

# Überprüfung der Test-Retest-Reliabilität der Clinical Learning Environment, Supervision and Nurse Teacher Scale (CLES+T-Skala)

Ein Instrument zur internen Qualitätssicherung praktischer Pflegeausbildung

Katharina Lienbacher<sup>1</sup>, MScN; Petra Schumacher<sup>2</sup>, MScN; Mario Milojevic<sup>3</sup>, MScN; Jan Kellerer<sup>4</sup>, MScN; Univ.-Prof. Dr. Gerhard Müller<sup>4</sup>

*Die Qualitätssicherung der praktischen Pflegeausbildung sollte mit einem validierten Evaluierungsinstrument erfolgen. Ziel der Untersuchung war daher, die Überprüfung der Test-Retest-Reliabilität der österreichischen CLES+T-Skala bei AusbildungsteilnehmerInnen der Gesundheits- und Krankenpflege. Das Retest-Intervall betrug zwischen zehn und vierzehn Tagen. Eine Datenanalyse fand auf Itemebene mittels Kendalls Tau-b und auf Subskalenniveau mit Spearman-Rho statt. 18 von 25 Items klassierten sich in dem Bereich mittlere Korrelation. Die verbleibenden sieben Items ordneten sich in die Klassifikationsstufe geringe Korrelation ein. Auf Ebene der Subskala konnten für alle Subskalen, mit Ausnahme der Subskala Pflegeverständnis auf der Station, hohe Korrelationen ermittelt werden. Mit diesen Ergebnissen kann die CLES+T-Skala als ein ausreichend stabiles Messinstrument hinsichtlich der Retest-Reliabilität bezeichnet werden.*

## Verification of the Test-Retest Reliability of the Clinical Learning Environment, Supervision and Nurse Teacher Scale (CLES+T-Scale)

An instrument for Internal Quality Assurance of Practical Nursing Education

*The test retest reliability verification of the Austrian Clinical Learning Environment, Supervision and Nursing Teacher Scale (CLES+T) used for quality assurance of practical nursing education, proofed to be a sufficiently stable measuring instrument in terms of retest reliability. The retest was performed between ten and fourteen days. The test-retest-reliability was determined for all items by rank correlation Kendall's tau-b and for data analysis of each subdimension Spearman-Rho was used. 18 of 25 items showed a moderate correlation and for seven items a low correlation was computed. The correlation coefficients for all subdimensions were high with exception the subdimension Premises of nursing on the ward. Results of this study show that the Austrian version of the CLES+T scale is a stable instrument through time.*

## Korrespondenzadresse

Katharina Lienbacher, MScN  
Kardinal Schwarzenberg Akademie  
Baderstraße 10  
A-5620 Schwarzach im Pongau  
katharina.lienbacher@gmx.at

Eingereicht am 19.08.2019  
Akzeptiert am 08.09.2020

DOI: 10.3936/1780

<sup>1</sup> Kardinal Schwarzenberg Akademie, Schwarzach im Pongau

<sup>2</sup> IMC Fachhochschule Krems, Department Health Sciences

<sup>3</sup> Gesundheits- und Krankenpflegeschule Feldkirch

<sup>4</sup> UMIT – Private Universität für Gesundheitswissenschaften, Medizinische Informatik und Technik, Hall in Tirol, Department für Pflegewissenschaft und Gerontologie, Institut für Pflegewissenschaft

## EINLEITUNG

Das Erlernen pflegerischen Handelns im klinischen Umfeld stellt für Auszubildende in der Gesundheits- und Krankenpflege einen integralen Teil ihrer Berufsausbildung dar. Die Bedeutung des Lernens im klinischen Setting spiegelt sich in vielen europäischen Ländern in der Summe der zu absolvierenden Praxisstunden wider (Warne et al., 2010). Mindestens 2480 Stunden müssen AusbildungsteilnehmerInnen in Österreich gemäß der Gesundheits- und Krankenpflege-Ausbildungsverordnung (GuK-AV) in der Praxis bewältigen (BGBl. II Nr. 179 / 1999 idgF. BGBl. II Nr. 296/2010, §18). Unterhalb dieses Stundenausmaßes ordnet sich die Summe der zu absolvierenden theoretischen Ausbildungsstunden ein – mindestens 2000 Stunden (BGBl. II Nr. 179/1999 idgF. BGBl. II Nr. 296/2010, §15).

Ausbildungseinrichtungen, welche die Qualifizierung zum gehobenen Dienst in der Gesundheits- und Krankenpflege anbieten, müssen dafür Sorge tragen, dass Auszubildende Wissen und Fähigkeiten für eine angemessene Patientenversorgung akquirieren. Hierfür kooperieren Bildungsinstitutionen mit verschiedenen Gesundheitsdienstleistern, die zur Realisierung einer praxisorientierten Ausbildung beitragen. Derartige Einrichtungen bzw. Organisationen, die den Erwerb von Handlungskompetenzen und beruflicher Identitätsfindung beeinflussen, definieren sich als *Clinical Learning Environment (CLE)*. Denn hier treffen Auszubildende auf ein Lern- und Arbeitsumfeld, die sich auf die duale Ausbildung fördernd auswirkt (Tomietto et al., 2016).

Um der Komplexität auf Metaebene zu begegnen, stellt die Einbeziehung der Lernenden in alltägliche Arbeitsabläufe eine unabdingbare Forderung dar. Die Integration der Lernenden folgt dem Ansatz des informellen Lernens. Ein Selbstständig werden in der Pflegepraxis, was ein zentrales Ziel des informellen Lernensatzes darstellt, verlangt ein Lernumfeld und Anleitungspersonen, die diesen Prozess ermöglichen. Nur ein kontinuierlicher Austausch zwischen dem Auszubildenden und der Anleitungsperson im Lernfeld Praxis erlaubt dem Lernenden Verantwortung zu übernehmen, Selbstvertrauen zu gewinnen und Unabhängigkeit zu erlangen (Bohrer, 2013). Dadurch trägt die klinische Lernumgebung zur Entwicklung einer beruflichen Identität bei (Bohrer & Walter, 2015). Auszubildende durchlaufen im Rahmen der Praktika einen professionsbezogenen Sozialisationsprozess. Dieser Sozialisationsprozess unterliegt jedoch den Kontextfaktoren der klinischen Lernumgebung (Baraz, Memarian & Vanaki, 2015).

