

Complexity of clinical cases in simulated learning environments: proposal for a scoring system

Leah Theresa Braun¹

Benedikt Lenzer²

Martin R. Fischer²

Ralf Schmidmaier¹

1 Ludwig-Maximilians-
University (LMU) Munich,
Klinikum der Universität
München, Medizinische
Klinik und Poliklinik IV,
Munich, Germany

2 Ludwig-Maximilians-
University (LMU) Munich,
Klinikum der Universität
München, Institut für
Didaktik und
Ausbildungsforschung in der
Medizin, Munich, Germany

Introduction

Diagnosing can be understood as the purposeful collection and evaluation of information for the purposes of uncertainty reduction to make a clinical decision. In medical education, students can be introduced to this task through simulations: Case-based learning has long been an important part of medical education [1]. In most curricula and in many exam situations, it is essential to work with clinical cases. Clinical cases are also used in medical education research for training and measurement of diagnostic competence. In studies, standardized cases in the form of virtual patients [2], [3] are used frequently. To improve diagnostic skills, learners are supported in case-based learning with virtual patient cases through various prompts (so-called scaffolds). Examples of scaffolds are structured reflection [4], feedback in different forms [5] or representation prompts [6]. However, the results of these studies regarding the effectiveness of the various interventions on diagnostic competence are contradictory. One reason for this could be that the degree of complexity of the used cases varies and therefore, the different scaffolding methods are of different value. The term complex (Latin: complexus, Part.adj. to Latin *complecti* “embrace”) means comprehensive, diverse, intertwined, multi-layered [https://www.dwds.de/wb/komplex, last accessed on 18.11.2018]. Complexity refers to systems with many components that can interact in different ways. In common usage, complexity is often

equated with difficulty. In the diagnostic context, however, we propose a clear separation of the two terms (see figure 1):

An example: A arithmetical problem (for example, a task in which Bayes's theorem has to be applied in order to solve the task correctly) can be of varying complexity. For example, there are Bayesian tasks as 1-test cases [7] (e.g., breast cancer is diagnosed by mammography) and 2-test cases [7] (e.g., breast cancer is diagnosed by mammography and ultrasound). 2-test cases are more complex than 1-test tasks, because they consist of more levels and are thus more complex, they have more features. The difficulty is not synonymous per se with complexity. Bayesian tasks, whether they are 1-test cases or 2-test cases, are easy to solve for math students - they are not difficult for this group of people. For medical students, on the other hand, Bayesian tasks can hardly be solved - regardless of whether they are 1-test or 2-test cases. The tasks are difficult for medical students and easy for math students. The task itself has a certain degree of complexity (1-test-case=simple/non-complex and 2-test-cases=complex), the probability of solution and thus the difficulty depends on the recipient [8]. The pre-knowledge of the person trying to solve the task influences the level of difficulty much more than the complexity does.

Complexity relates solely to features of the clinical case and is independent of the person diagnosing. It can be defined as the complexity of the clinical problem. Difficulty, on the other hand, can be understood as a dynamic

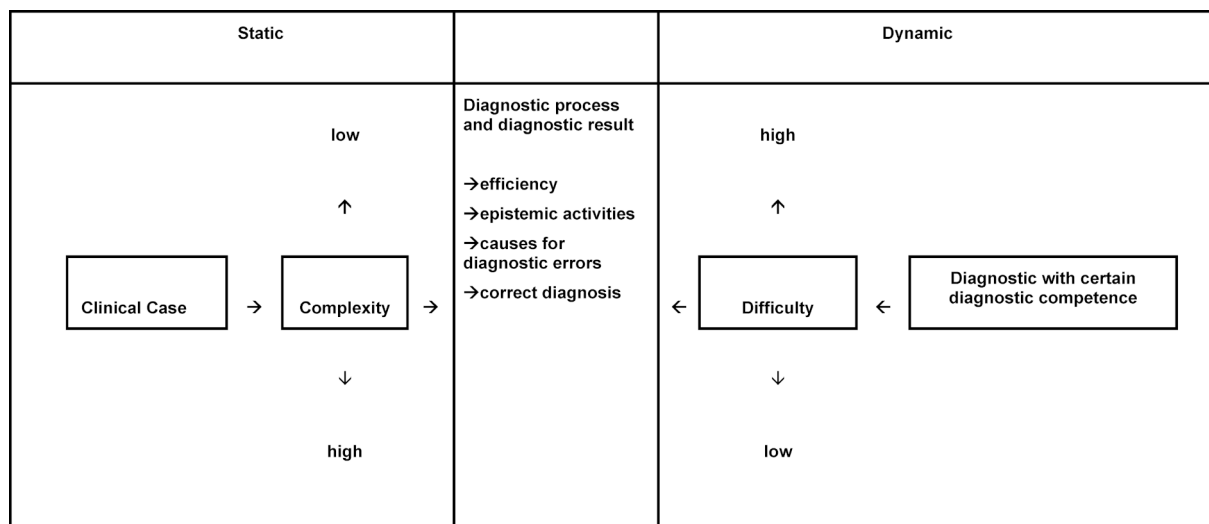


Figure 1: Differentiation: complexity and difficulty

concept, since it depends directly on the person diagnosing: There are several factors that contribute to the difficulty: In addition to the prior knowledge and experience of a person [9], for example, a case with the diagnosis Endocarditis is probably difficult for a medical student in the first clinical year, but possibly very easy for a cardiologist with several years of working experience – also the motivation, the nature of the day and ultimately pure coincidence is crucial. If the doctor has recently seen a similar case, the clinical problem may be very easy for him to solve [9]. Comparative studies of different rating systems are already available [10] to assess the difficulty of the case.

