



Does mode really matter?

Längsschnittliche Untersuchung von
Moduswechseln im GESIS Panel

Hintergrund

- Zunehmender Umstieg mixed-mode Befragungen (De Leeuw, 2005)
- Frage: Ändern sich Antwortverhalten bzw. Messtrukturen durch den Wechsel des Erhebungsmodus?
- Ansatz: Messinvarianzanalysen zwischen den Erhebungsmodi (Hox et al., 2015)
- Skalare Invarianz als Voraussetzung für aussagekräftige Mittelwertvergleiche (Leitgoeb et al., 2022)
- Bisherige Analysen ohne Berücksichtigung von Moduswechselnden im Längsschnitt

Datenbasis: GESIS Panel

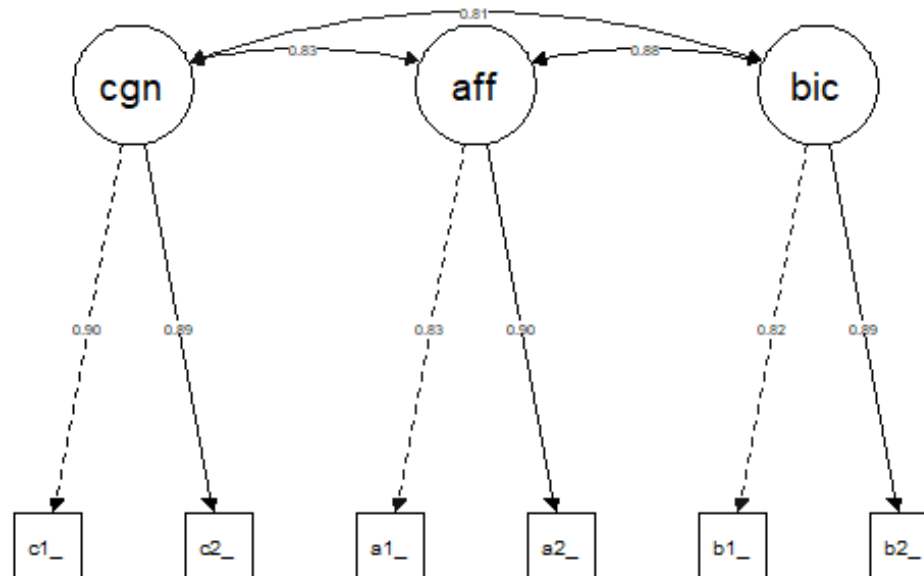
- Teilnahme am GESIS Panel sowohl offline als auch online möglich
- Push-to-Web Experiment 2018 → Wechsel von offline zu online möglich (Bretschi et al., 2023)
- Drei Gruppen vergleichbar: offline only / online only / mode switchers

Beispiel: Geflüchtetenindex

- 6 Items auf 3 Dimensionen (Kotzur et al., 2022):
 - Kognitive Ebene:
 - c1_: Stellen Flüchtlinge eher eine Bedrohung oder eine Bereicherung für die Werte in Deutschland dar?
 - c2_: Stellen Flüchtlinge eher eine Bedrohung oder eine Bereicherung für den Wohlstand in Deutschland dar?
 - Affektive Ebene:
 - a1_: Wie stark fühlen Sie mit Flüchtlingen mit?
 - a2_: Wie sympathisch sind Ihnen Flüchtlinge?
 - Behaviorale-intentionale Ebene:
 - b1_: Ich kann mir vorstellen mit Flüchtlingen zusammenzuarbeiten.
 - b2_: Ich hätte nichts gegen Flüchtlinge in meinem Wohnumfeld.
- Antwortoptionen: 5-Punkte Likert-Skala (quasi-metrisch)

Beispiel: Geflüchtetenindex

- 6 Items auf 3 Dimensionen: Kognitive, affektive und behavioral-intentionale Ebene (Kotzur et al., 2022)



Beispiel: Geflüchtetenindex

- 6 Items auf 3 Dimensionen: Kognitive, affektive und intentional-behaviorale Ebene
- Antwortoptionen: 5-Punkte Likert-Skala (quasi-metrisch)
- 8 Erhebungszeitpunkte von 2017 bis 2021
 - T1: Feb/Mär 2017 (Welle ea) n = 3868
 - T2: Jun/Jul 2017 (Welle ec) n = 3835
 - T3: Dez/Jan 2017/18 (Welle ef) n = 3825
 - T4: Dez/Jan 2018/19 (Welle ff) n = 3801
 - T5: Jun/Jul 2019 (Welle gc) n = 3694
 - T6: Jun/Jul 2020 (Welle hc) n = 3420
 - T7: Jun/Jul 2021 (Welle ic) n = 3294

Push-to-Web

Vorgehen

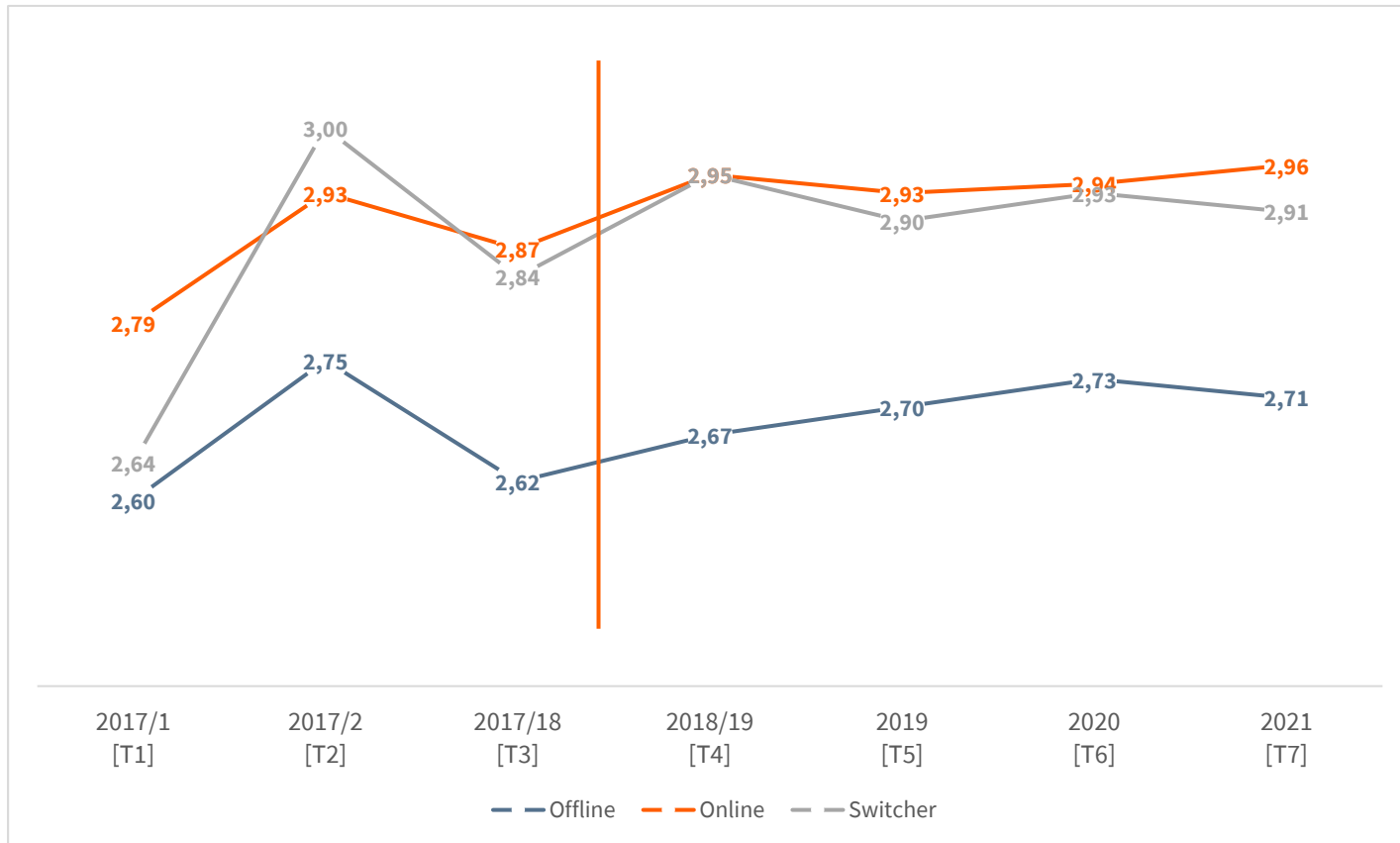
- **Deskription**

1. Fokus auf die Personengruppe, die min. einmal **vor** und einmal **nach** Push-to-Web teilgenommen haben ($N_{\text{mode}}=3995$)
2. Aufteilung in drei Gruppen:
 1. „offline only“; min. eine Erhebung T1-T3 offline & min. eine Erhebung T4-T7 offline ($n_{\text{offline}} = 1089$)
 2. „online only“; min. eine Erhebung T1-T3 online & min. eine Erhebung T4-T7 online ($n_{\text{online}} = 2717$)
 3. „switchers“; min. eine Erhebung T1-T3 offline & min. eine Erhebung T4-T7 online ($n_{\text{switch}}=189$)

