

25. Jahrestagung der DeGEval  
**Machtwissen? Evaluation zwischen Evidenz und (Mikro-)Politik**  
14. – 16. September 2022 (Linz)

Session B3:

«Formen von Evaluationen in Organisationen der Sozialen Arbeit»

**«Evaluationsaktivitäten in Einrichtungen der Sozialen Arbeit – Formen und Einflussfaktoren»**

## Gliederung

- Einführung
- Projekt EVALSA:
  - Fokus und Methode
  - Resultate
- Ausblick und Zusammenfassung

***Projekt EVALSA: «Auf dem Weg zur ‘evidenzbasierten Praxis’? Evaluationen in Organisationen der Sozialen Arbeit» (gefördert durch Schweiz. Nationalfonds SNF)***

# Einführung

**«Die höchste Vorgabe, die wir uns geben,  
ist: Wir möchten etwas bewirken mit  
unserer Intervention.»**

(K.S., 19.06.2022)

# Entwicklungen innerhalb & ausserhalb der Sozialen Arbeit



Transformation des Sozialstaats

Zunehmende Bedeutsamkeit von  
«Wirkungsorientierung» & «evidenzbasierte  
Praxis»

## Studien – einige Beispiele

- **NPO-Organisationen der Verbände CURAVIVA, INSOS und SPITEX, Schweiz (N=412; Bernet/Gmür 2015):**
  - 54% überprüfen die langfristige Zielerreichung; u.a. auf Basis von Selbstbeurteilungen. Schätzung: ca. 20% der Organisationen nutzen für die Wirkungsmessung wissenschaftliche Methoden
- **NPO-Organisationen, Deutschland (N=125; Phineo 2013):**
  - 54% stimmen der Aussage (eher) zu «Es findet eine quantitative und qualitative Erhebung zur Überprüfung der tatsächlich erreichten Erfolge statt.»
- **Organisationen im Sozialbereich, Deutschschweiz (2019, N=180; Hochuli et al 2020):**
  - 70% der Führungskräfte ist es «wichtig», die Wirkungen zu kennen
  - 75% setzen interne Aktivitäten ein, um Wirkungen Ihrer Leistungen und Angebote zu erfassen

# Begriffsverständnis - evaluieren & Wirkung?

- **versch. Formen v. Evaluationen:** Kosten-Nutzen-Analyse, Prozessevaluation, Wirkungsevaluation, etc...
  - ➔ Fokus: **interne** Evaluation
- **Evaluationsaktivitäten: Aufteilung von Aktivitäten in...**
  - **Erfassung** von Daten (was, wie und wann)
  - **Auswertung** der Daten
  - **Nutzung** (z.B. Legitimieren, Reflexion etc.)
  - ➔ versch. Formen der «**Bündelung von Aktivitäten**»
- **Ergebnis:** die Situation nach der erbrachten DL. Dies schliesst ein:
  - Outcome: unmittelbar, fallbezogen;
  - Impact: weitreichend, über Einzelfall hinausgehend



in Anlehnung an  
Haunberger, 2018, S. 6

## Was beeinflusst Evaluationsaktivitäten?

### Organisationsexterne Faktoren (Umwelt)

- Externe Vorgaben (Zwang, Gesetz, Leistungsauftrag, ...)
- Zertifizierungs- und Qualitätsmanagementsysteme
- Professionelle Standards

### Organisationsinterne Faktoren (evaluation capacity)

- **Leitung:** Commitment / Unterstützung von Evaluationsaktivitäten
- **Kultur:** Evaluationskultur, Lernkultur
- **Ressourcen:** finanzielle und zeitliche Mittel, Infrastruktur
- **Kompetenzen:** Fähigkeiten und Kompetenzen zur Durchführung von Evaluationen
- **Organisationsmerkmale:** Anzahl Mitarbeitende, Grösse der Organisation
- **methodologische Herausforderungen:** klare und messbare Ziele, individualisierte Praxis
- **Interne Vorgaben:** Verbindlichkeit, Art der Datenerhebung, verwendete Instrumente etc.