Causes that can contribute to the complexity of a patient's therapy have been studied in several studies [11]. The general increase in the complexity of medical roles in increasingly complex care systems has also been discussed extensively [12]. However, there is no applied system to determine the complexity of a case-based learning scenario. This raises the problem that study results due to different or not clearly-defined case complexity are comparable only with restrictions; and also the implementation of scientific findings in teaching practice is difficult. Even a differentiated use of cases of different degrees of complexity for different levels of expertise cannot be done without a quantification of the complexity.

The purpose of this article is to answer the following question: How can the degree of complexity of a clinical case be determined in a practicable and standardized way for simulation environments?

Methodology: Development of a literature-based model of complexity

In a fundamental study on decision-making in complex scenarios, Payne states two key features of a complex system: “The number of alternatives available” and “number of dimensions of information available” [13].

This definition as well as a case complexity classification, form the basis of our complexity model. According to our research this is the only classification in medical education or medical teaching research literature directly related to the case complexity in medical education. Hennen [14] suggests a score, consisting of five items: symptom complexity, complexity of clinical and technical findings, socioeconomic/behavioral difficulties, diagnostic complexity and management complexity. The addition of these subcategories results in a degree of complexity for the entire clinical problem [14]. Although the categorization itself includes many important aspects, the individual levels of each category are uni-dimensional and do not address the compounds of the individual elements named in the concept of complexity. Furthermore, this model [14] assesses the complexity of the final diagnosis as well as the difficulties of appropriate therapy. These aspects play no role in the definition of diagnostics used here.

The model therefore had to be adapted to the definitions of clinical diagnostics with a partial reorganization of the categories to create a general model for clinical cases in simulated learning environments. In the following, a proposal with five categories is described, which adapts the model of Hennen to the mentioned requirements. Methodologically, this is a first content validation through group discussion with expert consensus as the basis for external validity studies. As a result, the Physical Examination and Technical Findings categories were separated, while the two categories of Diagnostic and Management Complexity were combined. The original three-level matrix model was adapted to Payne, taking into account the complexity definition.

The modified scoring consists of five different categories:

1. anamnesis,
2. physical examination,
3. technical findings,
4. psychosocial aspects, and
5. secondary diagnoses (see table 1).

Table 1: Summary of scoring

Anamnesis	Information available	1 point
Physical examination	Information available	1 point
Technical examinations	Information available	1 point
Psychosocial aspects	Information available	1 point
Secondary diagnosis	existing	1 point

A clinical case of a patient usually consists of the following three basic elements: the medical history, a physical examination and technical diagnostic findings (for example, a laboratory, an X-ray or an electrocardiogram (ECG)). In addition, in clinical cases, the psychosocial context, for example the behavior and social involvement of the patient, is important. The degree of complexity can be varied at all levels. A case increases in complexity as it contains distracting information which increases the number of connections between the individual elements of the system and thus makes further differential diagnoses more likely. If the number of components in the model increases [13], which in turn interact with each other, the complexity increases.

We propose a 5 + 5 model: The more points, the more complex; that is the more complex the case. There are two steps in order to score the cases: First, it is assessed if the case contains information from all five categories. The exception here is the category diagnosis, this is not awarded a point per se but only for each further secondary diagnosis. A case gets a point in a category as soon as information is mentioned. In the second step, the linearity of the information is evaluated: If information is mentioned that is not linearly associated with the first information, then the case gets a second point. Under linear we understand that the information fits together, thus does not cause multi-layering. If a third (fourth, fifth) information level is added, which is not linked to the previous levels, then another point is assigned. The second step is independent of the previously mentioned categories, for example, a case could receive 3 additional points in the "history" category. Generally speaking, a new aspect does not generate a point if it does not repeal the linearity of a case, that is, the new information does not create a new level. On the other hand, more points are awarded when information is not linearly related to previous information but represent additional component that interacts with many other components of the system. These may be pathological / conspicuous findings as well as normal findings.

1. Linear (to each other and to a suitable diagnosis) information from the categories (see table 1)
2. Further levels of complexity within the 5 categories, additional information that is not linearly linked to the first level and are not linked to each other (maximum 6 additional levels)

A case can therefore contain a minimum of one point (only a history) and a maximum of 10 points. Although theoretically conceivable, it is unlikely that a case used for student teaching will contain more than six additional levels of information relevant to the clinical decision-making situation in the above definition of diagnosing. In addition, it can be assumed that, even in real patient cases, the degree of complexity does not increase infinitely.

Exemplary scoring of a clinical case

In the following, the scoring of the complexity will be shown on the basis of an example case of pneumothorax. The solution accuracy of this case was 45% in a study with 150 medical students in the clinical phase of studies (see attachment 1).

Discussion and outlook

Target of the scoring

The degree of complexity of a clinical case is fundamental for case developers, as clinical cases can be deliberately varied in complexity to suit a particular target audience. With increasing levels of education, students should be confronted with increasingly complex cases and diminishing supportive measures. For all studies in which clinical cases are used, it should be a requirement that the degree of complexity of the case be described in a standardized way. The aim should be that results from intervention studies are comparable across research groups.