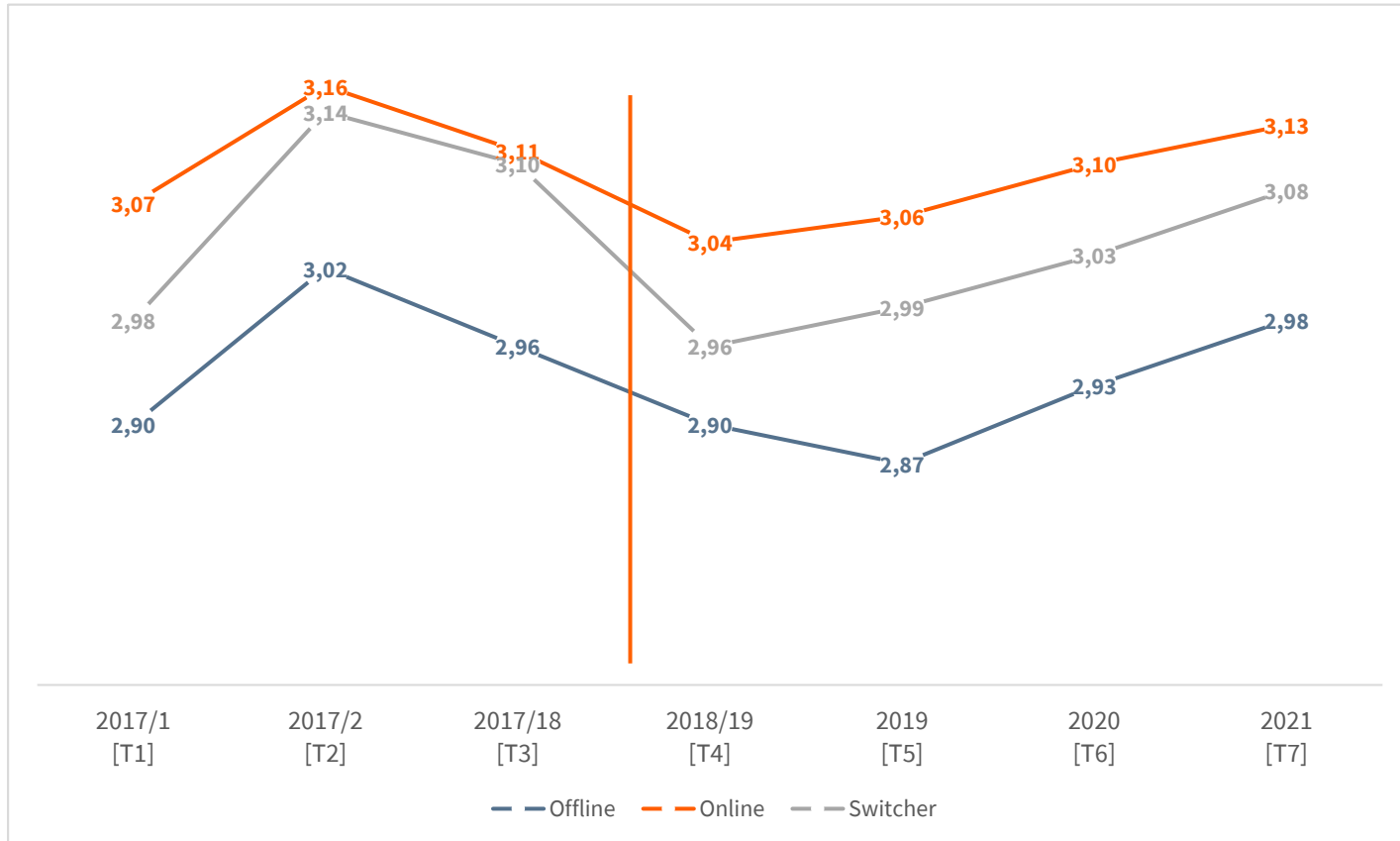
- **Messinvarianztestungen**

1. Separate Analyse der drei Gruppen zu jedem Erhebungszeitpunkt
2. Messinvarianz über der Gesamtgruppe über die Zeit ($N_{\text{mode}}=3995$)
3. Kombination der MI-Analysen der drei Gruppen und 7 Erhebungszeitpunkte