Die Notwendigkeit einer angemessenen klinischen Lernumgebung wird in zahlreichen Studien wiederholend debattiert (Flott & Linden, 2016; Bergjan & Hertel, 2013). Diese Diskussionen scheinen sich damit zu rechtfertigen, dass die klinische Lernumgebung maßgeblich die Kompetenzentwicklung des lernenden Individuums beeinflusst (Baraz et al., 2015; Salminen et al., 2010). Nur die direkte Nutzung von theoretischem Wissen im Rahmen von berufsrealen Anwendungssituationen erlaubt die Entwicklung von Handlungskompetenzen (Witte, Greis & Darmann-Finck, 2016). Um die Kontextfaktoren der klinischen Lern-

umgebung und damit den Lernerfolg der Auszubildenden in der klinischen Praxis sicherzustellen, müssen validierte Instrumente eingesetzt werden (Hooven, 2014).

Die Clinical Learning Environment, Supervision and Nurse Teacher Scale (CLES+T) (Saarikoski, Isoaho, Warne & Leino-Kilpi, 2008) kann als Goldstandard für die Bewertung der Qualität der klinischen Lernumgebung aus Sicht der Lernenden gesehen werden (Tomietto et al., 2012).

Die CLES+T-Skala wurde in Europa mehrfach auf ihre psychometrischen Eigenschaften getestet. Dabei wurden Inhaltsvalidität, Konstruktvalidität und Übereinstimmungsvalidität sowie die interne Konsistenz der Items geprüft (Bergjan & Hertel, 2013; Warne et al., 2010; Saarikoski et al., 2008). Vor dem Einsatz der CLES+T-Skala in Österreich müssen deren Gütekriterien überprüft werden. Müller, Mylonas und Schumacher (2018) testeten die Struktur der CLES+T-Skala mittels einer Hauptkomponentenanalyse mit Promax-Rotation in Österreich. Mit der vorliegenden Vier-Faktor-Lösung können 73% der Gesamtvarianz der mehrdimensionalen Skala erklärt werden. Als Ergebnis für die interne Konsistenz der einzelnen vier Subskalen wurden Cronbachs Alphawerte zwischen 0,83 und 0,95 ausgewiesen (Müller et al., 2018). Die Test-Retest-Reliabilität der CLES+T-Skala wurde mehrfach in internationalen Untersuchungen überprüft (Žvanut et al., 2018; Lovrić et al., 2016; Gustafsson, Blomberg & Holmefur, 2015; Schäfer et al., 2012). Für die finnische Originalversion (Saarikoski et al., 2008) konnten keine Angaben zur Test-Retest-Reliabilität identifiziert werden. Die Tabelle 1 zeigt die Höhe der ermittelten Koeffizienten auf Itemebene bzw. auf Ebene der Subskalen.

## PROBLEMBESCHREIBUNG UND ZIELSETZUNG

Um der Entwicklung von Professionen im Gesundheitswesen Nachhaltigkeit zu verleihen, darf die Sichtweise von Auszubildenden nicht außer Acht gelassen werden (Kachur & Krajic, 2006, zit. aus Bergjan, 2011, S. 427). Denn die subjektiven Erfahrungen von Auszubildenden in der Gesundheits- und Krankenpflege dienen für die Beschreibung der Qualität praktischer Pflegeausbildung (Chan, 2001). Als Anknüpfungspunkt erweist sich daher im Kontext der Darlegungen der Einsatz der CLES+T-Skala (Saarikoski et al., 2008), welche Auszubildenden eine Evaluierung der klinischen Lernumgebung, Anleitung und Begleitung ermöglicht. Die CLES+T-Skala greift mit den Befragungssubskalen pädagogisch relevante Einflussfaktoren aus dem klinischen Aneignungsprozess auf. Zudem wird sie als ein validiertes Instrument im europäischen Raum beschrieben (Bergjan & Hertel, 2013). Um zu unterscheiden, ob Messergebnisse, welche anhand der CLES+T-Skala eingeschätzt werden, auf wahren Merkmalsunterschieden oder auf einem Messfehler beruhen, bedarf es der Überprüfung der Test-Retest-Reliabilität. Derzeit liegen zur Test-Retest Reliabilität der CLES+T-Skala in Österreich keine Ergebnisse vor. Daher war das Ziel der Untersuchung, die Test-Retest-Reliabilität der deutschsprachigen und an die österreichischen Gegebenheiten adaptierten CLES+T-Skala zu überprüfen.

Subskalen	Item	Schäfer et al. (2012) n=54 <sup>1</sup> ICC	Gustafsson et al. (2015) n=42 <sup>1</sup> kw, ICC <sub>1,1</sub>	Lovrić et al. (2016) n=50 <sup>1</sup> r <sub>s</sub>	Žvanut et al. (2018) n=232 r <sub>s</sub>
<b>Qualität der Anleitung und Begleitung (gesamt)</b>		-	ICC <sub>1,1</sub> =0,96	-	-
	1	0,74	-	-	0,779
	2	0,72	-	-	0,605
	3	0,60	-	-	0,800
	4	0,74	-	-	0,795
	5	0,63	-	-	0,830
	6	0,71	-	-	0,789
	7	0,84	-	-	0,698
	8	0,85	-	-	0,776
sonstige Angaben (Items)			kw=0,77-0,89	r <sub>s</sub> =0,51-0,76	
<b>Lernatmosphäre (gesamt)</b>		-	ICC <sub>1,1</sub> =0,91	-	-
	1	0,57	-	-	0,524
	2	0,68	-	-	0,499
	3	0,65	-	-	0,529
	4	0,56	-	-	0,694
	5	0,49	-	-	0,571
	6	0,55	-	-	0,547
	7	0,74	-	-	0,523
	8	0,54	-	-	0,434
	9	0,56	-	-	0,384
sonstige Angaben (Items)			kw=0,61-0,85	r <sub>s</sub> =0,32-0,76	
<b>Rolle der praxisbegleitenden Lehrkräfte (gesamt)<sup>2</sup></b>		-	ICC <sub>1,1</sub> =0,88	-	-
	1	0,71	-	-	0,352
	2	0,46	-	-	0,515
	3	0,74	-	-	0,275
	4	0,63	-	-	0,346
	5	0,67	-	-	0,413
	6	0,65	-	-	0,350
	7	0,81	-	-	0,488
	8	0,58	-	-	0,545
	9	0,50	-	-	0,218
sonstige Angaben (Items)			kw=0,55-0,94	r <sub>s</sub> =0,43-0,79	
<b>Führungsstil der Stationsleitung (gesamt)</b>		-	ICC <sub>1,1</sub> =0,70	-	-
	1	0,74	-	-	0,583
	2	0,49	-	-	0,489
	3	0,74	-	-	0,396
	4	0,69	-	-	0,483
sonstige Angaben (Items)			kw=0,45-0,57	r <sub>s</sub> =0,28-0,55	
<b>Pflegeverständnis auf der Station (gesamt)</b>		-	ICC <sub>1,1</sub> =0,75	-	-
	1	0,60	-	-	0,569
	2	0,63	-	-	0,576
	3	0,64	-	-	0,722
	4	0,80	-	-	0,651
sonstige Angaben (Items)			kw=0,53-0,76	r <sub>s</sub> =0,29-0,52	

