)	(.25)	(.26)	(.59)	(.31)	(.35)	(.48)	(1.35)	(.61)	(1.47)	(.01)	(.79)	
CFT	5122	−.60	−.011	–	2.19	3.41	5.10	−3.86	−1.56	6.11	3.57	−.48	−5.26	−4.95	−3.69	49.62	.06	18.77	–
	(.40)	(.02)	(.001)		(.16)	(.25)	(.35)	(.26)	(.28)	(.60)	(.33)	(.37)	(.53)	(1.53)	(.70)	(1.41)	(.01)	(.49)	
WAISV	48.84	−.17	−.005	–	5.04	1.96	4.07	−1.14	−.15	7.44	5.71	−.28	−6.18	−7.00	−5.01	58.28	–	21.68	–
	(.42)	(.02)	(.001)		(.17)	(.25)	(.34)	(.28)	(.30)	(.64)	(.35)	(.39)	(.55)	(1.61)	(.75)	(1.57)		(.48)	
VS	46.51	−.49	−.012	–	−.04	2.27	2.29	.68	2.65	4.29	2.13	−.17	−2.54	−.90	−3.33	64.80	.10	19.45	–
	(.47)	(.02)	(.001)		(.18)	(.28)	(.39)	(.31)	(.33)	(.70)	(.39)	(.43)	(.62)	(1.82)	(1.08)	(1.86)	(.01)	(.55)	
ACT	48.64	−.55	−.012	1.12	–	–	–	−2.05	2.50	3.74	2.37	−1.25	−7.13	−7.44	−5.04	64.93	.06	10.21	3.12
	(.43)	(.02)	(.001)	(.11)				(.28)	(.30)	(.65)	(.36)	(.40)	(.56)	(1.64)	(.77)	(1.93)	(.02)	(.30)	(.82)
SR	48.45	−.44	−.003	–	–	1.29	.88	−2.32	2.99	3.28	4.37	.03	−3.88	−2.86	−.86	61.47	.03	23.38	–
	(.61)	(.03)	(.002)			(.23)	(.35)	(.28)	(.41)	(.84)	(.48)	(.54)	(.84)	(2.21)	(1.58)	(2.64)	(.01)	(.89)	
IVFR	49.02	−.42	−.007	–	.36	1.28	1.85	−4.99	2.81	3.18	2.65	.35	−2.38	−3.04	−1.60	36.87	.08	49.42	–
	(.44)	(.02)	(.001)		(.23)	(.33)	(.47)	(.29)	(.30)	(.64)	(.36)	(.40)	(.58)	(1.68)	(1.00)	(1.68)	(.02)	(1.22)	
DVR	47.29	−.45	−.006	.93	–	–	–	−3.24	3.21	4.30	2.93	.93	−1.03	−1.98	−2.43	48.63	.03	37.21	2.73
	(.45)	(.02)	(.001)	(.14)				(.30)	(.32)	(.67)	(.37)	(.42)	(.60)	(1.75)	(1.04)	(2.35)	(.01)	(1.0[illegible])	(.26)

Generalizable findings

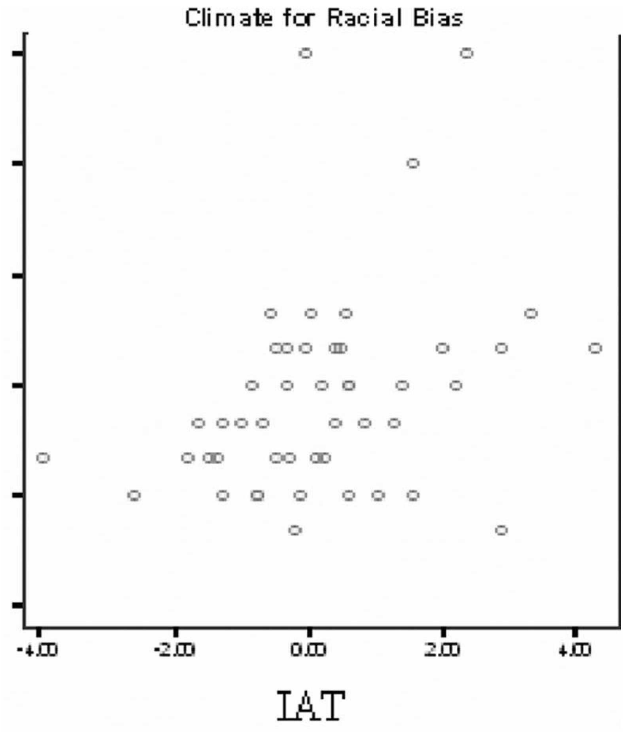
Table 3
Factor loadings of the two-factor exploratory analysis on the covariance matrix of variables' intercepts and slopes (with Promax rotation; the two factors correlate .19).

		Factor1	Factor2
Intercepts	AH4-1	−0.12	0.88
	AH4-2	0.01	0.77
	MH-A	−0.15	0.72
	MH-B	−0.19	0.75
	VFR	0.02	0.69
	CVR	0.18	0.75
	PR	0.04	0.45
	CFT	−0.05	0.83
	WAISV	−0.17	0.80
	VS	0.10	0.50
	ACT	0.03	0.77
	SR	0.11	0.61
	IVFR	0.11	0.68
	DVR	0.03	0.68
	ShL	−0.20	0.69
	SpL	0.20	0.70
	ShSpL	0.16	0.72
	PaP	0.02	0.66
	MO	0.17	0.64
	MOP	0.14	0.69
Slopes	AH4-1	0.75	0.09
	AH4-2	0.90	0.01
	MH-A	0.51	0.41
	VFR	0.84	−0.36
	CVR	1.02	0.00
	PR	0.79	−0.06
	CFT	0.84	0.23
	VS	0.68	−0.06
	ACT	0.84	0.06
	SR	0.73	0.05
	IVFR	1.03	0.10
	DVR	1.10	−0.09
	SpL	0.38	0.31
	ShSpL	0.47	0.28
	PaP	0.88	−0.21
	MO	0.84	0.09
	MOP	0.85	−0.09
Linear retest	ACT	−0.04	0.06
	DVR	0.24	0.13

Note. First block denotes Intercepts, the second Slopes, the third linear retest effects (only present for ACT and DVR).

Consider using robust statistics

Re-analysis of Ziegert and Hanges (2005) by Blanton et al. (2009)



In sum

- Greater methodological rigor
- More openness
- Advanced analytical tools

Methodological and analytical considerations in behavioral intervention research

Paolo Ghisletta

Dept. of Psychology, Faculty of Psychology and Educational Sciences, University of Geneva
Swiss Distance Learning University
Swiss National Centre of Competence in Research LIVES - Overcoming vulnerability: Life course perspectives, Universities of Lausanne and Geneva

March 26, 2017