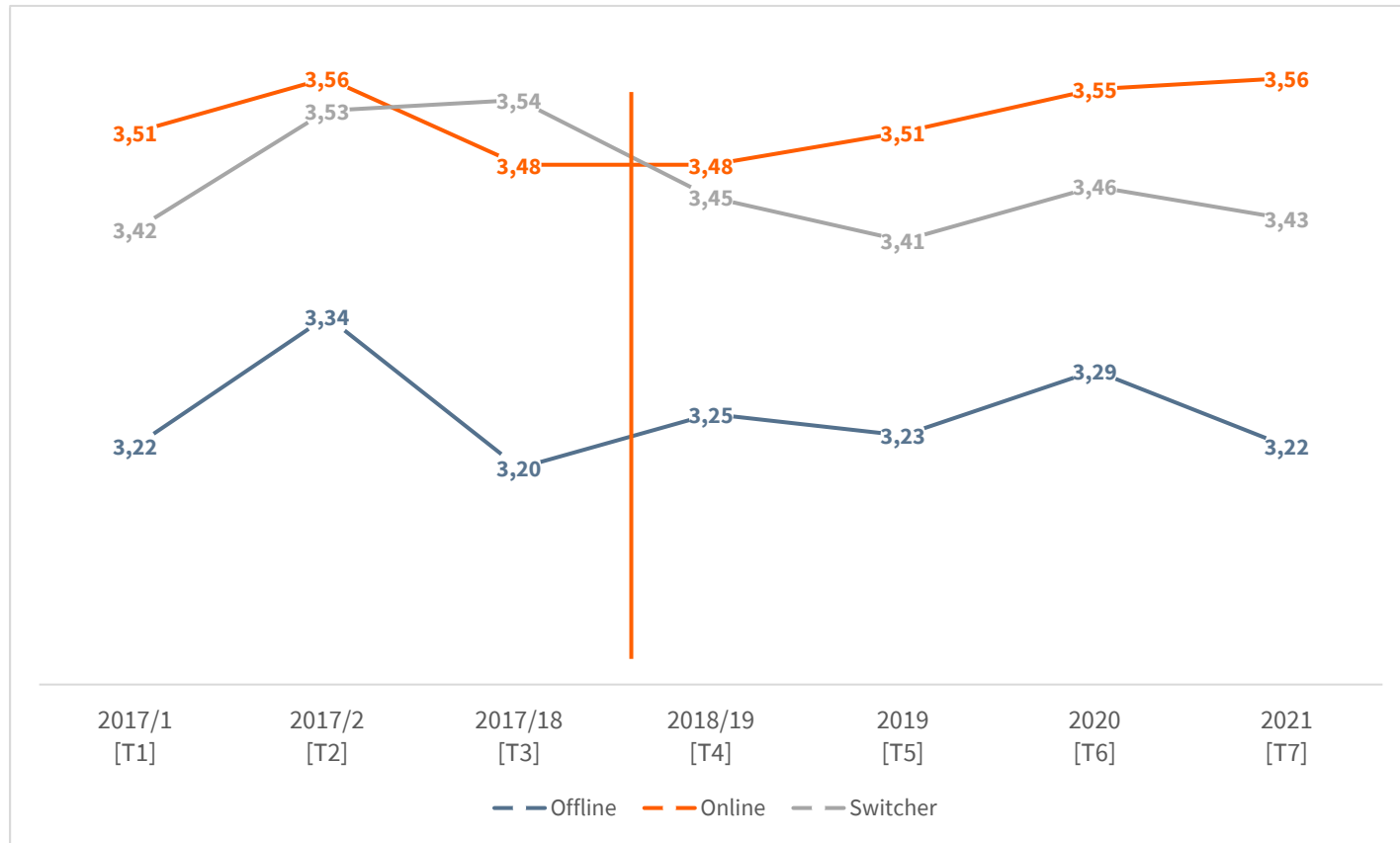
Deskription: Kognitive Dimension



Deskription: Affektive Dimension



Deskription: Intentional-Behaviorale Dimension



Messinvarianzanalyse

- Grundfrage: ist die Fragebogenstruktur zwischen Gruppen bzw. über die Zeit hinweg vergleichbar?
 - Vorteil: Vergleich der latenten Mittelwerte unter Berücksichtigung von Messfehlern
- Vorgehen: Testung der Struktur mithilfe von zunehmend restriktiveren, genesteten Modellen
 1. **Baseline:** Kann in jeder Gruppe / zu jedem Messzeitpunkt ein akzeptabler Modelfit erreicht werden?
 2. **Configural Invariance:** Ist die Messstruktur über die Gruppen / die Zeit hinweg akzeptabel?
 3. **Metric Invariance:** Gleichsetzung der Itemladungen zwischen den Gruppen
 4. **Scalar Invariance:** Gleichsetzung der Ladungen und Itemintercepts
 5. **Latent Mean Invariance:** Zusätzliche Gleichsetzung der latenten Mittelwerte

→ Modellvergleiche mittels Veränderungen der Fit-Indices

(Δ RMSEA \leq .015; Δ CFI $>$.01; Δ SRMR \leq .03

bzw. Δ RMSEA \leq .01; Δ CFI $>$.005; Δ SRMR \leq .025 bei ungleicher Gruppengröße)

Baseline

	Offline	Online	„Switcher“
T1			
T2			
...			
T7			

Messinvarianzanalyse

- Grundfrage: ist die Fragebogenstruktur zwischen Gruppen bzw. über die Zeit hinweg vergleichbar?
 - Vorteil: Vergleich der latenten Mittelwerte unter Berücksichtigung von Messfehlern
- Vorgehen: Testung der Struktur mithilfe von zunehmend restriktiveren, genesteten Modellen
 1. **Baseline:** Kann in jeder Gruppe / zu jedem Messzeitpunkt ein akzeptabler Modelfit erreicht werden?
 2. **Configural Invariance:** Ist die Messstruktur über die Gruppen / die Zeit hinweg akzeptabel?
 3. **Metric Invariance:** Gleichsetzung der Itemladungen zwischen den Gruppen
 4. **Scalar Invariance:** Gleichsetzung der Ladungen und Itemintercepts
 5. **Latent Mean Invariance:** Zusätzliche Gleichsetzung der latenten Mittelwerte

→ Modellvergleiche mittels Veränderungen der Fit-Indices

(Δ RMSEA \leq .015; Δ CFI $>$.01; Δ SRMR \leq .03

bzw. Δ RMSEA \leq .01; Δ CFI $>$.005; Δ SRMR \leq .025 bei ungleicher Gruppengröße)

Configural

	Offline	Online	„Switcher“
T1			
T2			
...			
T7			

Messinvarianzanalyse

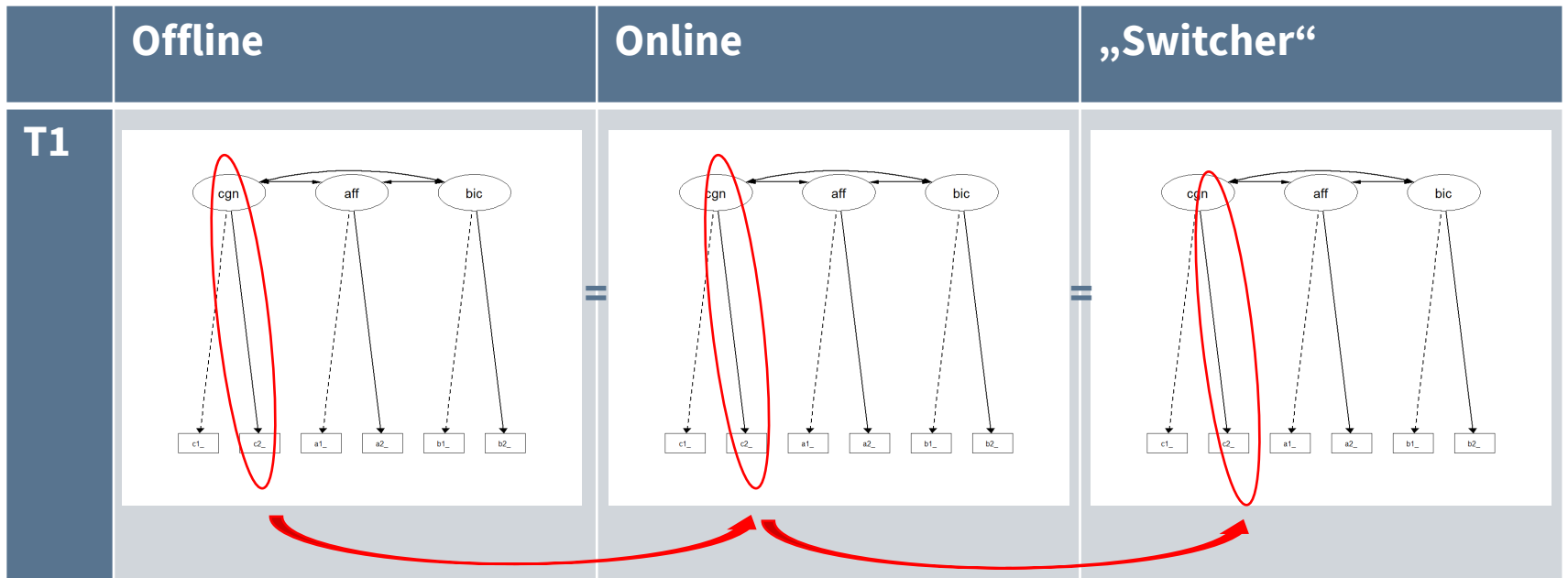
- Grundfrage: ist die Fragebogenstruktur zwischen Gruppen bzw. über die Zeit hinweg vergleichbar?
 - Vorteil: Vergleich der latenten Mittelwerte unter Berücksichtigung von Messfehlern
- Vorgehen: Testung der Struktur mithilfe von zunehmend restriktiveren, genesteten Modellen
 1. **Baseline:** Kann in jeder Gruppe / zu jedem Messzeitpunkt ein akzeptabler Modelfit erreicht werden?
 2. **Configural Invariance:** Ist die Messstruktur über die Gruppen / die Zeit hinweg akzeptabel?
 3. **Metric Invariance:** Gleichsetzung der Itemladungen zwischen den Gruppen
 4. **Scalar Invariance:** Gleichsetzung der Ladungen und Itemintercepts
 5. **Latent Mean Invariance:** Zusätzliche Gleichsetzung der latenten Mittelwerte