(vgl. Bieri/Studer 2016; Cousins et al. 2014; Carman/Fredericks 2010)





# Projekt EVALSA: Methode & Resultate

## Projektdetails

- **Methode:** standardisierte Online-Befragung (März – Mai 2021)
  - Fokus 1: «Evaluationsaktivitäten in Einrichtungen der SA»
  - Fokus 2: «organisationsinterne / -externe Einflussfaktoren»
- **Zielgruppe:** Leitungspersonen von Einrichtungen der SA in der Deutschschweiz
- **Samples:**
  - A) «Sozialwesen»: Betriebs- und Unternehmensregister (BUR;  $N \approx 900$ );
  - B) drei Arbeitsfelder der SA (KSA, SSA, BSA;  $N \approx 950$ ; Basis: Adressen Verbände, Daten aus Kantonen, eigene Recherchen etc.);
- **Rücklauf:** 38 % ( $N = 739$ )

BUR:	395
Schulsozialarbeit (SSA):	168
Gesundheitsbezogene SA (KSA):	130
Betriebliche Soziale Arbeit (BSA):	46

## Formen der erfassten Informationen in % (N)

→ Einrichtungen fokussieren bei der Erfassung mehrheitlich sowohl auf Leistungs- & Qualitätsparametern (Anzahl Klienten/Klientinnen, Ressourcen, Zufriedenheit) sowie auf Ergebnisse

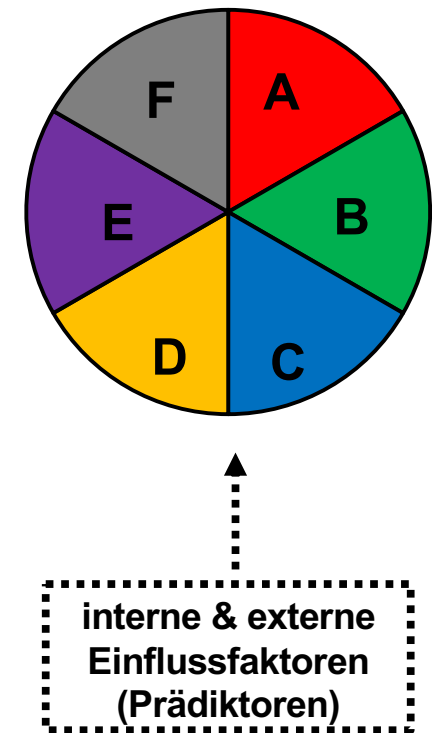
Antwort	'Ja'
Inputs (Ressourcen)	<b>89.5%</b> (624)
Outputs	
– Anzahl Dienstleistungen	<b>81.4%</b> (579)
– Anzahl Klientinnen	<b>92.0%</b> (664)
Zufriedenheit	
– Klient*in	<b>53.8%</b> (396)
– Angehörige	<b>19.3%</b> (139)
– Finanzgeber*in	<b>38.2%</b> (278)
– Kooperationspartner*in	<b>25.0%</b> (180)
Outcome-Erfassung (Ergebnisse)*	<b>80.6%</b> (596)
* im grössten Dienstleistungsbereich	