Complexity and difficulty

Further studies should investigate the impact of complexity on case difficulty, with accuracy being described not only by the accuracy of the diagnostic result but also by other dependent variables such as diagnostic efficiency or types of errors [6], [15]. Diagnosing is difficult. Every tenth diagnosis is a misdiagnosis [16], [17]. In general, there are two sources of error: lack of knowledge and lack of information processing. In particular the degree of complexity could cause errors due to incorrect information processing. This is especially important if in one case several differential diagnoses are conceivable or the patient has several clinical problems that need to be

weighted correctly. To what extent the degree of complexity of a clinical case contributes to the difficulty is unclear.

Outlook: validation of the model and modulation of complexity

The aim of this work and of the working group is to develop an instrument for assessing the case complexity in order to be able to correctly describe this important case characteristic and to be able to use corresponding cases specifically for teaching and research. Whether the proposed scoring system can reliably measure the complexity of clinical cases for simulation-based learning environments must be verified in validation studies. According to Kane's validation tool by Cook et. al. this commentary describes the beginning of a validation process ("articulating the claims and assumptions associated with the proposed decision (the interpretation/use argument)") [18]. Evidence must now be generated in the further course, because: "Just as one can never prove a hypothesis, validity can never be proven"; [19]. The generalizability and practicability of the scoring system will be reviewed in a multistage expert review using the Angoff method [20].

The goal would be to divide the now 10-part scale into sections: less complex, moderately complex, very complex. In this context, it should also be examined whether the difficulty is influenced by the complexity modulation of the case given the same level of expertise. It is also unclear whether there are other aspects that are included in the degree of complexity but are not yet covered by the model. Whether and to what extent the model can also be applied to other topics and professional groups or in general to problem-solving tasks is another important aspect of the future validation study.

We would be grateful for feedback from other researchers on the practicability and validity of our assessment model.

Competing interests

The authors declare that they have no competing interests.

Attachments

Available from

<https://www.egms.de/en/journals/zma/2019-36/zma001288.shtml>

1. Attachment_1.pdf (101 KB)
Exemplary scoring of a clinical case

References

1. Cabot RC. Case teaching in medicine. Lexington: DC Heath & Company; 1906.
2. Huwendiek S, De Ieng BA, Zary N, Fischer MR, Ruiz JG, Ellaway R. Towards a typology of virtual patients. *Med Teach*. 2009;31(8):743-748. DOI: 10.1080/01421590903124708
3. Kononowicz AA, Zary N, Edelbring S, Corral J, Hege I. Virtual patients—what are we talking about? A framework to classify the meanings of the term in healthcare education. *BMC Med Educ*. 2015;15:11. DOI: 10.1186/s12909-015-0296-3
4. Mamede S, Van Gog T, Sampaio AM, De Faria RMD, Maria JP, Schmidt HG. How can students' diagnostic competence benefit most from practice with clinical cases? The effects of structured reflection on future diagnosis of the same and novel diseases. *Acad Med*. 2014;89(1):121-127. DOI: 10.1097/ACM.0000000000000076
5. Kopp V, Stark R, Fischer MR. Fostering diagnostic knowledge through computer-supported, case-based worked examples: effects of erroneous examples and feedback. *Med Educ*. 2008;42(8):823-829. DOI: 10.1111/j.1365-2923.2008.03122.x
6. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, Fischer MR, Schmidmaier R. Representation scaffolds improve diagnostic efficiency in medical students. *Med Educ*. 2017;51(11):1118-1126. DOI: 10.1111/medu.13355
7. McDowell M, Jacobs P. Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychol Bull*. 2017;143(12):1273-1312. DOI: 10.1037/bul0000126
8. Binder K, Krauss S, Bruckmaier G, Marienhagen J. Visualizing the Bayesian 2-test case: The effect of tree diagrams on medical decision making. *PLoS One*. 2018;13(3):e0195029. DOI: 10.1371/journal.pone.0195029
9. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving: an analysis of clinical reasoning. Cambridge (Mass): Harvard University Press; 1978. DOI: 10.4159/harvard.9780674189089
10. Koskinen H. Evaluation of the level of difficulty of patient cases for veterinary problem-solving examination: a preliminary comparison of three taxonomies of learning. *J Vet Med Educ*. 2007;34(2):106-111. DOI: 10.3138/jvme.34.2.106
11. Safford MM, Allison JJ, Kiefe CI. Patient complexity: more than comorbidity. The vector model of complexity. *J Gen Intern Med*. 2007;223 Suppl 3:382-390. DOI: 10.1007/s11606-007-0307-0
12. Plsek PE, Greenhalgh T. Complexity science: The challenge of complexity in health care. *BMJ*. 2001;323(7313):625. DOI: 10.1136/bmj.323.7313.625
13. Payne JW. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ Behav Human Perform*. 1976;16(2):366-387. DOI: 10.1016/0030-5073(76)90022-2
14. Hennen BK. Measuring the complexity of clinical problems. *Acad Med*. 1984;59(6):487-493. DOI: 10.1097/00001888-198406000-00005
15. Braun LT, Zwaan L, Kiesewetter J, Fischer MR, Schmidmaier R. Diagnostic errors by medical students: results of a prospective qualitative study. *BMC Med Educ*. 2017;17(1):191. DOI: 10.1186/s12909-017-1044-7
16. Graber ML, Carlson B. Diagnostic error: the hidden epidemic. *Physician Exec*. 2011;37(6):12-14, 16, 18-19.
17. Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf*. 2013;22 Suppl 2:ii21-ii27. DOI: 10.1136/bmjqs-2012-001615
18. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ*. 2015;49(6):560-575. DOI: 10.1111/medu.12678

19. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. *Adv Simul* (London). 2016;1:31. DOI: 10.1186/s41077-016-0033-y
20. Jalili M, Hejri SM, Norcini JJ. Comparison of two methods of standard setting: the performance of the three-level Angoff method. *Med Educ*. 2011;45(12):1199-1208. DOI: 10.1111/j.1365-2923.2011.04073.x

Corresponding author:

Leah Theresa Braun
Ludwig-Maximilians-University (LMU) Munich, Klinikum
der Universität München, Medizinische Klinik und
Poliklinik IV, Ziemssenstr. 1, D-80336 Munich, Germany
Leah.Braun@med.uni-muenchen.de

Please cite as

Braun LT, Lenzer B, Fischer MR, Schmidmaier R. Complexity of clinical cases in simulated learning environments: proposal for a scoring system. *GMS J Med Educ*. 2019;36(6):Doc80. DOI: 10.3205/zma001288, URN: urn:nbn:de:0183-zma0012880

This article is freely available from

<https://www.egms.de/en/journals/zma/2019-36/zma001288.shtml>

Received: 2018-04-15

Revised: 2018-11-24

Accepted: 2019-01-22

Published: 2019-11-15

Copyright

©2019 Braun et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

Komplexität klinischer Fälle in simulierten Lernumgebungen: Vorschlag für ein Bewertungsverfahren

Leah Theresa Braun¹

Benedikt Lenzer²

Martin R. Fischer²

Ralf Schmidmaier¹

1 Ludwig-Maximilians-Universität (LMU) München, Klinikum der Universität München, Medizinische Klinik und Poliklinik IV, München, Deutschland

2 Ludwig-Maximilians-Universität (LMU) München, Klinikum der Universität München, Institut für Didaktik und Ausbildungsforschung in der Medizin, München, Deutschland

Einführung

Diagnostizieren kann als das zielgerichtete Sammeln und Bewerten von Informationen zum Zwecke der Unsicherheitsreduktion verstanden werden, um eine klinische Entscheidung zu treffen. In der medizinischen Ausbildung können Studierende durch **Simulationen** an diese Aufgabe herangeführt werden: Fallbasiertes Lernen ist schon seit langem ein wichtiger Bestandteil der medizinischen Ausbildung [1]. In den meisten Curricula und in vielen Prüfungssituation ist die Arbeit an und mit klinischen Fällen ein elementarer Bestandteil. Klinische Fälle werden auch in der medizindidaktischen Forschung zum Training und zur Messung der Diagnosekompetenz eingesetzt. In Studien werden häufig standardisierte Fälle in Form virtueller Patienten [2], [3] genutzt. Zur Verbesserung der Diagnosekompetenz werden Lernende beim fallbasierten Lernen mit virtuellen Patientenfällen durch verschiedene Aufforderungen und Hilfen (sogenannte Scaffolds) unterstützt. Beispiele für Scaffoldings sind die strukturierte Reflexion [4], Feedback in unterschiedlichen Formen [5] oder Repräsentationsprompts [6]. Die Ergebnisse dieser Studien in Bezug auf die Effektivität hinsichtlich der Diagnosekompetenz der verschiedenen Interventionsmaßnahmen sind allerdings widersprüchlich. Ein Grund dafür könnte sein, dass der Komplexitätsgrad der verwendeten

Fälle variiert und daher die genannten Unterstützungsmaßnahmen unterschiedlich wirksam sind.

Der Begriff **komplex** (aus dem Lateinischen hergeleitet von *complexus*, Part.adj. zu lat. *complecti* „umschlingen, umfassen“) bedeutet umfassend, vielfältig, verflochten, vielschichtig [https://www.dwds.de/wb/komplex, zuletzt aufgerufen am 18.11.2018]. Komplexität bezieht sich auf Systeme mit vielen Komponenten, die auf verschiedene Weisen miteinander interagieren können. Im allgemeinen Sprachgebrauch wird Komplexität oft mit Schwierigkeit gleichgesetzt. Im diagnostischen Kontext schlagen wir allerdings eine saubere Trennung der beiden Begrifflichkeiten vor (siehe Abbildung 1):

Ein Beispiel: Eine Rechenaufgabe (Beispielsweise eine Aufgabe, in welcher der Satz von Bayes angewendet werden muss, um zum richtigen Ergebnis zu gelangen) kann unterschiedlich komplex sein. So gibt es Bayes-Aufgaben als 1-Test-Fälle [7] (z.B. Brustkrebs wird mithilfe einer Mammographie diagnostiziert) und als 2-Test-Fälle [7] (z.B. Brustkrebs wird mithilfe einer Mammographie und einer Ultraschalluntersuchung diagnostiziert). 2-Test-Fälle sind komplexer als 1-Test-Aufgaben, da sie aus mehr Ebenen bestehen und damit vielschichtiger sind, also mehr Merkmale aufweisen. Die Schwierigkeit ist aber nicht per se gleichbedeutend mit der Komplexität. So sind Bayes-Aufgaben, egal ob sie 1-Test-Fälle oder 2-Test-Fälle sind, für Mathematikstudenten problemlos zu lösen – sie sind für diese Personengruppe *nicht schwierig*. Für Medizinstudierende hingegen sind Bayes-Aufgaben – und