→ Modellvergleiche mittels Veränderungen der Fit-Indices

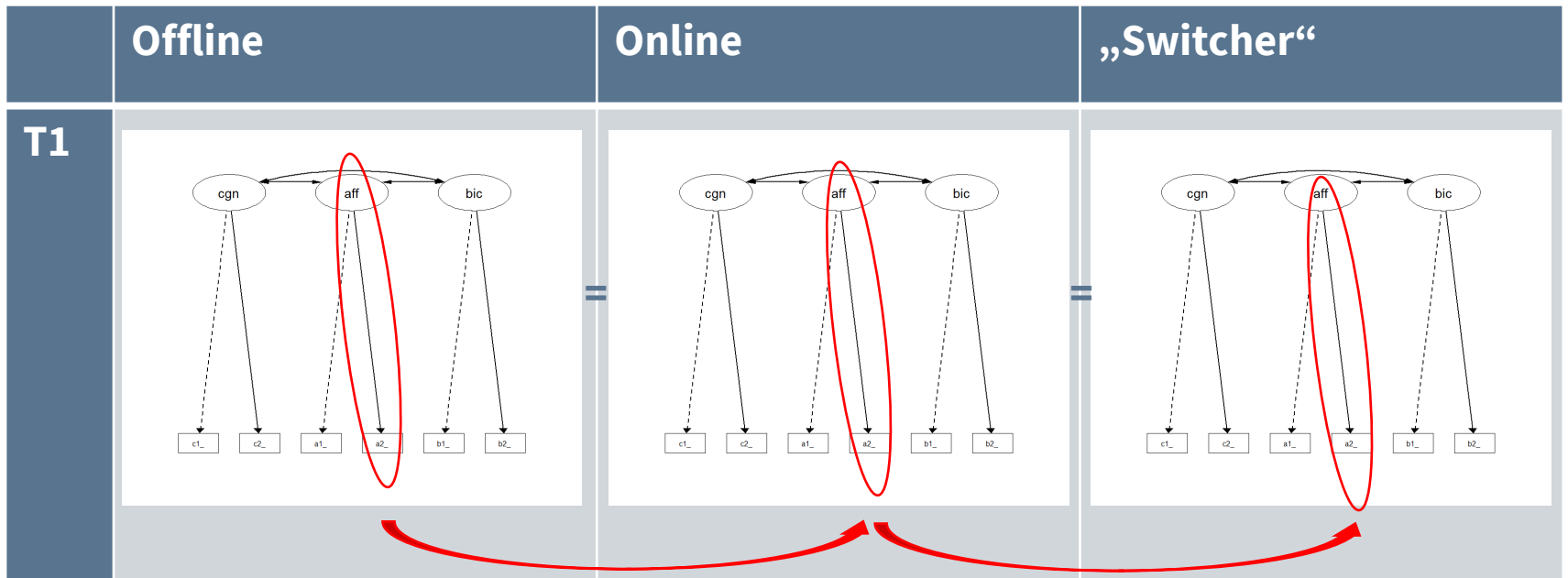
(Δ RMSEA \leq .015; Δ CFI $>$.01; Δ SRMR \leq .03

bzw. Δ RMSEA \leq .01; Δ CFI $>$.005; Δ SRMR \leq .025 bei ungleicher Gruppengröße)

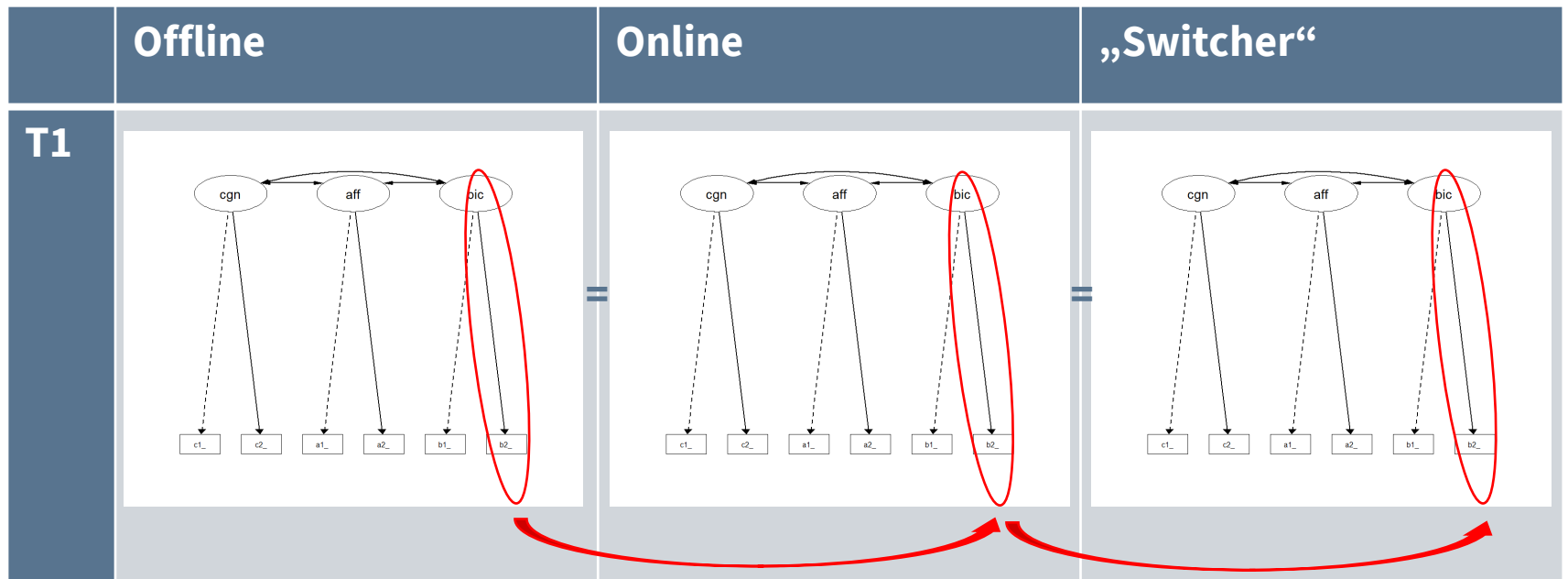
Metric invariance



Metric invariance



Metric invariance



Messinvarianzanalyse

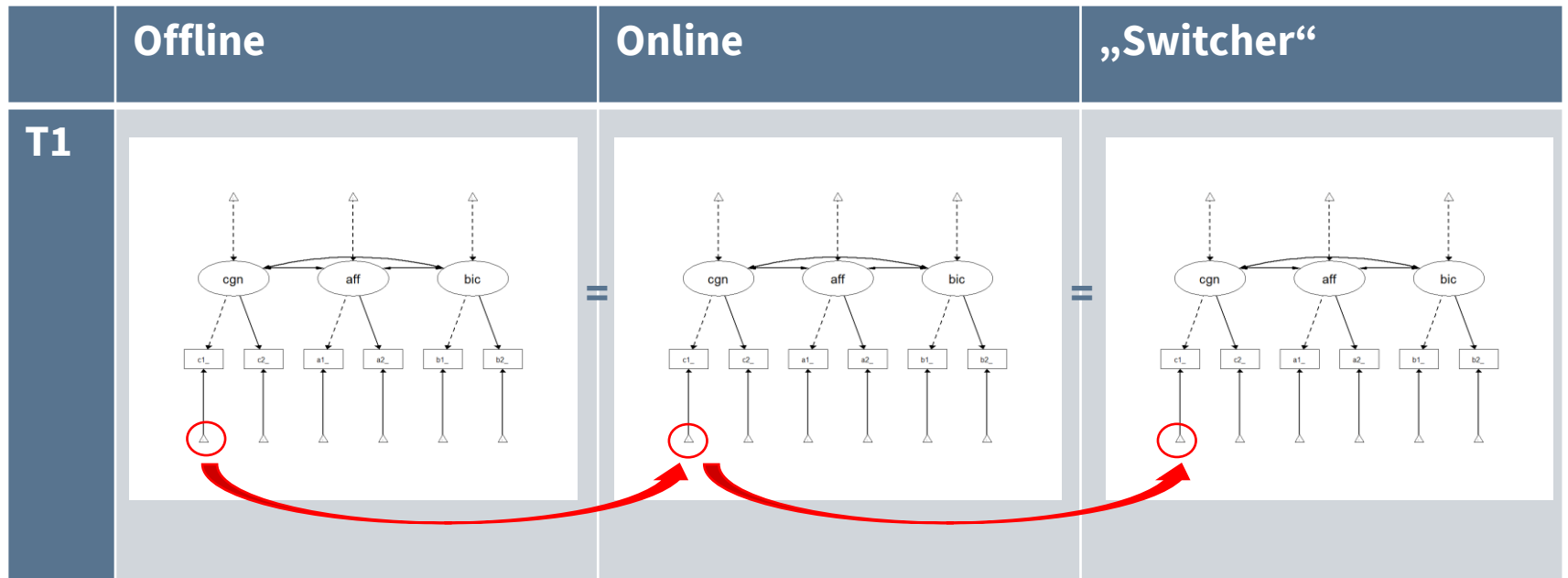
- Grundfrage: ist die Fragebogenstruktur zwischen Gruppen bzw. über die Zeit hinweg vergleichbar?
 - Vorteil: Vergleich der latenten Mittelwerte unter Berücksichtigung von Messfehlern
- Vorgehen: Testung der Struktur mithilfe von zunehmend restriktiveren, genesteten Modellen
 1. **Baseline:** Kann in jeder Gruppe / zu jedem Messzeitpunkt ein akzeptabler Modelfit erreicht werden?
 2. **Configural Invariance:** Ist die Messstruktur über die Gruppen / die Zeit hinweg akzeptabel?
 3. **Metric Invariance:** Gleichsetzung der Itemladungen zwischen den Gruppen
 4. **Scalar Invariance:** Gleichsetzung der Ladungen und Itemintercepts
 5. **Latent Mean Invariance:** Zusätzliche Gleichsetzung der latenten Mittelwerte