in Anlehnung an  
Hau Berger, 2018, S. 6

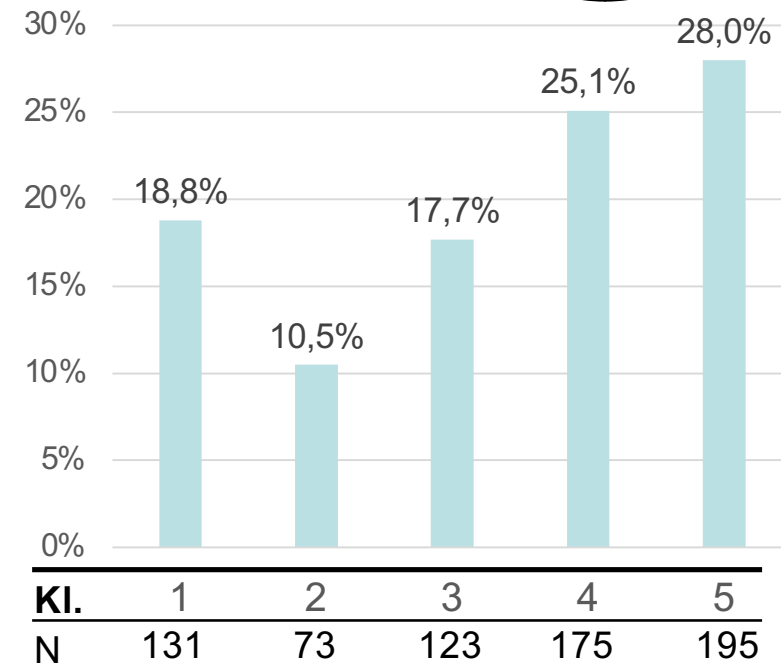
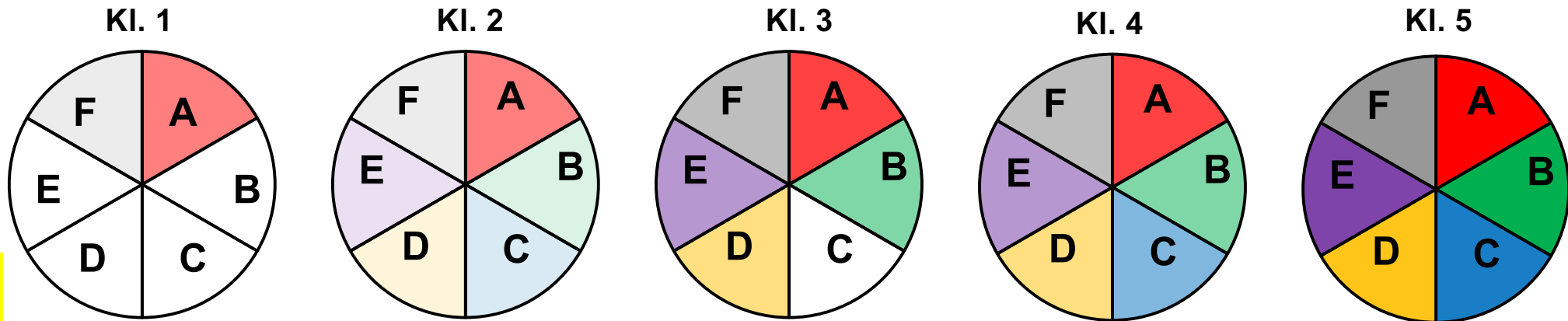
## Typen der Evaluierung – Ergebnis einer LCA

Aktivitätenbündel	verwendete Indikatoren
A: Erfassung Input & Output	Ressourcen; Anz. Klient*innen; Anz. Dienstleistungen; Zufriedenheit
B: Erfassung Ergebnisse*	sys. Ergebniserfassung im gr. / anderen Bereich; syst. Methodenanwendung; stand. Instrumente; multiperspektivisch; MZP / Design; Impact-Erfassung
C: Auswertung*	Syst. Auswertung; Vergleiche (Einrichtungs- / Einzelfallziele / Ressourcen / Vorjahres- / Leistungsdaten / Gruppen / Standorte, andere Organisationen)
D: Nutzung Org-Ebene*	<b>OE:</b> Zielerreichung, Grundsatzentscheide, Optimierung <b>Wissen:</b> Fortbildung; interne Wissensbasis; Wissensaufbau Profession <b>Legitim:</b> Legitimierung; Reporting; Informationszwecke
E: Nutzung DL-Erbringung*	Entscheidungen; Zielprüfung; Vorgehen; Reflexion
F: ext. / weiterf. Eval-Aktivitäten	Prozessmonitoring; ext. Evaluationen (Fokus Wirkung); weiterführende Evaluationen (Fokus Wirkung)
*auf grössten Dienstleistungsbereich bezogen (= meisten Ressourcen aufgewendet)	

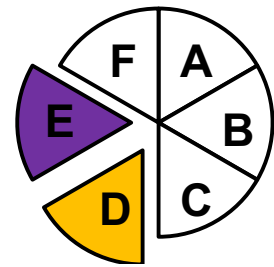


# 5-Klassen Modell

- A** Erfassung Inputs & Outputs
- B** Erfassung Ergebnisse
- C** Auswertung
- D** Nutzung Org-Ebene
- E** Nutzung DL-Erbringung
- F** Externe / Weiterführende Eval-Aktivitäten