<sup>1</sup> ICC= Intraklassen-Korrelationskoeffizient, kw=gewichteter Kappakoeffizient, r<sub>s</sub>=Spearman Rangkorrelationskoeffizient

<sup>2</sup> Diese Subskala ist nicht in der österreichischen Version der CLES-T-Skala (Müller et al., 2018) enthalten.

**Tabelle 1: Höhe der Koeffizienten – Test-Retest-Reliabilität**

## METHODE

Das methodische Vorgehen orientiert sich an einem nicht-experimentellen, korrelationellen, longitudinalen Untersuchungsdesign.

### Fragestellung

Basierend auf der Zielformulierung orientierte sich die psychometrische Testung an folgender Forschungsfrage: Wie stabil ist die deutschsprachige und an die österreichischen Gegebenheiten adaptierte Clinical Learning Environment, Supervision and Nurse Teacher Scale (CLES+T) bei der Erfassung der Qualität der klinischen Lernumgebung, Anleitung und Begleitung in der praktischen Pflegeausbildung aus Lernendenperspektive, über zwei Erhebungszeitpunkte?

### Erhebungsinstrument

Basierend auf einer umfassenden Literaturrecherche entwickelten Saarikoski und Leino-Kilpi (2002) die fünf Subskalen der Clinical Learning Environment and Supervision (CLES) Scale. Saarikoski et al. (2008) betrachteten als einen weiteren Aspekt der praktischen Pflegeausbildung die Rolle von Lehrpersonen im klinischen Aneignungsprozess. Im Jahr 2008 erweiterten Saarikoski et al. (2008) die CLES um die Subskala *Role of the nurse teacher* und die Skala wurde folgend als CLES+T bezeichnet. Die aktuelle Version der CLES+T-Skala (Saarikoski et al., 2008) besteht aus 34 Items, welche sich auf fünf Subskalen verteilen. Dabei wird das Thema der Lernumgebung durch die Subkategorien *Lernatmosphäre* (neun Items), *Führungsstil der Stationsleitung* (vier Items) und *Pflegeverständnis auf der Station* (vier Items) verkörpert. Weitere Subkategorien der CLES+T-Skala sind *Qualität der Anleitung und Begleitung* (acht Items) sowie *Rolle der praxisbegleitenden Lehrkräfte* (neun Items). Darüber hinaus impliziert die CLES+T-Skala Fragen zur begleitenden Fachperson. Alle Items der angeführten Subkategorien werden anhand einer fünfstufigen Likert-Skala (1=stimme gar nicht zu bis 5=stimme voll zu) bewertet (Bergjan, 2011). Nach Saarikoski et al. (2013) ist das Instrument in über 29 Sprachen verfügbar.

### Stichprobe

Die Stichprobenauswahl erfolgte in Kooperation mit Ausbildungsverantwortlichen von Einrichtungen, wo eine Ausbildung zum gehobenen Dienst in der Gesundheits- und Krankenpflege offeriert wird. Eingeschlossen wurden AusbildungsteilnehmerInnen, welche die Qualifizierung zum gehobenen Dienst in der Gesundheits- und Krankenpflege anstrebten, ein Mindestalter von 18 Jahren erfüllten und freiwillig an der Erhebung teilnahmen. Da die Ausbildung zum gehobenen Dienst in der Gesundheits- und Krankenpflege in Österreich derzeit sowohl auf tertiärem Niveau (Fachhochschulen und Universitäten) als auch auf Sekundarniveau (Schulen für Gesundheits- und Krankenpflege) erfolgt, wurden beide Ausbildungsniveaus gleichermaßen für die Datenerhebung vorgesehen. Auf eine Eingrenzung des Ausbildungsstandes wurde verzichtet, was sich mit der Zielintention des Erhebungsinstruments begründen lässt, eine Evaluierung der Lernumgebung, Anleitung und Begleitung unabhängig vom Ausbildungsstand. Eine weitere

Voraussetzung für die Teilnahme an der Erhebung stellte ein absolviertes Praktikum in einer Krankenanstalt gemäß der GuK-AV (BGBl. II Nr. 179 / 1999 idgF. BGBl. II Nr. 296/2010) bzw. FH-GuK-AV (BGBl. II Nr. 200/2008) dar. Der Einsatz eines Erhebungsinstruments, welches österreichische Ausbildungsstrukturen in der Gesundheits- und Krankenpflege widerspiegelt, forderte dergleichen Anwendungsbedingungen. Evaluierungen von Praktika, die Auszubildende außerhalb Österreichs absolvierten, blieben daher unberücksichtigt. Wie aus der GuK-AV (BGBl. II Nr. 179/1999 idgF. BGBl. II Nr. 296/2010, §18) hervorgeht, muss jeder Praktikumeinsatz mindestens 160 Stunden umfassen, wodurch auf die Nennung einer Mindestdauer verzichtet wurde.

Döring und Bortz (2016, S. 466) empfehlen bei der Anwendung der Retest-Methode eine besonders große Startstichprobe zu generieren. Die Testpersonen müssen zweimal für die Teilnahme an der Befragung motiviert werden, was die Gefahr einer höheren Drop-out-Rate birgt. Ein Stichprobenumfang von 120 (60 doppelte Datensätze) wurde für die vorliegende Untersuchung angestrebt.