→ Modellvergleiche mittels Veränderungen der Fit-Indices

(Δ RMSEA \leq .015; Δ CFI $>$.01; Δ SRMR \leq .03

bzw. Δ RMSEA \leq .01; Δ CFI $>$.005; Δ SRMR \leq .025 bei ungleicher Gruppengröße)

Scalar invariance



Messinvarianzanalyse

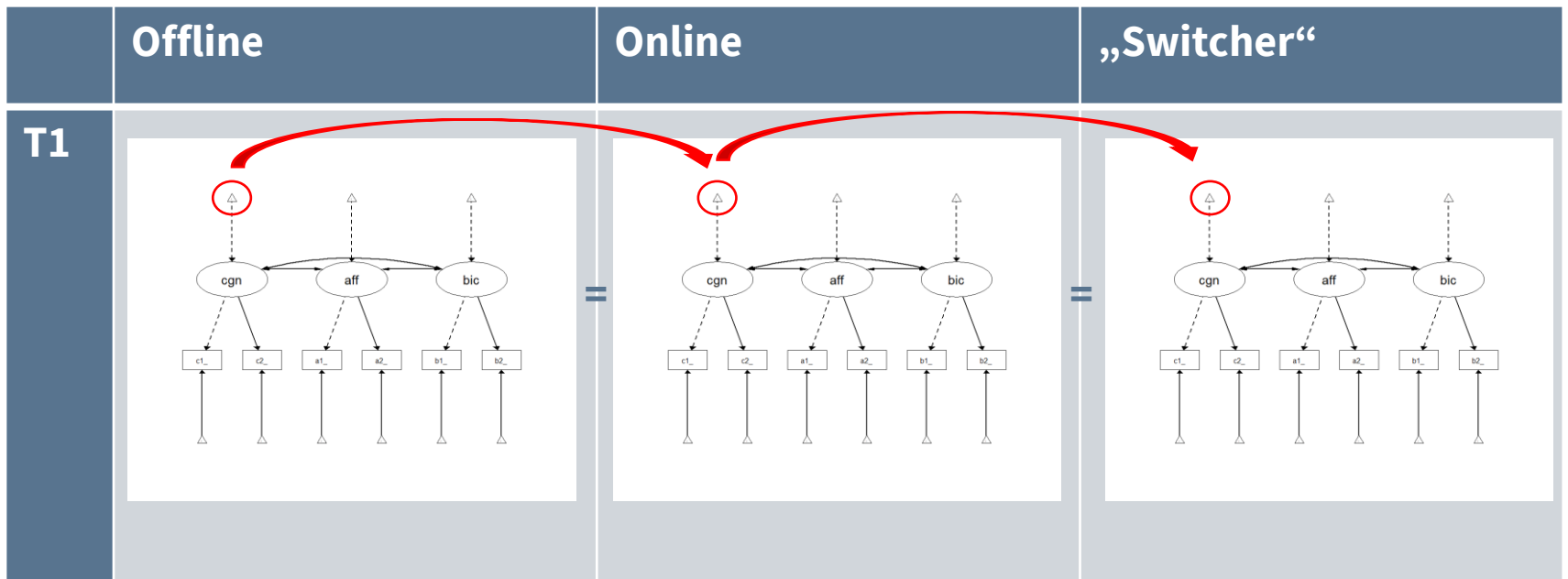
- Grundfrage: ist die Fragebogenstruktur zwischen Gruppen bzw. über die Zeit hinweg vergleichbar?
 - Vorteil: Vergleich der latenten Mittelwerte unter Berücksichtigung von Messfehlern
- Vorgehen: Testung der Struktur mithilfe von zunehmend restriktiveren, genesteten Modellen
 1. **Baseline:** Kann in jeder Gruppe / zu jedem Messzeitpunkt ein akzeptabler Modelfit erreicht werden?
 2. **Configural Invariance:** Ist die Messstruktur über die Gruppen / die Zeit hinweg akzeptabel?
 3. **Metric Invariance:** Gleichsetzung der Itemladungen zwischen den Gruppen
 4. **Scalar Invariance:** Gleichsetzung der Ladungen und Itemintercepts
 5. **Latent Mean Invariance:** Zusätzliche Gleichsetzung der latenten Mittelwerte

→ Modellvergleiche mittels Veränderungen der Fit-Indices

(Δ RMSEA \leq .015; Δ CFI $>$.01; Δ SRMR \leq .03

bzw. Δ RMSEA \leq .01; Δ CFI $>$.005; Δ SRMR \leq .025 bei ungleicher Gruppengröße)

Latent mean invariance



Messinvarianzanalyse

- Grundfrage: ist die Fragebogenstruktur zwischen Gruppen bzw. über die Zeit hinweg vergleichbar?
 - Vorteil: Vergleich der latenten Mittelwerte unter Berücksichtigung von Messfehlern
- Vorgehen: Testung der Struktur mithilfe von zunehmend restriktiveren, genesteten Modellen
 1. **Baseline:** Kann in jeder Gruppe / zu jedem Messzeitpunkt ein akzeptabler Modelfit erreicht werden?
 2. **Configural Invariance:** Ist die Messstruktur über die Gruppen / die Zeit hinweg akzeptabel?
 3. **Metric Invariance:** Gleichsetzung der Itemladungen zwischen den Gruppen
 4. **Scalar Invariance:** Gleichsetzung der Ladungen und Itemintercepts
 5. **Latent Mean Invariance:** Zusätzliche Gleichsetzung der latenten Mittelwerte

→ Modellvergleiche mittels Veränderungen der Fit-Indices (Chen, 2007)

(Δ RMSEA \leq .015; Δ CFI $>$.01; Δ SRMR \leq .03

bzw. Δ RMSEA \leq .01; Δ CFI $>$.005; Δ SRMR \leq .025 bei ungleicher Gruppengröße)

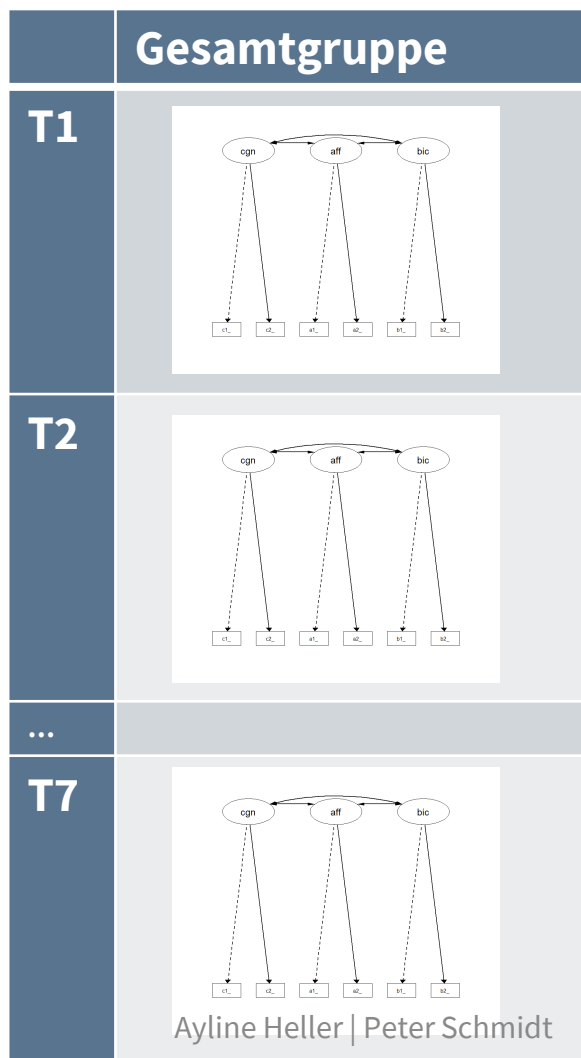
Messinvarianzanalyse

Reminder:
 $\Delta RMSEA \leq .01$; $\Delta CFI < .005$; $\Delta SRMR \leq .025$
 (ungleiche Gruppengröße)