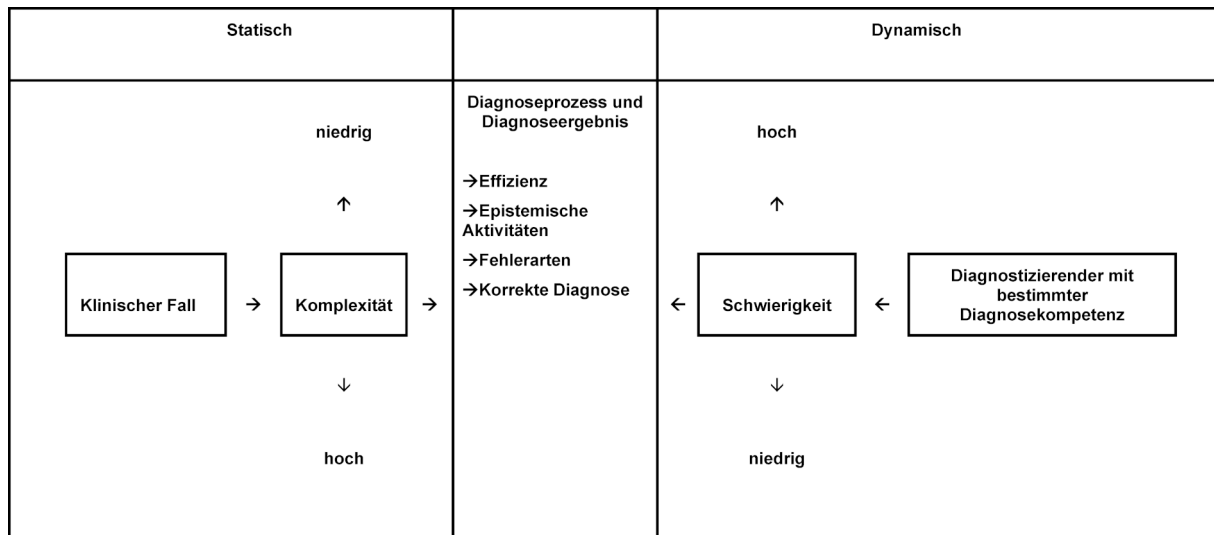


Abbildung 1: Differenzierung: Komplexität und Schwierigkeit

zwar egal, ob es sich um 1-Test oder 2-Test-Fälle handelt, kaum zu lösen. Die Aufgaben sind *schwierig* für Medizinstudierende und *einfach* für Mathematikstudierende. Die Aufgabe an sich hat also einen bestimmten Komplexitätsgrad (1-Test-Fall = simpel/nicht-komplex und 2-Test-Fälle = komplex), die Lösungswahrscheinlichkeit und damit die Schwierigkeit ist allerdings rezipientenabhängig [8]. Das Vorwissen der bearbeitenden Person beeinflusst den Schwierigkeitsgrad viel mehr als die Komplexität.

Komplexität bezieht sich also alleine auf Merkmale des klinischen Falls und ist unabhängig von der diagnostizierenden Person. Er kann als die Vielschichtigkeit des klinischen Problems definiert werden. Schwierigkeit hingegen kann als dynamischer Begriff verstanden werden, da er unmittelbar von der diagnostizierenden Person abhängig ist: Es gibt verschiedene Faktoren, die zur Schwierigkeit beitragen: Neben dem Vorwissen und der Vorerfahrung einer Person [9] – so ist zum Beispiel ein Fall mit der Diagnose Endokarditis wahrscheinlich schwierig für einen Medizinstudierenden im ersten klinischen Jahr, aber möglicherweise sehr einfach für einen Kardiologen mit mehreren Jahren Berufserfahrung – ist auch die Motivation, die Tagesform und letztendlich auch der reine Zufall entscheidend. Hat der Arzt/die Ärztin einen ähnlichen Fall erst vor kurzem so gesehen, dann ist das klinische Problem eventuell sehr einfach für ihn zu lösen [9]. Zur Einschätzung der Fallschwierigkeit liegen schon Vergleichsstudien verschiedener Ratingsysteme vor [10].

Ursachen, die zur Komplexität der Therapie eines Patienten beitragen können, wurden schon in einigen Studien untersucht [11]. Auch die generelle Komplexitätszunahme der ärztlichen Rollen in komplexer werdenden Versorgungssystemen wurde bereits umfassend diskutiert [12]. Es gibt bisher aber kein angewandtes System zur Bestimmung der Komplexität eines klinischen Falls im fallbasierten Lernen. Dadurch ergibt sich die Problematik, dass Studienergebnisse aufgrund unterschiedlicher bzw. nicht klar definierter Fallkomplexität nur mit Einschränkungen vergleichbar sind und auch die Umsetzung der wissenschaftlichen Erkenntnisse in die Lehrpraxis erschwert

wird. Auch ein differenzierter Einsatz von Fällen unterschiedlichen Komplexitätsgrades für verschiedene Expertisestufen kann ohne eine Quantifizierung der Komplexität nicht erfolgen.

In diesem Artikel soll daher die folgende Frage beantwortet werden: Wie kann der Komplexitätsgrad eines klinischen Falles für Simulationsumgebungen praktikabel und standardisiert bestimmt werden?