### Datenerhebung

Für die Datenerhebung kam die österreichische Version der CLES+T-Skala von Müller et al. (2018) zum Einsatz. Analog zu Müller et al. (2018) fand die Erhebung ohne *Rolle der praxisbegleitenden Lehrkräfte* (neun Items) statt. Die Befragung erfolgte mithilfe von SoSci Survey (Leiner, 2014). SoSci Survey (Leiner, 2014) stellt ein Online-Tool dar, welches eine webbasierte und damit eine digitale Befragung ermöglichte. Sowohl zum ersten Erhebungszeitpunkt ( $t_1$ ) als auch zum zweiten Erhebungszeitpunkt ( $t_2$ ) wurden Serienmails angelegt und verschickt. Die E-Mail-Adressen der TeilnehmerInnen erhielten beim Import eine individuelle Kennung durch eine sechsstellige Zahlenkombination. Dadurch konnten die erhobenen Daten anonymisiert werden. Unter der Annahme, dass sich die im Kontext eines Praktikumeinsatzes gemachten Erfahrungen nicht grundlegend verändern, wurde unter Berücksichtigung der Empfehlung von Streiner, Norman und Cairney (2015, S. 172) und der Empfehlung von Bühner (2011, S. 159) ein Retest-Intervall von mindestens 10 Tagen, jedoch höchstens 14 Tagen anvisiert.

### Datenanalyse

Die Datenanalyse erfolgte mithilfe des Programms IBM SPSS Statistics Version 24. Das Signifikanzniveau wurde auf  $\alpha \leq 5\%$  festgelegt.

Die in der Online-Umfrage generierten Daten wurden konvertiert und anschließend analysiert. Soziodemografische Daten wurden mithilfe deskriptiver Statistik analysiert. Für die Reliabilitätsschätzung kam auf Itemebene die Rangkorrelation Kendalls Tau-b ( $\tau$ ) zum Einsatz. Auf Subskalenebene fand der Spearman Rangkorrelationskoeffizient ( $r_s$ ) Anwendung (Leonhart, 2013, S. 284ff). Ergänzend wurden die Korrelationen dem Wilcoxon-Vorzeichen-Rangtest für verbundene Stichproben (Leonhart, 2013, S. 246) unterzogen. Die Berechnung der Korrelationskoeffizienten erfolgte für jedes Item sowie für jede Subskala. Die Basis für die

Korrelationsanalyse auf Ebene der Subskalen bildete die Aufsummierung der Itemscores. Die resultierende Summe jeder Subskala zum ersten Erhebungszeitpunkt wurde mit der entsprechenden Summe zum zweiten Erhebungszeitpunkt korreliert (Schermerlell-Engel & Werner, 2011). Für die Interpretation der Höhe der Korrelationskoeffizienten fand das Klassifikationsschema nach Zöfel (2003, S. 151) Anwendung.

**Ethische Aspekte**

Ethische Überlegungen zur geplanten Untersuchung basieren auf den von Polit, Beck und Hungler (2004, S. 98ff) ausformulierten Prinzipien: Achtung vor der Würde des Menschen, Prinzip des Nutzens, Prinzip der Gerechtigkeit. Die Forschung mit einer nicht vulnerablen Personengruppe, wie Auszubildende der Gesundheits- und Krankenpflege, erforderte eine Meldung des Studienplans beim Committee for Scientific and Ethical Questions (RCSEQ) der Privaten Universität für Gesundheitswissenschaften, Medizinische Informatik und Technik (UMIT) in Hall in Tirol. Nach positivem Votum im Oktober 2017 wurde der Studienplan umgesetzt.

**ERGEBNISSE**

Die beschriebenen Ergebnisse basieren auf einer Längsschnittuntersuchung, welche in einem Abstand von mindestens zehn, aber maximal vierzehn Tagen in den Monaten Jänner und Februar 2018 in Österreich durchgeführt wurde. Im Rahmen der Erstbefragung wurden 319 Serienmails an die Auszubildenden verschickt; bei der Zweitbefragung waren es 220. Nach dem Aussortieren unvollständiger Daten und einem Datenabgleich, welche aus der Erst- und Zweitbefragung hervorgingen, konnten 160 vollständige Datensätze (Rücklaufquote 50%) für die Analyse herangezogen werden.

Beim überwiegenden Anteil der Befragten (n=131, 81,9%) lag das Alter zum Erhebungszeitpunkt zwischen 18 und 25 Jahren. Das Durchschnittsalter betrug 23,7 (SD±2,1) Jahre. Mehr als die Hälfte der UntersuchungsteilnehmerInnen (n=97, 60,6%) befand sich zum Zeitpunkt der Befragung im letzten Ausbildungsabschnitt (drittes Ausbildungsjahr bzw. fünftes oder sechstes Semester). Die Ausbildung zum gehobenen Dienst absolvierten rund zwei Drittel (n=99, 61,9%) der TeilnehmerInnen an einer Schule für allgemeine Gesundheits- und Krankenpflege. Das verbleibende Drittel (38,1%) der Einschätzungspersonen beschränkt die Ausbildung an Fachhochschulen oder Universitäten.

Hinsichtlich des zuletzt absolvierten Praktikums zeigte sich der Anteil (n=86, 53,8%) für den akut operativen Fachbereich etwas höher als der Anteil (n=74, 46,3%) für den akut konservativen Fachbereich.

Bei der deskriptiven Auswertung der Items (n=25) dominierte sowohl bei der Erst- als auch bei der Zweitbefragung weitgehend das Antwortverhalten *stimme eher zu* und *stimme voll zu*. Der überwiegende Anteil der AusbildungsteilnehmerInnen (n=113, 70,6%) erlebte eine individuelle Betreuung während des Praktikumseinsatzes. Beinahe alle TeilnehmerInnen (n=151, 94,4%) erhielten Betreuung durch eine diplomierte Gesundheits- und Krankenpflegeperson, wobei davon 71 (44,4%) Personen angaben, dass der gehobene Dienst eine Zusatzausbildung zur Praxisanleitung besaß. Nur ein marginaler Anteil (n=6, 3,8%) negierte eine funktionierende Beziehung zur Begleitperson. Die Zufriedenheit mit der erlebten klinischen Praxisphase lässt sich basierend auf diesen Ergebnissen insgesamt als hoch einstufen.

