

Replication – The ugly duckling of science?

Götz Fabry^{1,2}

Martin R. Fischer^{3,4}

1 Albert-Ludwig-Universität
Freiburg, Abt. für Med.
Psychologie, Freiburg/Brg.,
Deutschland

2 GMS Zeitschrift für
Medizinische Ausbildung,
stellv. Schriftleiter, Erlangen,
Deutschland

3 Klinikum der Ludwig-
Maximilians-Universität
München, Institut für
Didaktik und
Ausbildungsforschung in der
Medizin, München,
Deutschland

4 GMS Zeitschrift für
Medizinische Ausbildung,
Schriftleiter, Erlangen,
Deutschland

Editorial

This August “Science” published a much-noticed paper by a collective of authors demonstrating that the results of many studies in the field of psychology could not be replicated [1]. With great methodological effort the Open Science Collaboration that incorporates 270 scientists from all over the world selected 100 up-to-date experimental studies from three top-ranked psychological journals in order to assign them for replication to designated and qualified research groups. In these replication studies material and instruments from the original studies were used and the authors of these studies were also consulted during the preparation phase. The results are sobering: While 97% of the original studies reported significant results only 36% of the replication studies did so. Furthermore, the reported effect sizes of the replication studies were only half as large as the original ones and even when the original data and the replicated data were analyzed conjointly only 68% of the results turned out to be significant.

What do these findings mean? First of all the problem itself is known for a long time and is not restricted to psychology. Quite recently, we witnessed a sometimes heated discussion circled around the question whether too much “research waste” is produced in the biomedical sciences [2]. As a matter of fact, the replication of many studies in this field fails, even if they are published in top-ranked

journals. It is estimated that the proportion of non-replicable studies within the biomedical domain is actually larger than in psychology (approximately 75–90%) and even frequently cited studies make no exception [3].

Replication in the educational sciences

Against this background one must strongly assume that medical education research is neither exempt from this calamity. Despite the fact that no empirical findings exist to know this for sure, there is some evidence suggesting an urgent need for action here too. A recent study published by the Educational Researcher – the organ of the American Educational Research Association (AERA) – inquired how often replication studies are published in the domain of educational sciences and what kind of evidence they provide [4]. An analysis of all studies published over the period of five years in the 100 top-ranked educational science journals revealed that the proportion of replication studies was 0.13% (221 of 164 589) only. Remarkably, almost two thirds of these studies replicated the results of the original studies. This relatively large proportion however is put into perspective by the fact that more than half of the replication studies were published by the same authors who were also responsible for the original studies. When only those studies were analyzed that had no overlap of authors the proportion of successful replications declined to about 50%. Thus, if

the different approaches are taken into account we see approximately the same picture here as in the current publication on psychological studies, which comes as no surprise given the close proximity of both domains. Furthermore, the psychological replication study provides some additional evidence that also in medical education not everything that glitters is gold. Successful replication there was more likely when the original P values were smaller (i.e. stricter than .05) and the effect sizes were larger. Unfortunately, both conditions are rather hard to find in medical education studies [5]. In addition, successful replication was less likely when studies used rather complex procedures, which in turn are quite common in educational studies [6]. These findings suggest the alarming assumption that many insights from medical education research would not withstand a more specific inquiry. Thus, do we need more replication studies in medical education? Why are these studies so rare and what has to change [7]?

Replication is more than just history repeating

Some might be thinking back here to the latest reviewer comments that criticized the submitted manuscript for being not interesting enough and just repeating what is already known. How does this critique align with the demand for more replication studies? To answer that question it is necessary to take a closer look on the function and characteristics of replication studies.