Methodik: Entwicklung eines literaturbasierten Modells zur Komplexität

In einer Grundlagenstudie zur Entscheidungsfindung in komplexen Szenarien nennt Payne zwei wesentliche Merkmale eines komplexen Systems: „The number of alternatives available“ sowie „number of dimensions of information available“ [13]. Diese Definition sowie eine Klassifikation zur Fallkomplexität, welche nach unseren Recherchen bisher die einzige Klassifikation in der medizinischen oder medizinischen Lehrforschungsliteratur ist, die sich direkt auf die Fallkomplexität in der ärztlichen Ausbildung bezieht, sind Grundlage unseres Komplexitätsmodells: Hennen [14] schlägt einen Score vor, der aus fünf Items besteht: Symptomkomplexität, Komplexität der klinischen und technischen Befunde, sozioökonomische/verhaltensspezifische Schwierigkeiten, Diagnosekomplexität sowie Managementkomplexität. Aus einer Addition dieser Unterkategorien ergibt sich ein Komplexitätsgrad für das gesamte klinische Problem [14]. Obwohl die Kategorisierung an sich viele wichtige Aspekte umfasst, sind die einzelnen Stufen jeder Kategorie eindimensional und gehen nicht auf die im Komplexitätsbegriff benannten Verbindungen der Einzelelemente ein. Weiterhin bewertet dieses Modell [14] die Komplexität der finalen Diagnose sowie die Schwierigkeiten einer angemessenen Therapie. Diese Aspekte spielen bei der

hier angewandten Definition des Diagnostizierens keine Rolle.

Das Modell musste deshalb an die Begriffsbestimmungen des klinischen Diagnostizierens angepasst werden mit einer teilweisen neuen Einteilung der Kategorien, sodass ein allgemein anwendbares Modell für klinische Fälle in simulierten Lernumgebungen entsteht. Im Folgenden wird ein Vorschlag mit fünf Kategorien beschrieben, der das Modell von Hennen an die genannten Erfordernisse anpasst. Methodisch handelt es sich dabei um eine erste Inhaltsvalidierung durch Gruppendiskussion mit Expertenskonsens als Grundlage für Studien zur externen Validität. Als Ergebnis wurden die Kategorien Körperliche Untersuchung und Technische Befunde getrennt, die beiden Kategorien Diagnose- und Managementkomplexität hingegen zusammengefasst. Das ursprünglich dreistufige Matrix-Modell wurde unter Berücksichtigung der Komplexitätsdefinition nach Payne angepasst.

Das modifizierte Scoring besteht aus fünf verschiedenen Kategorien:

1. Anamnese,
2. Körperliche Untersuchung,
3. Technische Befunde,
4. psychosoziale Aspekte und
5. Nebendiagnosen (siehe Tabelle 1).

Ein klinischer Patientenfall besteht in der Regel aus den folgenden drei Grundelementen: der Anamnese, einer körperlichen Untersuchung sowie technischer diagnostischer Befunde (beispielsweise einem Labor, einem Röntgenbild oder einem Elektrokardiogramm (EKG) etc.). Zudem ist bei klinischen Fällen der psychosoziale Kontext, also das Verhalten und die soziale Einbindung des Patienten, wichtig. Nun kann der Komplexitätsgrad auf allen Ebenen variiert werden. Ein Fall nimmt an Komplexität zu, sobald er distrahierende Informationen enthält, welche die Anzahl an Verbindungen zwischen den Einzelelementen des Systems erhöht und damit weitere Differentialdiagnosen wahrscheinlicher macht. Erhöht sich die Anzahl der Komponenten im Modell („number of dimensions of information available“) [13], die dann wiederum miteinander interagieren, erhöht sich die Komplexität.

Wir schlagen ein **5+5-Modell** vor: Je mehr Punkte, desto komplexer, also vielschichtiger ist der Fall. Das Scoring der Fälle umfasst zwei Schritte: Zuerst wird bewertet, ob der Fall Informationen aus allen fünf Kategorien erhält. Ausnahme ist hier die Kategorie Diagnose, hier wird nicht per se ein Punkt vergeben, sondern nur für jede -neben der Hauptdiagnose- weitere Nebendiagnose. Ein Fall erhält einen Punkt in einer Kategorie, sobald eine Information dazu genannt wird. Im zweiten Schritt wird die Linearität der Informationen bewertet: Falls eine Information genannt wird, die nicht linear mit der ersten Information verknüpft ist, dann erhält der Fall einen zweiten Punkt. Unter linear verstehen wir, dass die Information in der gleichen Fallebene zusammenpassen, also keine Mehr-/Vielschichtigkeit hervorrufen. Kommt nun eine dritte (vierte, fünfte) Informationsebene hinzu, die jeweils nicht mit den vorherigen Ebenen verknüpft ist, dann wird wie-

derum ein Punkt vergeben. Der zweite Schritt ist unabhängig von den vorher genannten Kategorien, so könnte beispielsweise ein Fall 3 zusätzliche Punkte in der Kategorie „Anamnese“ erhalten. Allgemein gesagt: Ein neuer Aspekt generiert keinen Punkt, wenn er die Linearität eines Falls nicht aufhebt, also durch die neue Information keine neue Fallebene geschaffen wird. Mehr Punkte werden hingegen vergeben, wenn eine Information nicht linear mit vorherigen Informationen verknüpft ist, sondern eine zusätzliche Komponente darstellt, die mit vielen anderen Komponenten des Systems interagiert. Hierbei kann es sich sowohl um pathologische/auffällige Befunde als auch um Normalbefunde handeln.