Für die Subskala *Lernatmosphäre* ergaben sich mittlere Korrelationskoeffizienten von  $\tau=0,501$  bis  $\tau=0,669$ . Jene Subskala beinhaltete das Item „Ich ging gerne auf dieser Station ins Praktikum“ mit der höchsten Rangkorrelation. Die Koeffizienten für die Subskala *Führungsstil der Stationsleitung* lagen zwischen geringer ( $\tau=0,454$ ) und mittlerer Korrelation ( $\tau=0,618$ ). Den kleinsten Wertebereich wies die Subskala *Pflegeverständnis auf der Station* auf. Für diese Subskala ergaben sich geringere Korrelationswerte von  $\tau=0,434$  bis  $\tau=0,474$ . Darüber hinaus befand sich in der Subskala *Pflegeverständnis auf der Station* die niedrigste Korrelation von  $\tau=0,434$  bis  $\tau=0,474$ . Die Subskala *Qualität der Begleitung und Anleitung* präsentierte sich mit geringen bis mittleren Korrelationskoeffizienten von  $\tau=0,499$  bis  $\tau=0,638$ . Alle Korrelationen waren bei einer Irrtumswahrscheinlichkeit von  $p=0,01$  signifikant (zweiseitig).

Keines der 25 Items klassierte sich in den Bereichen sehr hohe Korrelation oder hohe Korrelation. Auch der niedrigsten Klassifikationsstufe (sehr geringe Korrelation [ $r \leq 0,2$ ]) ließ sich kein ermittelter Koeffizient von den Items zuordnen. Folglich distribuierten sich die 25 berechneten Korrelationswerte auf die beiden Klassifikationsstufen geringe Korrelation und mittlere Korrelation (Tabelle 2).

Auffällige Unterschiede zwischen der Erst- und Zweitbefragung ergaben sich im Rahmen der deskriptiven Auswertung für sieben Items, wobei ein signifikanter Unterschied nur für drei Items mittels einer zweiseitigen Signifikanzprüfung festgestellt werden konnte.

Subskala	Items (n)	Geringe Korrelation (0,2 < r ≤ 0,5)	Mittlere Korrelation (0,5 < r ≤ 0,7)
1. Lernatmosphäre	9	-	9
2. Führungsstil der Stationsleitung	4	2	2
3. Pflegeverständnis auf der Station	4	4	-
4. Qualität der Begleitung und Anleitung	8	1	7
<b>Summe der Items</b>	<b>25</b>	<b>7</b>	<b>18</b>

Tabelle 2: Verteilung der Items in den Klassifikationsstufen

Subskala	Spearman (r <sub>s</sub> )*	95% KI	Sig. (2-seitig)**
1. Lernatmosphäre	0,788	0,771; 0,802	<0,001
2. Qualität der Begleitung und Anleitung	0,750	0,731; 0,770	0,001
3. Führungsstil der Stationsleitung	0,709	0,687; 0,731	<0,001
4. Pflegeverständnis auf der Station	0,546	0,519; 0,567	0,003

\*p<0,01, \*\* Wilcoxon-Vorzeichen-Rangtest für verbundene Stichproben (p<0,05)

Tabelle 3: Korrelationswerte der Subskalen (n=160)

fung durch den Wilcoxon-Vorzeichen-Rang-Test für verbundene Stichproben Untermauerung fand (Item 3:  $\tau$ -b=0,501,  $p<0,01$ ;  $z=-2,194$ ,  $p=0,028$ ; Item 5:  $\tau$ -b=0,571,  $p<0,01$ ;  $z=-2,168$ ,  $p=0,030$  und Item 10:  $\tau$ -b=0,486,  $p<0,01$ ;  $z=-2,338$ ,  $p=0,019$ ). Demnach verhielt sich die Mehrheit der Items (n=22) über den festgelegten Befragungszeitraum stabil.

Die erzielten Korrelationskoeffizienten der Subskalen *Lernatmosphäre*, *Führungsstil der Stationsleitung* und *Qualität der Begleitung und Anleitung* implizierten die Interpretation eines signifikanten hohen positiven Zusammenhangs ( $r_s=0,709-0,788$ ,  $p<0,01$ ). Ausschließlich die Subskala *Pflegeverständnis auf der Station* wies eine signifikante mittelstarke positive Korrelation ( $r_s=0,546$ ,  $p<0,01$ ) auf. Die Stabilität zwischen  $t_1$  und  $t_2$  der Subskalen ließ sich anhand der Ergebnisse des Wilcoxon-Vorzeichen-Rangtests nicht bestätigen (Tabelle 3).

## DISKUSSION

Die Entwicklung von Kompetenzen im klinischen Umfeld hängt von Wissen und Erfahrung ab. In Österreich findet die CLES+T-Skala derzeit keine Anwendung, um die Qualität des klinischen Praktikums zu gewährleisten. Dies gibt Anlass zur Sorge, nachdem der klinische Lernort maßgeblich an der Kompetenzentwicklung beteiligt ist (Bergjan & Hertel, 2013). Die Zufriedenheit der StudentInnen mit dem klinischen Praktikum spiegelt somit auch die Qualität der klinischen Lernerfahrung wider (Materne, Henderson & Eaton, 2017), die für die zukünftige Berufstätigkeit unerlässlich ist (Lamont, Brunero & Woods, 2015). Daher war das Ziel der Untersuchung die Überprüfung der Test-Retest-Reliabilität der deutschsprachigen und an die österreichischen Gegebenheiten adaptierten CLES+T-Skala.

Die Mehrheit der Items (n=18) ordnet sich gemäß der Klassifikation nach Zöfel (2003, S. 151) in dem Bereich mittlere Korrelation ( $\tau=0,454-0,669$ ) ein. Auf Ebene der Subskala, mit Ausnahme der Subskala *Pflegeverständnis auf der Station* ( $\tau=0,434-0,474$ ), präsentiert sich das Instrument mit hohen Reliabilitätswerten. Darüber hinaus konnten systematische Veränderungen bei drei Items und jeder Subskala durch den Wilcoxon-Vorzeichen-Rang-Test identifiziert werden. Die Mehrheit der Items (n=22) verhielt sich demnach über den festgelegten Befragungszeitraum stabil.