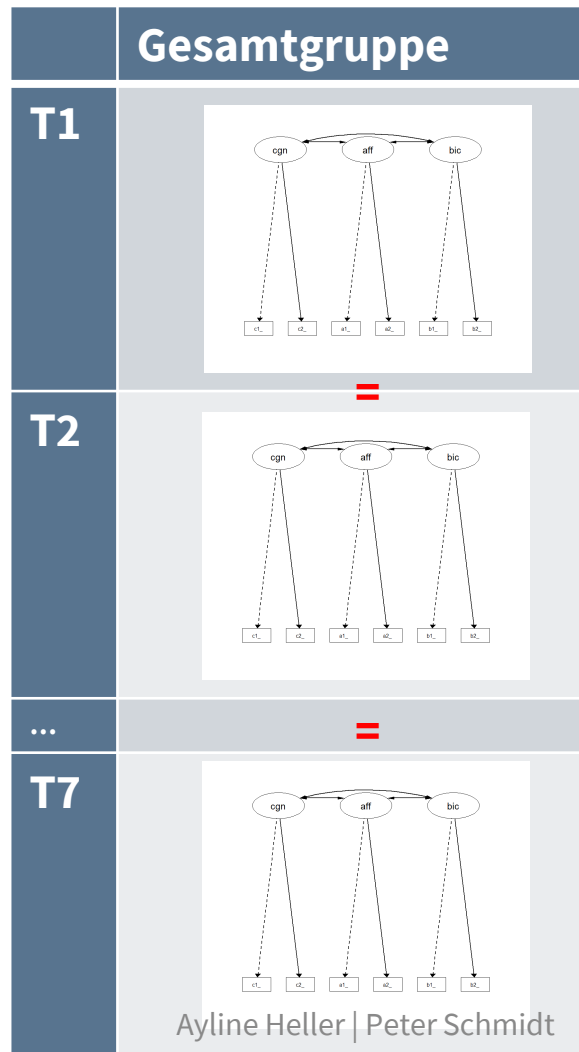
1. Zwischen den drei Gruppen zu jedem Erhebungszeitpunkt

	T1	T2	T3	T4	T5	T6	T7	T8
Skalare Invarianz	gegeben	$\Delta RMSEA = .011$	gegeben	-	$\Delta RMSEA = .012$	$\Delta RMSEA = .019$	$\Delta RMSEA = .011$	$\Delta RMSEA = .017$
Latente Mittelwerte	gleich	gleich	gleich	-	gleich	$\Delta SRMR = .028$	gleich	$\Delta SRMR = .028$
N	3868	3835	3825	3801	3801	3694	3420	3294
(offline/online / switchers)	(1046/2642 / 180)	(1058/2593/ 184)	(1021/2621/ 183)		(1014/2599/ 188)	(996/2514/ 184)	(883/2358/ 179)	(815/2305/ 174)