Overall, replication studies are done to verify scientific evidence. In a review article on replication in the social sciences that is definitely worth reading, Schmidt explained this function in greater detail [8]: Controlling for sampling error and chance (e.g. due to selection bias), insufficient internal validity (e.g. due to intervening variables, regression to the mean, testing effects, etc.) and fraud. Furthermore, replication studies can be used to clarify whether the results of a certain study can be generalized to larger or other populations or to test the hypotheses of the primary study. Considering the function of replication studies more specifically is important because the study design is determined by the purpose of the study [8]. This is especially true because a successful replication study cannot be a simple “clone” of the primary study. On the one hand this would not be possible at all in typical medical education studies (or generally in psychological or social science studies) as they involve individuals as participants as well as researchers and both cannot be identical at two different points in time. On the other hand, an identical copy of the primary study would not make sense because a verification of research evidence usually requires just to replicate the results at a different point in time at a different place by a different person to accumulate evidence in favor of the transferability and generalizability of the effects found in the primary study. Thus, when designing a replication study it is of utmost importance to reflect precisely which aspects should be held identical and which ones are to be changed in order to gain significant and meaningful res-

ults. If, for instance, the study aims at controlling for sampling error or chance it will be necessary that variables and context are as identical as possible to the primary study while the study sample will be different. Typically, this happens when the same researcher replicates a survey or an experiment with a different or a larger sample. If, in contrast, a replication study is done because of doubts regarding the internal validity of the primary study, it will be necessary to replicate the intervention or the measuring procedure as exactly as possible while all other context variables will be different. This usually happens when a different researcher repeats a study with a different sample at a different place under different circumstances. Thus, all studies that are primarily controlling for chance, internal validity, fraud or generalizability might be described as *direct replications*, because certain aspects of the experimental and context conditions of the primary study are repeated as precisely as possible [8].

Things are different however, with studies testing hypotheses. These studies explicitly search for alternative experimental or methodological approaches to gain additional evidence to support the respective construct. Thus, these studies might be described as *conceptual replications* [8]. They are especially relevant since they contribute to completing a theory or to a more comprehensive understanding of constructs or concepts [9]. However, a drawback of these studies is that a failure of the replication does not allow to conclude that this is due to flaws or biases in the primary study as these can only be revealed by direct replications [8]. Thus, a meaningful replication of insights that have already been described elsewhere distinguishes itself by the fact that the replication is the a priori aim of the study and – in accordance with that – that its function is carefully considered against the conceptual background and the preexisting evidence.

Duckling or Swan?

Is replication then really the ugly duckling of science that looks rather grey and remains so compared to studies that promise innovative insights? Taking the perspective of the individual researcher it seems that the answer is yes since replication studies are unattractive at most regarding publication prospects and frequency of citations. The review regarding the educational sciences mentioned above reported that 43 of the 100 top journals did not publish any replication study. With regard to citation frequency the original studies yielded 31, the replication studies just 5 citations on average (4). Even if this might at least partly be explained by the fact that the replication studies were published some time after the original studies, the difference remains vast.

However, the perspective of the scientific community on replication studies is different. The structure of a paper already reminds us that replication is one of the defining core principles of science. Good scientific papers delineate the background, methods and results transparently so that other researchers can reconstruct the study not

just *in sensu* but – at least as a matter of principle – also *in vivo* (whether we live up to this standard remains open, cf. [10]). Current developments in science already take it a step further so that not just the publications are openly accessible but also the underlying primary data by depositing them in designated repositories. By this means, they can be reviewed independently anytime. This is also possible for publications in our journal because as a part of the platform GMS we are a member of Dryad, an international repository for research data (more information cf. [11]). Against this background replication studies are the touchstone that determines whether the noble principles of science withstand a tangible reality check. While innovations illustrate what might be possible, replications point out what is likely or valid. Scientific progress needs both. Thus, we should not wait any longer to turn the replication duckling into a swan – and that also holds true for medical education research. However, to facilitate this, the incentives for publishing sound replication studies must change. The GMS Journal for Medical Education will contribute to this development by providing a forum for such studies.

Competing interests

The authors declare, that they have no competing interests.