Die im Rahmen der Untersuchung ermittelten Rangkorrelationen wiesen signifikant hohe positive Zusammen-

hänge ( $r_s=0,709-0,788$ ,  $p<0,01$ ) für drei der vier Subskalen auf. Ausschließlich für die Subskala *Pflegeverständnis auf der Station* wurde eine signifikante mittelstarke positive Korrelation ( $r_s=0,546$ ,  $p<0,01$ ) festgestellt. Auch Lovrić et al. (2016) erzielten mittelstarke bis hohe Korrelationen ( $r_s=0,51-0,76$ ) für die Subskala *Qualität der Anleitung und Begleitung*. Jedoch geringe bis hohe Korrelationen für die Subskala *Lernatmosphäre* ( $r_s=0,32-0,76$ ), *Führungsstil der Stationsleitung* ( $r_s=0,28-0,55$ ) und *Pflegeverständnis auf der Station* ( $r_s=0,29-0,52$ ) (Lovrić et al., 2016). Ähnliche Ergebnisse spiegeln die Untersuchungsergebnisse von Žvanut et al. (2018). Auch hier wies die Subskala *Qualität der Anleitung und Begleitung* mittlere bis hohe Korrelationen ( $r_s=0,605-0,830$ ) auf und ebenso die Subskala *Pflegeverständnis auf der Station* ( $r_s=0,569-0,722$ ). Die Subskalen *Lernatmosphäre* ( $r_s=0,384-0,694$ ) und *Führungsstil der Stationsleitung* ( $r_s=0,396-0,583$ ) zeigten geringe bis mittlere Korrelationen auf. Zudem unterliegt die Ergebnisinterpretation mit dem verwendeten Klassifikationsschema von Zöfel (2003, S. 151) einerseits einer strengen Norm, andererseits stellen jene Vorgaben im literarischen Vergleich einen milden Rahmen dar. Polit et al. (2004, S. 296) sowie Streiner et al. (2015, S. 10) beschreiben einen Reliabilitätskoeffizienten über 0,7 als zufriedenstellend. Gemäß dem Klassifikationsschema, welches im Rahmen dieser Untersuchung Anwendung fand, präsentiert ein Betrag über 0,7 eine hohe Reliabilität (Zöfel, 2003, S. 151). Das Klammern an Interpretationsvorgaben steht einer Interpretation auf inhaltlicher Basis entgegen. Basierend auf einem Ergebnisvergleich von gleichartigen Studien sollte ein Reliabilitätsgrenzwert individuell festgelegt werden (Wirtz & Caspar, 2002, S. 25).

Ein weiterer zu berücksichtigender Faktor bei der Ergebnisinterpretation stellt die Zusammensetzung der Stichprobe dar (Bühner, 2011, S. 81). Fehlende Varietät in der Stichprobe führt zu einer Abschwächung der Reliabilität (Polit, 2014). Dominierende Stichprobencharakteristika liefern einen möglichen Erklärungsansatz für das Ausbleiben hoher bzw. sehr hoher Korrelationen. Einen hohen Stellenwert räumt auch Polit (2014) der Stichprobenzusammensetzung in Reliabilitätsanalysen ein. Eine maximale Varietät und folglich eine hohe Reliabilität lässt sich beim Einsatz von fokussierten Instrumenten nur bedingt erreichen (Polit, 2014). Das Fehlen von hohen bzw. sehr hohen Korrelationen in der vorliegenden psychometrischen Testung findet durch gerade erwähnte Aussage weitere Argumentation. Auch Gustafsson et al. (2015) diskutieren die Zusammensetzung ihrer Stichprobe vor dem Hintergrund erzielter Ergebnisse. Kontrovers zur vorliegenden Studie begrenzt-

te sich die Rekrutierung der BefragungsteilnehmerInnen in zuletzt genannter Untersuchung auf nur eine Ausbildungsstufe. Da definierte Ein- und Ausschlusskriterien im Vergleich zu Gustafsson et al. (2015) durchaus Variabilität boten, mutet die eher homogene Stichprobe im geschilderten Kontext jedoch nicht als ein Versäumnis der Studienplanung an. Im internationalen Vergleich mit Test-Retest-Reliabilitätsstudien unter Anwendung der CLES+T-Skala (Žvanut et al., 2018; Lovrić et al., 2016; Gustafsson et al., 2015; Schäfer et al., 2013) charakterisiert sich der Stichprobenumfang in der vorliegenden Untersuchung jedoch als groß.

Neben bisherig erläuterten Einflussfaktoren wird in Studien, welche eine Überprüfung der Test-Retest-Reliabilität verfolgen, dem gewählten Zeitintervall ein hoher Stellenwert beigemessen (Bühner, 2011, S. 159). Im Konnex mit dem Zeitintervall und der Länge des Testverfahrens wird insbesondere dem Erinnern von Antworten Bedeutung gezollt (Döring & Bortz, 2016, S. 466). Die CLES+T-Skala wurde mit 25 Items nicht als kurzes Testverfahren deklariert, wodurch der gewählte Zeitabstand von mindestens zehn Tagen und maximal vierzehn Tagen als angemessen erschien. In vergleichbaren Studien betrug der Zeitabstand zwischen den beiden Erhebungen eine Woche (Žvanut et al., 2018; Gustafsson et al., 2015). Lovrić et al. (2016) führten die Zweiterhebung bereits nach drei Tagen durch. Dass auch ein kürzerer Befragungszeitraum nicht zwingend einen Erinnerungseffekt provoziert, lassen die Ergebnisse der kroatischen Reliabilitätsprüfung vermuten (Lovrić et al., 2016). Die Beträge der erzielten Korrelationskoeffizienten auf Itemebene spiegeln weitgehend eine stringente Wahl des Befragungsintervalls wider. Obwohl auf Itemebene hohe bzw. sehr hohe Korrelationen fehlten, gestatten die präsentierten Ergebnisse keine Ignorierung des Erinnerungseffektes. Das Erinnern an Items und Antworten könnte zu einer Überschätzung der Ergebnisse geführt haben. Inwieweit der Erinnerungseffekt für erzielte Ergebnisse eine Rolle spielte, bleibt aufgrund einer fehlenden Kontrolle dieses Phänomens im Rahmen der Untersuchung derzeit offen. Ausschließlich Item 19 „Ich erhielt individuelle Begleitung“, welches sich ebenso durch eine Ich-Botschaft charakterisiert, zählte in keiner Reliabilitätsprüfung zu den stabilsten Items. Dies steht möglicherweise damit in Verbindung, dass die Betreuung während der Praxisphase variierte und das Item damit keine konstante Voraussetzungsbedingung für die Testung der Test-Retest-Reliabilität darstellt. Žvanut et al. (2018) und Lovrić et al. (2016) schildern identes Szenario jedoch im Kontext der erzielten Ergebnisse in der Subskala *Rolle der praxisbegleitenden Lehrkräfte*.