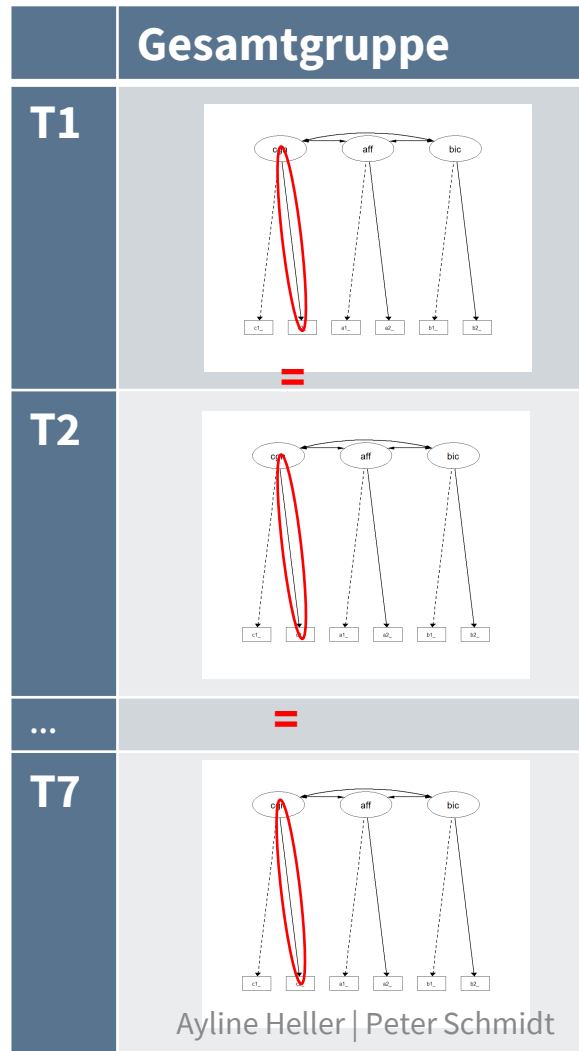
Baseline über die Zeit



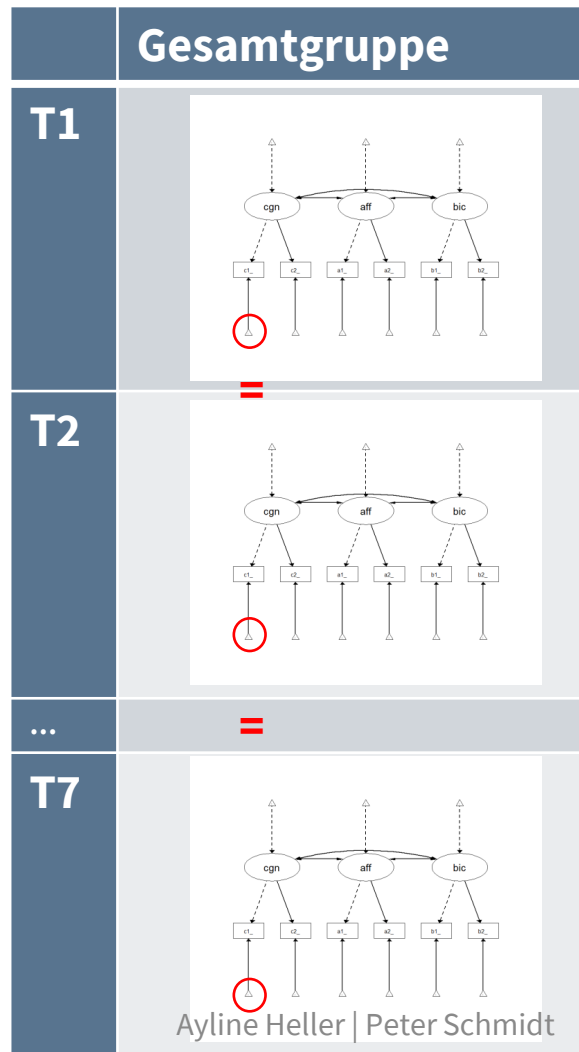
Configural über die Zeit



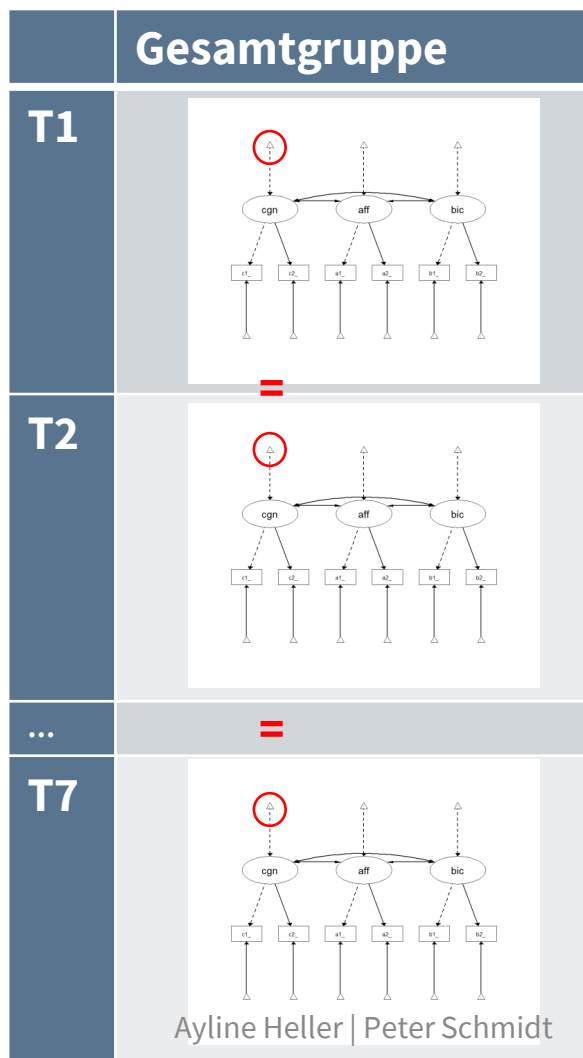
Metric über die Zeit



Scalar über die Zeit



Latent means über die Zeit



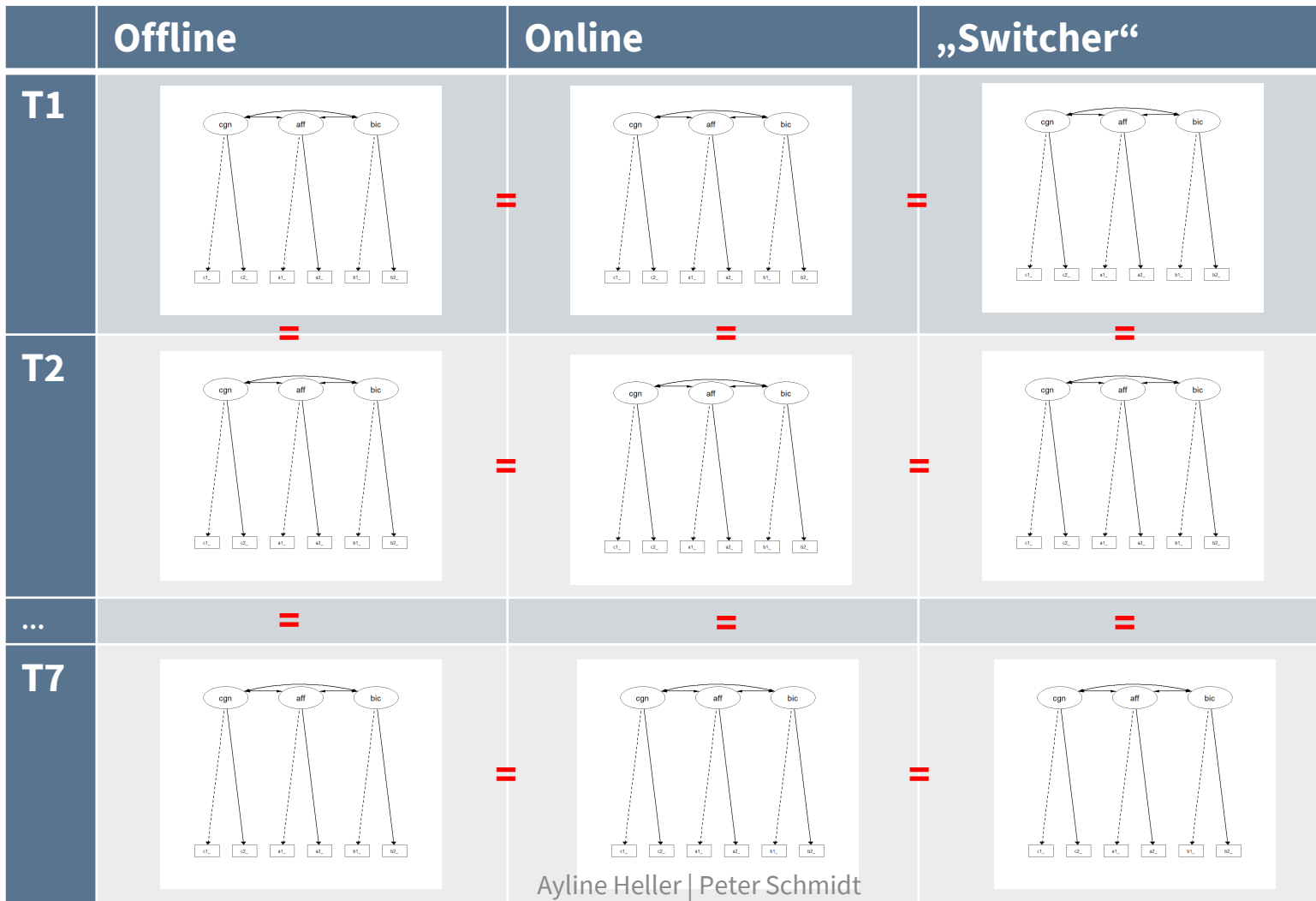
Messinvarianzanalyse

Reminder:
 Δ RMSEA \leq .015; Δ CFI $>$.99; Δ SRMR \leq .03
 (Gruppengrößen ähnlich)

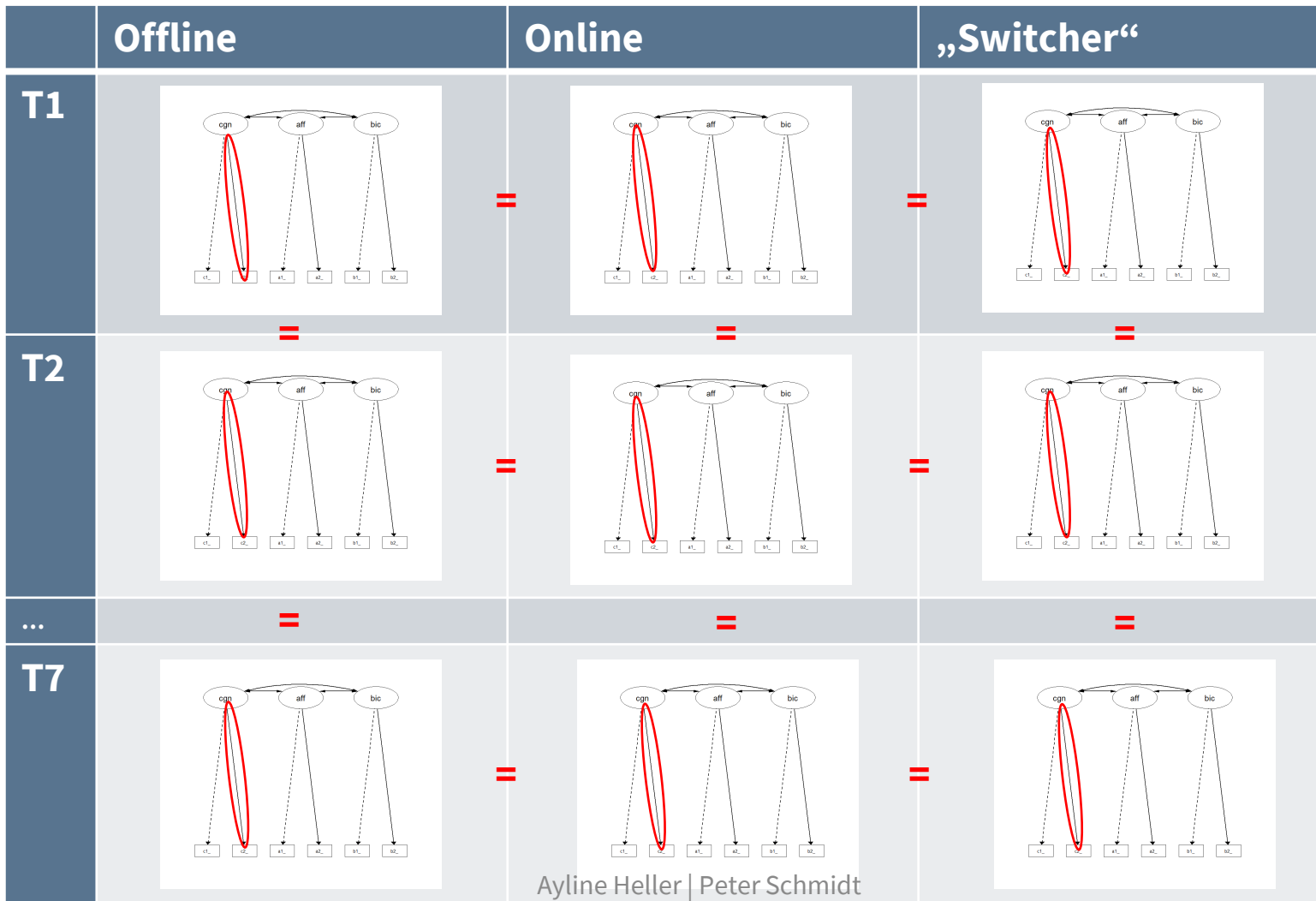
2. Für die relevanten Gruppen gemeinsam über die Zeit (Baseline nicht abgebildet)