Ein Fall kann daher minimal einen Punkt enthalten (nur eine Anamnese) und maximal 10 Punkte. Obwohl theoretisch denkbar, ist es unwahrscheinlich, dass ein Fall, der zur studentischen Lehre eingesetzt wird, mehr als sechs zusätzliche Informationsebenen enthält, die für die klinische Entscheidungssituation in oben genannter Definition des Diagnostizierens relevant sind. Zudem ist davon auszugehen, dass auch bei realen Patientenfällen der Komplexitätsgrad nicht unendlich steigt.

Beispielhaftes Scoring eines klinischen Falls

Im Folgenden soll anhand eines Beispielfalls Pneumothorax das Scoring der Komplexität gezeigt werden. Die Lösungsrichtigkeit dieses Falls lag in einer Studie mit 150 Medizinstudierenden im klinischen Studienabschnitt bei 45% (siehe Anhang 1).

Diskussion und Ausblick

Ziel des Scorings

Der Komplexitätsgrad eines klinischen Falls ist für Fallentwickler elementar, da klinische Fälle bewusst in ihrer Komplexität variiert werden können, um sie für eine bestimmte Zielgruppe passend zu gestalten. Mit zunehmendem Ausbildungsgrad sollten Studierende mit immer komplexeren Fällen und abnehmenden Unterstützungsmaßnahmen konfrontiert werden. Für alle Studien, in denen klinische Fälle verwendet werden, wäre zu fordern: Der Komplexitätsgrad des Falls sollte in einer standardisierten Weise beschrieben werden. Ziel sollte sein, dass Ergebnisse von Interventionsstudien Forschergruppenübergreifend vergleichbar sind.

Komplexität und Schwierigkeit

In weiteren Studien sollte der Einfluss der Komplexität auf die Fallschwierigkeit genauer untersucht werden, wobei die Richtigkeit nicht nur durch Genauigkeit des Diagnoseergebnisses, sondern auch durch andere abhängige Variablen wie Diagnoseeffizienz oder Fehlerarten beschrieben werden sollte [6], [15]. Diagnostizieren ist schwierig. Jede zehnte Diagnose ist eine Fehldiagnose [16], [17]. Generell gibt es zwei Fehlerquellen: mangeln-

Tabelle 1: Zusammenfassung des Scorings

Anamnese	Informationen vorhanden	1 Punkt
Körperliche Untersuchung	Informationen vorhanden	1 Punkt
Technische Untersuchungen	Informationen vorhanden	1 Punkt
Psychosoziale Aspekte	Informationen vorhanden	1 Punkt
Nebendiagnose	vorhanden	1 Punkt

des Wissen und mangelnde Informationsverarbeitung. Der Komplexitätsgrad könnte insbesondere Fehler aufgrund von falscher Informationsverarbeitung bedingen. Dies ist insbesondere wichtig, wenn in einem Fall mehrere Differentialdiagnosen denkbar sind oder der Patient verschiedene klinische Probleme hat, die korrekt gewichtet werden müssen. Inwiefern der Komplexitätsgrad eines klinischen Falls zur Schwierigkeit beiträgt, ist unklar.

Ausblick: Validierung des Modells und Modulation der Komplexität

Ziel dieser Arbeit und der Arbeitsgruppe ist es, ein Instrument zur Bewertung der Fallkomplexität zu entwickeln, um diese wichtige Falleigenschaft korrekt beschreiben und entsprechende Fälle gezielt für Lehre und Forschung einsetzen zu können. Ob das vorgeschlagene Scoring-System die Komplexität klinischer Fälle für simulationsbasierte Lernumgebungen zuverlässig messen kann, muss in Validierungsstudien überprüft werden. Laut dem an Kane angelehnten Validierungsinstrument von Cook et. al. beschreibt dieser Kommentar den Beginn eines Validierungsprozess („articulating the claims and assumptions associated with the proposed decision (the interpretation/use argument)“) [18]. Im weiteren Verlauf muss nun Evidenz generiert werden, denn es gilt: „Just as one can never prove a hypothesis, validity can never be proven; but evidence can, as it accumulates, support or refute the validity argument.“ [19]. Generalisierbarkeit und Praktikabilität des Scoringsystems sollen mit einem mehrstufigen Experten-Review mit Hilfe der Angoff-Methode [20] überprüft werden.

Ziel wäre es, die jetzt 10-teilige Skala in Abschnitte zu gliedern: wenig komplex, mäßig komplex, sehr komplex. In diesem Zusammenhang ist auch zu prüfen, ob bei gleichbleibender Expertisestufe durch Komplexitätsmodulation des Falles die Schwierigkeit beeinflusst wird. Offen bleibt außerdem, ob es weitere Aspekte gibt, die in den Komplexitätsgrad einfließen, von dem Modell allerdings noch nicht erfasst werden. Ob und inwiefern das Modell auch auf andere Themenfelder und Berufsgruppen, beziehungsweise generell auf Problemlöseaufgaben übertragbar ist, ist ein weiterer wesentlicher Aspekt für zukünftige Validierungsstudie.

Für Rückmeldungen anderer Forscher zur Praktikabilität und Validität unseres Bewertungsmodells wären wir dankbar.

Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel haben.