Der Anspruch auf eine Generalisierbarkeit der vorliegenden Ergebnisse erweist sich vor dem Hintergrund, dass die UntersuchungsteilnehmerInnen ausschließlich aus drei Bundesländer kamen, als unzulässig. Eine Übertragbarkeit der Ergebnisse unter dem Aspekt, dass die Minderheit eine Ausbildung auf tertiärem Niveau absolvierte, muss zudem in einem kritischen Diskurs erfolgen. Darüber hinaus erschwert das eingesetzte Analyseverfahren eine Vergleichbarkeit der Ergebnisse mit internationalen Studien. Denn

keine identifizierte Studie, welche eine Überprüfung der Test-Retest-Reliabilität der CLES+T-Skala beschreibt, führte die Korrelationsanalyse auf Basis der Rangkorrelation Kendalls Tau-b durch. Zwar wurden Parallelen hinsichtlich der Platzierung von Items bzw. Subskalen teilweise identifiziert, jedoch basierend auf divergenten Reliabilitätsmaßen. Als eine weitere Limitation der vorliegenden Untersuchung muss in Betracht gezogen werden, dass Ergebnisse durch einen Selektionsbias systematisch verzerrt wurden. Die Motivation an der Befragung teilzunehmen, um die Qualität der klinischen Lernumgebung, Begleitung und Anleitung zu evaluieren, steht möglicherweise in Verbindung mit der Ausbildungsstufe. Auszubildende, welche sich im letzten Abschnitt ihrer Ausbildung befinden, schöpfen aus einem größeren Pool an Erfahrungen, was mehr Evaluierungsperspektiven öffnet und möglicherweise folglich eine höhere Teilnahmebereitschaft mit sich bringt.

## SCHLUSSFOLGERUNG

Ergebnisse der vorliegenden psychometrischen Testung erlauben die deutschsprachige und an die österreichischen Gegebenheiten adaptierte CLES+T-Skala als ein ausreichend stabiles Messinstrument hinsichtlich der Retest-Reliabilität für den akutstationären Bereich zu bezeichnen. Vor dem Hintergrund angeführter Limitationen kann eine Empfehlung für die Anwendung der Skala in Österreich ausgesprochen werden. Unter dem Blickwinkel mangelnder Stabilität besitzt die österreichische Version der CLES+T-Skala durchaus das Potential, sensitiv auf Veränderungen zu reagieren. Damit liefern Ergebnisse dieser Untersuchung ein Indiz dafür, dass das Instrument als ein geeignetes Evaluierungsinstrument in der praktischen Pflegeausbildung gilt.

Mit dem Ausschluss der Nurse Teacher Scale wurde der Prozess-Qualität wenig Beachtung geschenkt. Da gemäß der FH-GuK-AV sowie in Rechtsvorschriften zu neu etablierten Berufsgruppen in der Gesundheits- und Krankenpflege, Pflegeassistenz und Pflegefachassistenz, die Anleitung durch eine Lehrperson nicht als verpflichtend gilt, wäre eine auf den Praxisanleiter adaptierte Version dieser Subskala erstrebenswert. Die Bescheinigung der zufriedenstellenden Messzuverlässigkeit der österreichischen Version der CLES+T-Skala beschränkt sich auf das Akutsetting. Fehlende Ergebnisse für das Langzeitpflegesetting decken damit weiteren Forschungsbedarf auf.

## LITERATUR

- Baraz, S., Memarian, R., & Vanaki, Z. (2015). *Learning challenges of nursing students in clinical environments: A qualitative study in Iran. International Journal of Health Promotion and Education, 52 (4)*. doi: 10.4103/2277-9531.162345.
- Bergjan, M., & Hertel, F. (2013). *Evaluating students' perception of their clinical placements - Testing the clinical learning environment and supervision and nurse teacher scale (CLES+T scale) in Germany. Nurse Education Today, 33 (11), 1393-1398.*