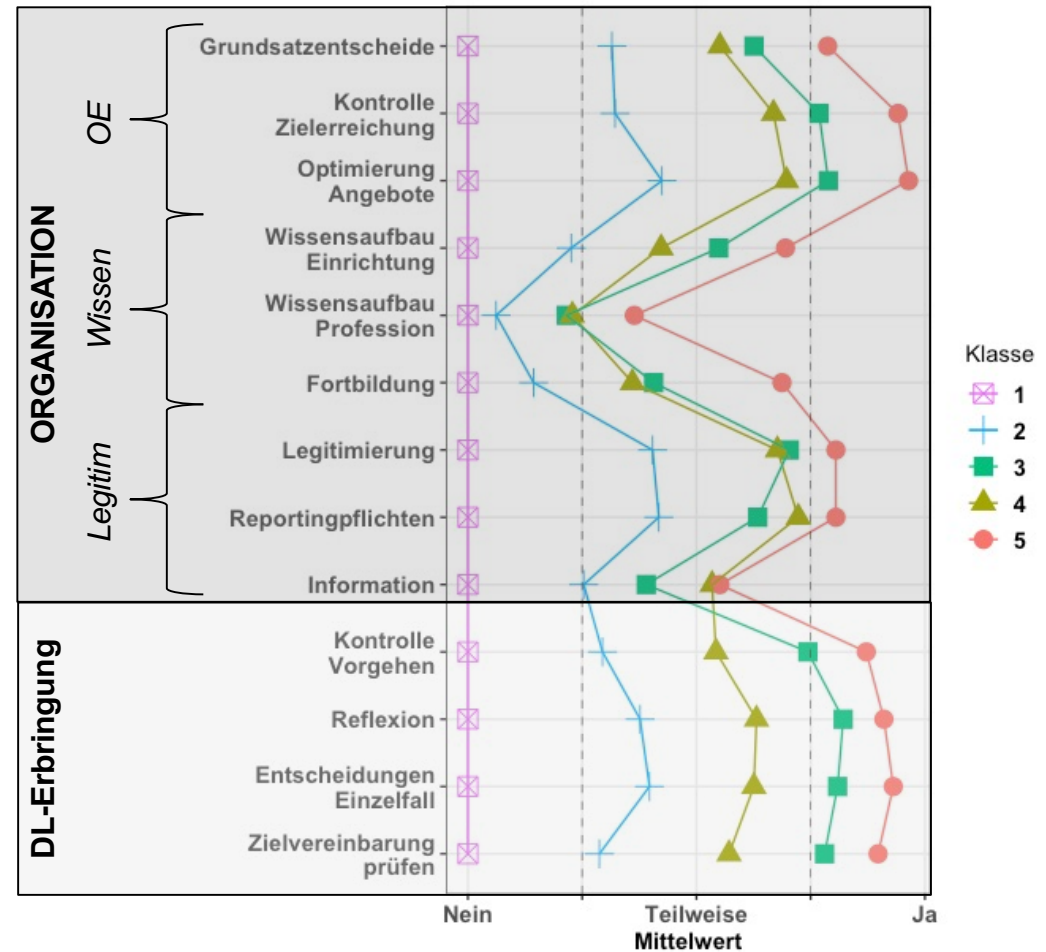
- Variation: wenig ausdifferenzierte Eval-Praxis innerhalb KI. 1 & 2 vs. weiter fortgeschrittene Praxis innerhalb KI. 3 & 4 vs. ausdifferenzierte Eval-Praxis KI. 5 (zunehmende Systematisierung, Häufigkeit von Aktivitäten, meth. Güte, vorgenommene Vergleiche, Nutzungsformen auf Organisations- und DL-Erbringungsebene etc.)
  - **KI. 1:** «keine Evaluationspraxis»
  - **KI. 2:** «beschränkte Evaluationspraxis»
  - **KI. 3:** «fortgeschrittene Evaluationspraxis, ohne weiterführende Auswertung»
  - **KI. 4:** «fortgeschrittene Evaluationspraxis»
  - **KI. 5:** «umfangreiche Erfassungs-, Auswertungs- und Nutzungspraxis»



# Nutzung auf Organisations- & DL-Erbringungsebene

- Einrichtung Kl. 5: OE > Legitimierung (Kl. 2,3,4: OE ≈ Legitimierung)
  - auch Nutzung auch auf DL-Erbringungsebene am wahrscheinlichsten
- Einrichtung Kl. 2: Tendenz Legitimierung > OE
- Einrichtung Kl. 3 & 4 bzgl. Nutzung auf Org.-Ebene ähnlicher (vgl. mit DL-Erbringungsebene)
- Generell:
  - organisations- & dienstleistungsbezogene Nutzung gekoppelt
  - Klassenübergreifend: Wissen < OE & Legitimierung
    - «Wissensaufbau Profession» am wenigsten bedeutsam