Anhänge

Verfügbar unter

<https://www.egms.de/de/journals/zma/2019-36/zma001288.shtml>

1. Anhang_1.pdf (98 KB)

Beispielhaftes Scoring eines klinischen Falls

Literatur

1. Cabot RC. Case teaching in medicine. Lexington: DC Heath & Company; 1906.
2. Huwendiek S, De Ieng BA, Zary N, Fischer MR, Ruiz JG, Ellaway R. Towards a typology of virtual patients. Med Teach. 2009;31(8):743-748. DOI: 10.1080/01421590903124708
3. Kononowicz AA, Zary N, Edelbring S, Corral J, Hege I. Virtual patients—what are we talking about? A framework to classify the meanings of the term in healthcare education. BMC Med Educ. 2015;15:11. DOI: 10.1186/s12909-015-0296-3
4. Mamede S, Van Gog T, Sampaio AM, De Faria RMD, Maria JP, Schmidt HG. How can students' diagnostic competence benefit most from practice with clinical cases? The effects of structured reflection on future diagnosis of the same and novel diseases. Acad Med. 2014;89(1):121-127. DOI: 10.1097/ACM.0000000000000076
5. Kopp V, Stark R, Fischer MR. Fostering diagnostic knowledge through computer-supported, case-based worked examples: effects of erroneous examples and feedback. Med Educ. 2008;42(8):823-829. DOI: 10.1111/j.1365-2923.2008.03122.x
6. Braun LT, Zottmann JM, Adolf C, Lottspeich C, Then C, Wirth S, Fischer MR, Schmidmaier R. Representation scaffolds improve diagnostic efficiency in medical students. Med Educ. 2017;51(11):1118-1126. DOI: 10.1111/medu.13355
7. McDowell M, Jacobs P. Meta-analysis of the effect of natural frequencies on Bayesian reasoning. Psychol Bull. 2017;143(12):1273-1312. DOI: 10.1037/bul0000126
8. Binder K, Krauss S, Bruckmaier G, Marienhagen J. Visualizing the Bayesian 2-test case: The effect of tree diagrams on medical decision making. PloS One. 2018;13(3):e0195029. DOI: 10.1371/journal.pone.0195029
9. Elstein AS, Shulman LS, Sprafka SA. Medical problem solving an analysis of clinical reasoning. Cambridge (Mass): Havard University Press; 1978. DOI: 10.4159/harvard.9780674189089

10. Koskinen H. Evaluation of the level of difficulty of patient cases for veterinary problem-solving examination: a preliminary comparison of three taxonomies of learning. *J Vet Med Educ.* 2007;34(2):106-111. DOI: 10.3138/jvme.34.2.106
11. Safford MM, Allison JJ, Kiefe CI. Patient complexity: more than comorbidity. The vector model of complexity. *J Gen Intern Med.* 2007;223 Suppl 3:382-390. DOI: 10.1007/s11606-007-0307-0
12. Plsek PE, Greenhalgh T. Complexity science: The challenge of complexity in health care. *BMJ.* 2001;323(7313):625. DOI: 10.1136/bmj.323.7313.625
13. Payne JW. Task complexity and contingent processing in decision making: An information search and protocol analysis. *Organ Behav Human Perform.* 1976;16(2):366-387. DOI: 10.1016/0030-5073(76)90022-2
14. Hennen BK. Measuring the complexity of clinical problems. *Acad Med.* 1984;59(6):487-493. DOI: 10.1097/00001888-198406000-00005
15. Braun LT, Zwaan L, Kiesewetter J, Fischer MR, Schmidmaier R. Diagnostic errors by medical students: results of a prospective qualitative study. *BMC Med Educ.* 2017;17(1):191. DOI: 10.1186/s12909-017-1044-7
16. Graber ML, Carlson B. Diagnostic error: the hidden epidemic. *Physician Exec.* 2011;37(6):12-14, 16, 18-19.
17. Graber ML. The incidence of diagnostic error in medicine. *BMJ Qual Saf.* 2013;22 Suppl 2:ii21-ii27. DOI: 10.1136/bmjqs-2012-001615
18. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ.* 2015;49(6):560-575. DOI: 10.1111/medu.12678
19. Cook DA, Hatala R. Validation of educational assessments: a primer for simulation and beyond. *Adv Simul (London).* 2016;1:31. DOI: 10.1186/s41077-016-0033-y
20. Jalili M, Hejri SM, Norcini JJ. Comparison of two methods of standard setting: the performance of the three-level Angoff method. *Med Educ.* 2011;45(12):1199-1208. DOI: 10.1111/j.1365-2923.2011.04073.x

Korrespondenzadresse:

Leah Theresa Braun
Ludwig-Maximilians-Universität (LMU) München, Klinikum der Universität München, Medizinische Klinik und Poliklinik IV, Ziemssenstr. 1, 80336 München, Deutschland
Leah.Braun@med.uni-muenchen.de

Bitte zitieren als

Braun LT, Lenzer B, Fischer MR, Schmidmaier R. Complexity of clinical cases in simulated learning environments: proposal for a scoring system. *GMS J Med Educ.* 2019;36(6):Doc80. DOI: 10.3205/zma001288, URN: urn:nbn:de:0183-zma0012880

Artikel online frei zugänglich unter

<https://www.egms.de/en/journals/zma/2019-36/zma001288.shtml>

Eingereicht: 15.04.2018

Überarbeitet: 24.11.2018

Angenommen: 22.01.2019

Veröffentlicht: 15.11.2019

Copyright

©2019 Braun et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.