- Bergjan, M. (2011). Klinische Lernumgebung evaluieren. Die Clinical Learning Environment Supervision Scale. *Pflegezeitschrift*, 64 (7), 426-429.
- Bohrer, A. (2013). Evaluation von Lernprozessen. Perspektiven auf der Basis einer empirischen Studie zum informellen Lernen in der praktischen Pflegeausbildung. *Pflegewissenschaft*, 15 (5), 353-366.
- Bohrer, A., & Walter, A. (2015). Entwicklung beruflicher Identität – empirische Erkenntnisse zum Lernen in der Berufspraxis. *Pädagogik der Gesundheitsberufe*, 2 (3), 23-31.
- Bühner, M. (2011). Einführung in die Test- und Fragebogenkonstruktion. 3. Aufl. München: Pearson Deutschland GmbH.
- Chan, D. (2001). Combining qualitative and quantitative methods in assessing hospital learning environments. *International Journal of Nursing Studies*, 38 (4), 447-459.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation*. 4. Aufl. Heidelberg: Springer
- FH-Gesundheits- und Krankenpflege-Ausbildungsverordnung BGBl. II Nr. 200 / 2008: FH-Gesundheits- und Krankenpflege-Ausbildungsverordnung – FH-GuK-AV
- Flott, E. A., & Linden, L. (2016). The clinical learning environment in nursing education: a concept analysis. *Journal of Advanced Nursing*, 72 (3), 501-513.
- Gesundheits- und Krankenpflege-Ausbildungsverordnung BGBl. II Nr. 296 / 2010: Änderung der Gesundheits- und Krankenpflege-Ausbildungsverordnung (GuK-AV)
- Gustafsson, M., Blomberg, K., & Holmefur, M. (2015). Test-retest reliability of the Clinical Learning Environment, Supervision and Nurse Teacher (CLES+T) Scale. *Nurse Education Today*, 15 (4), 253-257.
- Hooven, K. (2014). Evaluation of Instruments developed to measure the Clinical Learning Environment. *Nurse Educator*, 39 (6), 316-320.
- Lamont, S., Brunero, S., & Woods, K.P., 2015. Satisfaction with clinical placement—the perspective of nursing students from multiple universities. *Collegian*, 22 (1), 125–133.
- Leiner, D. J. (2014). SoSci Survey (Version 2.6.00-i). <https://www.sosicisurvey.de/> (30.03.2018).
- Leonhart, R. (2013). *Lehrbuch Statistik. Einstieg und Vertiefung*. 3. Aufl. Bern: Hans Huber.
- Lovrić, R., Piškorjanac, P., Pekić, V., Vujanić, J., Ratković, K. K., Luketić, S., Plužarić, J., Matijašić-Bodalec, D., Barać, I., & Žvanut, B. (2016). Translation and validation of the clinical learning environment, supervision and nurse teacher scale (CLES+T) in Croatian language. *Nurse Education in Practice*, 19, 48-53.
- Materne, M., Henderson, A., Eaton, E., 2017. Building workplace social capital: a longitudinal study of student nurses' clinical placement experiences. *Nurse Education Practice*, 26, 109-114. <http://dx.doi.org/10.1016/j.nepr.2017.07.007>.
- Mueller, G., Mylonas, D., & Schumacher, P. (2018). Quality assurance of the clinical learning environment in Austria: Construct validity of the Clinical Learning Environment, Supervision and Nurse Teacher Scale (CLES+T scale). *Nurse Education Today*. Published online. doi: 10.1016/j.nedt.2018.04.022.
- Polit, D. F., Beck, C. T., & Hungler, B. P. (2004). *Lehrbuch Pflegeforschung. Methodik, Beurteilung und Anwendung*. 1. Aufl. Bern: Hans Huber.
- Polit, D. F. (2014). Getting serious about test-retest reliability: a critique of retest research and some recommendations. *Quality of Life Research*. Published online. doi: 10.1007/s11136-014-0632-9.
- Saarikoski, M., & Leino-Kilpi, H. (2002). The clinical learning environment and supervision by staff nurses: developing the instrument. *International Journal of Nursing Studies*, 39 (3), 259-267.
- Saarikoski, M., Isoaho, H., Warne, T., & Leino-Kilpi, H. (2008). The nurse teacher in clinical practice: developing the new sub-dimension to the clinical learning environment and supervision (CLES) scale. *International Journal of Nursing Studies*, 45 (8), 1233-1237.
- Saarikoski, M., Kaila, P., Lambrinou, E., Perez Canaveras, R. M., Tichelaar, E., To-mietto, M., & Warne, T. (2013). Students' experiences of cooperation with nurse teacher during their clinical placements: An empirical study in a Western European context. *Nurse Education in Practice*, 13 (2), 78-82.
- Salminen, L., Stolt, M., Saarikoski, M., Suikkala, A., Vaartio, H., & Leino-Kilpi, H. (2010). Future challenges for nursing education – A European perspective. *Nurse Education Today*, 30 (3), 233–238.
- Schäfer, P., Scheidegger, M., Conca, A., Vossler, G., & Reutlinger, B. (2012). Qualitätsinstrument für die klinische Lernumgebung, Praxisbegleitung und Berufsschullehrerin/Berufsbildnerin (CLES+T). Überprüfung der Mess-zuverlässigkeit. (Jahresbericht). Kantonsspital Aarau, Aarau.
- Schermelleh-Engel, K., & Werner, C. S. (2011). SPSS-Beispiel zu Kapitel 6: Methoden der Reliabilitätsbestimmung. [https://lehrbuchpsychologie.springer.com/sites/default/files/atoms/files/moosbrugger\\_a2\\_978-3-642-20071-7\\_kapitel\\_6\\_spss.pdf](https://lehrbuchpsychologie.springer.com/sites/default/files/atoms/files/moosbrugger_a2_978-3-642-20071-7_kapitel_6_spss.pdf) (04.07.2018).
- Streiner, D. L., Norman, G. R., & Cairney, J. (2015). *Health Measurement Scales. A practical guide to their development and use*. 5. überarb. Aufl. Oxford: University Press.
- Tomietto, M., Comparcini, D., Simonetti, V., Pelusi, G., Troiani, S., Saarikoski, M., & Cicolini, G. (2016). Work-engaged nurses for a better clinical learning environment: a ward-level analysis. *Journal of Nursing Management*, 24 (4), 457-482.
- Tomietto, M., Saiani, L., Palese, A., Cunico, L., Cicolini, G., Watson, P., & Saarikoski, M. (2012). Clinical Learning Environment and Supervision plus nurse Teacher (CLES+T) scale: testing the psychometric characteristics of the Italian version. *Giornale Italiano di Medicina del Lavoro ed Ergonomia*, 34 (3), 49-55.
- Warne, T., Johansson, U. B., Papastavrou, E., Tichelaar, E., Tomietto, M., Van den Bossche, K., Vizcaya-Moreno, M. F., & Saarikoski, M. (2010). An exploration of the clinical learning experience of nursing students in nine European countries. *Nurse Education Today*, 30 (8), 809-815.
- Wirtz, M., & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität*. Göttingen: Hogrefe.
- Witte, S., Greis, G., & Darmann-Finck, I. (2016). Optimale Bedingungen für die klinische Kompetenzentwicklung auf Ausbildungsstationen/-bereichen. *Pädagogik für Gesundheitsberufe*, 3 (3), 20-27.
- Zöfel, P. (2003). *Statistik für Psychologen im Klartext*. München: Pearson Deutschland GmbH.
- Žvanut, B., Lovrić, R., Kolnika, T. Š., Šavle, M., & Pucer, P. (2018). A Slovenian version of the "clinical learning environment, supervision and nurse teacher scale (Cles+T)" and its comparison with the Croatian version. *Nurse Education in Practice*, 30, 27-34.