Nutzung auf Organisations- und Dienstleistungserbringungsebene



# Einfluss organisationsinterner und -externer Faktoren

Predictor	r-class $c_R$	class $c_x$	$b$	SE	$t$	$p$	OR	CI.lower	CI.upper	
<b>keine Erfassungspraxis vs. Erfassungspraxis</b>										
int. Vorgaben (verb.)	1	3	0,94	0,43	2,17	0,030	2,57	1,10	6,02	
ext. Vorgaben Typ 1	1	4	0,50	0,25	1,98	0,049	1,64	1,00	2,68	
int. Ressourcen	1	4	0,69	0,20	3,50	0,001	2,00	1,36	2,94	
int. Vorgaben (unverb.)	1	4	1,19	0,47	2,53	0,012	3,28	1,31	8,24	
int. Vorgaben (verb.)	1	4	0,99	0,37	2,68	0,008	2,70	1,31	5,59	
ext. Vorgaben Typ 1	1	5	0,58	0,29	2,01	0,046	1,78	1,01	3,12	
ext. Ressourcen	1	5	-0,50	0,24	-2,11	0,036	0,61	0,38	0,97	
Leitung	1	5	0,76	0,25	2,99	0,003	2,14	1,30	3,53	
int. Ressourcen	1	5	0,81	0,23	3,50	0,001	2,25	1,43	3,55	
Ziele	1	5	0,66	0,27	2,46	0,014	1,93	1,14	3,25	
int. Vorgaben (verb.)	1	5	1,90	0,62	3,07	0,002	6,67	1,98	22,42	
int. Vorgaben (unverb.)	1	5	1,82	0,52	3,50	0,001	6,19	2,23	17,19	
Anzahl MA dipl.	1	5	0,02	0,01	2,07	0,040	1,02	1,00	1,04	
<b>fortgeschrittene Erfassungspraxis (Ø Auswertung) vs. umfangreiche Erfassungspraxis</b>										
Leitung	3	5	0,66	0,31	2,12	0,035	1,93	1,05	3,54	
<b>fortgeschrittene Erfassungspraxis vs. umfangreiche Erfassungspraxis</b>										
Sektor (privat)	4	5	1,39	0,60	2,31	0,022	4,01	1,23	13,02	
Ziele	4	5	0,51	0,25	2,03	0,043	1,66	1,02	2,72	

Notiz: r-class: Referenzklasse; b: multinomial logit coefficient; SE: standard error; OR: Odds Ratio d. Klassenzugehörigkeit; CI.lower / CI.upper: 95% Konfidenzintervall (untere / obere) Grenze

Farben: intern; extern;

# Ausblick



## Ausblick: Qualitativer Teil

Klasse	Anzahl Organisationen	Anzahl Interviews	Arbeitsfelder
Klasse 1	2 Organisationen	5 Interviews	Betr. Sozialdienst & Kl. Sozialdienst
Klasse 3	3 Organisationen	7 Interviews	SSA, Kl. Sozialdienst & Sozialdienst
Klasse 5	3 Organisationen	7 Interviews	Kinder- und Familienhilfe & gesund. SA (Sucht)
<b>Total</b>	<b>8 Organisationen</b>	<b>19 Interviews</b>	

# Zusammenfassung

## Zusammenfassung

- Rund **80%** erfassen Ergebnisse ihrer Dienstleistungen.
- Es gibt eine **hohe Variation** von Evaluationsaktivitäten (von kaum / wenig ausgeprägt (Klasse 1) bis zu ausdifferenziert (Klasse 5))
- Eine **organisations- und dienstleistungsbezogene Nutzung** von Evaluationsergebnissen kommt **kombiniert** vor, schliesst sich nicht aus.
- Der **Wissensaufbau & -transfer** sind ein **untergeordneter Zweck** von Evaluationsresultaten (im Vergleich zu Legitimierung, Organisationsentwicklung).
- Die Evaluationsaktivitäten **positiv beeinflussen** generell **externe Vorgaben** (Finanzgeber\*in; Gesetzgeber\*in), **Ressourcen** (finanziell, zeitlich), die Haltung der **Leitung** sowie messbare **Ziele**.