Model	Chisq (scaled)	Df	Pvalue	RMSEA.rob	CFI.rob	TLI.rob	SRMR	AIC	BIC	N
Configural	901.723	636	.000	.012	.997	.997	.013	286781.305	290481.471	3995
Metric	1091.362	657	.000	.015 (Δ =.003)	.997 (Δ <.001)	.996	.016 (Δ =.003)	286794.395	290362.412	3995
Scalar	1291.420	678	.000	.017 (Δ =.002)	.996 (Δ =.001)	.994	.015 (Δ =.001)	286952.453	290388.321	3995
Latent means	2108.002	696	.000	.025 (Δ =.008)	.992 (Δ =.004)	.987	.020 (Δ =.005)	287733.035	291055.633	3995

Configural kombiniert



Metric kombiniert



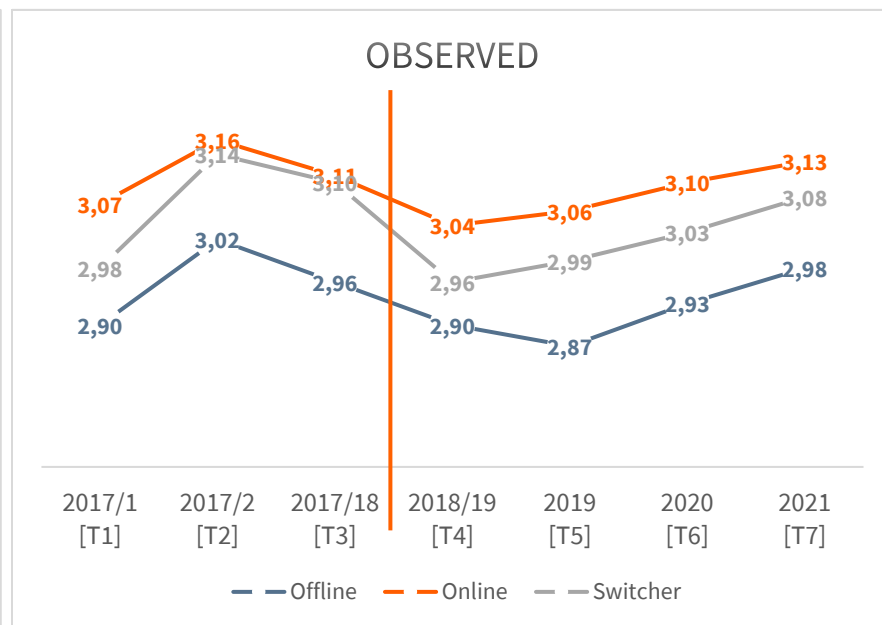
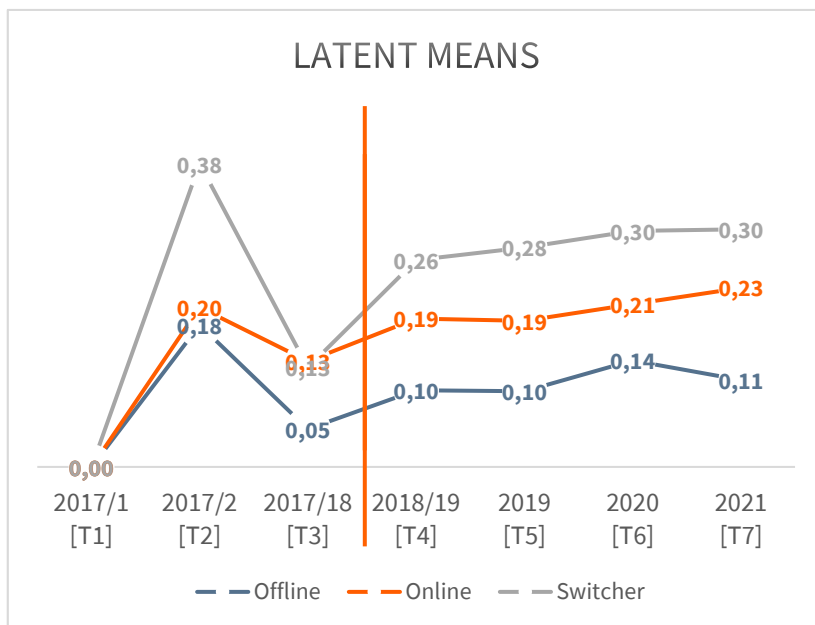
Messinvarianzanalyse

Reminder:
 Δ RMSEA \leq .01; Δ CFI $<$.005; Δ SRMR \leq .025
 (ungleiche Gruppengröße)

3. Kombinierte Analyse: Zwischen den Gruppen und über die Zeit hinweg (ohne T4)

Model	Chisq (scaled)	Df	Pvalue	RMSEA.rob	CFI.rob	TLI.rob	SRMR	AIC	BIC	N
Configural	1966.477	1449	.000	.021	.995	.992	.014	258748.593	267470.066	3994
Metric	2274.016	1509	.000	.024 (Δ =.003)	.994 (Δ =.001)	.989	.024 (Δ =.010)	258785.021	267128.941	3994
Scalar	2698.080	1575	.000	.028 (Δ =.005)	.991 (Δ =.003)	.986	.030 (Δ =.006)	259077.086	267005.697	3994
Latent means	3363.362	1626	.000	.032 (Δ =.004)	.988 (Δ =.003)	.981	.034 (Δ =.004)	259640.368	267248.059	3994

Latent means vs. observed scores: Kognitive Dimension



Fazit/Ausblick

- Bisherige Ergebnisse deuten darauf hin, dass in der betrachteten Gruppe **keine** relevanten Änderungen über die Zeit oder den Modus hinweg zu beobachten sind
- Offene Frage:
 - Ist die within-person MI im derzeitigen Modell ausreichend berücksichtigt? Wenn nicht: Wie könnte man das erreichen?

Literatur

- Bretschi, D., Schaurer, I., & Dillman, D. (2023). An Experimental Comparison of Three Strategies for Converting Mail Respondents in a Probability-Based Mixed-Mode Panel to Internet Respondents. *Journal of Survey Statistics and Methodology*, 11(1), 23–46. <https://doi.org/10.1093/jssam/smab002>
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- De Leeuw, E. D. (2005). To mix or not to mix data collections in surveys. *Journal of Official Statistics*, 21, 233–255.
- Hamaker, E.L., Kuiper, R.M., Grasman, R.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, 20(1). 102-16. <https://doi.org/10.1037/a0038889>
- Hox, J., de Leeuw, E. D., & Zijlmans, E. A. O. (2015). Measurement Equivalence in mixed mode surveys. *Frontiers in Psychology*, 6(87). <https://doi.org/10.3389/fpsyg.2015.00087>
- Kotzur, P. F., Friehs, M.-T., Schmidt, P., Wagner, U., Pötzschke, S., & Weiß, B. (2022). Attitudes towards refugees: Introducing a short three-dimensional scale. *British Journal of Social Psychology*, 61(4), 1305-1331. <https://doi.org/10.1111/bjso.12538>
- Leitgoeb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., Jak, S., Meitinger, K., Menold, N., Muthén, B., Rudnev, M., Schmidt, P., van de Schoot, R. (2023). Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social Science Research*, 110, 102805. <https://doi.org/10.1016/j.ssresearch.2022.102805>