# Literatur

## Quellen (I)

- Bernet, Pius/Gmür, Markus. (2015). Leistungs- und Erfolgsmessung in Sozialen und Gesundheits-NPO. In: *Verbandsmanagement*, 2(2015), 23-33.
- Bieri, Oliver/ Studer, Sibylle (2016). Wirkungsmessung bei NPO: Notwendiges Übel oder Teil der Strategie? Wirkungsmessung und Entwicklungspotenzial bei Spenden sammelnden Organisationen in der Schweiz. In: *Soziale Arbeit*, 65(6/7), 267-273.
- Carman, Joanne G./Fredericks, Kimberly A. (2010). Evaluation Capacity and Nonprofit Organizations. Is the Glass Half-Empty or Half-Full? In: *American Journal of Evaluation*, 31(1), 84-104.
- DiMaggio, P. J./Powell, W. W. (1983): The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields, in: *American Sociological Review*, 48, 147-160.
- Ellis, Paul D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. New York: Cambridge University Press.
- Holm, Sture (1979). A simple sequentially rejective multiple test procedure. In: *Scandinavian Journal of Statistics*, 65-70.
- Meyer, John W./Rowan, and Brian. (1977). Institutionalized Organizations: Formal Structure as Myth and Ceremony. *American Journal of Sociology*, 83(2), 340-363.
- Scott, W. Richard. (2001). *Institutions and Organizations*. (2. Aufl.). Thousand Oaks: Sage.
- Walgenbach, Peter/Meyer, Renate (2008): *Neoinstitutionalistische Organisationstheorie*. Stuttgart: Kohlhammer.

## Quellen (II)

- Alpizar, C. A. (2016). Alternative to classical latent class models selection methods for sparse binary data: An illustration with simulated data. *Revista de Matemática: Teoría y Aplicaciones*, 23(1), 199-220.
- Bolck, A./Croon, M./Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political analysis*, 12(1), 3-27.
- Collins, L. M./Fidler, P. L./Wugalter, S. E./Long, J. D. (1993). Goodness-of-fit testing for latent class models. *Multivariate Behavioral Research*, 28(3), 375-389. (Chi / G2 Fit bzw. Sparsness)
- Hagenaars, J. A., & McCutcheon, A. L. (Eds.). (2002). *Applied latent class analysis*. Cambridge University Press.
- van Kollenburg, G. H./Mulder, J./Vermunt, J. K. (2015). Assessing model fit in latent class analysis when asymptotics do not hold. *Methodology*. (Chi / G2 Fit bzw. Sparsness)
- Linzer, Drew A./Jeffrey Lewis. 2011. "poLCA: Polytomous Variable Latent Class Analysis." *Journal of Statistical Software* 42(10):1 29.
- Linzer, Drew A./Jeffrey Lewis. 2013. "poLCA: Polytomous Variable Latent Class Analysis." R package version 1.4. <https://github.com/dlinzer/poLCA>.
- Nielsen, A. M./Vach, W./Kent, P./Hestbaek, L./Kongsted, A. (2016). Using existing questionnaires in latent class analysis: should we use summary scores or single items as input? A methodological study using a cohort of patients with low back pain. *Clinical Epidemiology*, 8, 73.
- Wurpts, I. C./Geiser, C. (2014). Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Frontiers in Psychology*, 5, 920.

## Quellen Bilder

- Bild «Group work» (pietluk) <https://publicdomainvectors.org/en/free-clipart/Group-work/41429.html>
- Bild «Arbeitsplatz, Büro, Schreibtisch, Zimmer, Computer, Stuhl, Technik, Möbel, Tisch» (unbekannt) <https://pixnio.com/de/objekte/computer/arbeitsplatz-buro-schreibtisch-zimmer-computer-stuhl-technik-mobel-tisch#>
- Bild «Euro - Schweizer Franken» (Martin Abegglen) <https://www.flickr.com/photos/twicepix/16123918416>
- Bild «Einfache Uhr-Vektorgrafiken» (unbekannt) <https://publicdomainvectors.org/de/kostenlose-vektorgrafiken/Einfache-Uhr-Vektorgrafiken/12579.html>
- Bild «Diverse Meeting» (j4p4n) <https://publicdomainvectors.org/en/free-clipart/Diverse-meeting/60804.html>
- Bild «Concept of driving a car after alcohol consumption» (Marco Verch) <https://foto.wuestenigel.com/concept-of-driving-a-car-after-alcohol-consumption/>
- Bild «Risikoreicher Alkoholkonsum, 2017» (Bundesamt für Statistik) <https://www.bfs.admin.ch/bfs/de/home/statistiken/gesundheit/determinanten/alkohol.assetdetail.6466158.html>