References

- Open Science Collaboration. PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science*. 2015;349(6251):aac4716. DOI: 10.1126/science.aac4716
- Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374(9683):86-89. DOI: 10.1016/S0140-6736(09)60329-9
- Begley CG, Ioannidis JP. Reproducibility in science. Improving the standard for basic and preclinical research. *Circ Res*. 2015;116(1):116-126. DOI: 10.1161/CIRCRESAHA.114.303819
- Makel MC, Plucker JA. Facts are more important than novelty: Replication in the educational sciences. *Educ Res*. 2014;20(10):1-13. DOI: 10.3102/0013189x14545513
- Cook DA, Hatala R. Got power? A systematic review of sample size adequacy in health professions education research. *Adv Health Sci Educ Theory Pract*. 2015;20(1):73-83. DOI: 10.1007/s10459-014-9509-5
- Norman G. RCT= Results confounded and trivial: the perils of grand educational experiments. *Med Educ*. 2003;37(7):582-584. DOI: 10.1046/j.1365-2923.2003.01586.x
- Artino AR Jr. Why don't we conduct replication studies in medical education? *Med Educ*. 2013;47(7):746-747. DOI: 10.1111/medu.12204
- Schmidt S. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Rev Gen Psychol*. 2009;13(2):90-100. DOI: 10.1037/a0015108
- Bordage G. Conceptual frameworks to illuminate and magnify. *Med Educ*. 2009;43(4):312-319. DOI: 10.1111/j.1365-2923.2009.03295.x
- Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: A systematic review. *Med Educ*. 2007;41(8):737-745. DOI: 10.1111/j.1365-2923.2007.02777.x
- Arning U. GMS publishes your research findings – and makes the related research data available through Dryad. *GMS Z Med Ausbild*. 2015;32(3):Doc34. DOI: 10.3205/zma000976

Corresponding author:

Dr. med. Götz Fabry
 Albert-Ludwig-Universität Freiburg, Abt. für Med.
 Psychologie, Rheinstraße 12, 79107 Freiburg/Brg.,
 Deutschland, Tel.: +49 (0)761/203-5512, Fax: +49
 (0)761/203-5514
 goetz.fabry@mps.uni-freiburg.de

Please cite as

Fabry G, Fischer MR. Replication – The ugly duckling of science? *GMS Z Med Ausbild*. 2015;32(5):Doc57.
 DOI: 10.3205/zma000999, URN: urn:nbn:de:0183-zma0009993

This article is freely available from

<http://www.egms.de/en/journals/zma/2015-32/zma000999.shtml>

Received: 2015-10-28

Revised: 2015-10-30

Accepted: 2015-10-30

Published: 2015-11-16

Copyright

©2015 Fabry 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/>.

Replikation – Das hässliche Entlein der Wissenschaft?

Götz Fabry^{1,2}

Martin R. Fischer^{3,4}

1 Albert-Ludwig-Universität
Freiburg, Abt. für Med.
Psychologie, Freiburg/Brg.,
Deutschland

2 GMS Zeitschrift für
Medizinische Ausbildung,
stellv. Schriftleiter, Erlangen,
Deutschland

3 Klinikum der Ludwig-
Maximilians-Universität
München, Institut für
Didaktik und
Ausbildungsforschung in der
Medizin, München,
Deutschland

4 GMS Zeitschrift für
Medizinische Ausbildung,
Schriftleiter, Erlangen,
Deutschland

Leitartikel

Im August dieses Jahres erschien in „Science“ der vielbeachtete Artikel eines Autorenkollektivs, in dem gezeigt wurde, dass sich die Befunde vieler Studien aus dem Bereich der Psychologie nicht replizieren lassen [1]. Mit großem methodischen Aufwand hatten die Autoren der Open Science Collaboration, der 270 Wissenschaftler aus aller Welt angehören, aus drei der führenden psychologischen Fachzeitschriften 100 aktuelle experimentelle Studien ausgewählt und von jeweils ausgewiesenen Forschungsgruppen replizieren lassen. In den Replikationsstudien wurden jeweils Material und Instrumente aus den Originalstudien verwendet, außerdem wurden auch deren Autoren in die Vorbereitung mit einbezogen. Die Ergebnisse sind ernüchternd: Während in 97% der Originalstudien statistisch signifikante Ergebnisse gefunden wurden, war das nur in 36% der Replikationsstudien der Fall. Außerdem waren die replizierten Effektstärken nur etwa halb so groß wie die in den Erstveröffentlichungen und selbst wenn die Originaldaten mit den replizierten Daten gemeinsam ausgewertet wurden, konnten nur noch in 68% der Fälle signifikante Ergebnisse festgestellt werden. Was bedeuten diese Erkenntnisse? Zunächst einmal ist das Problem an sich bereits seit langem bekannt und betrifft nicht nur die Psychologie. So gibt es im Bereich

der biomedizinischen Forschung gerade in den letzten Jahren eine teilweise sehr pointiert geführte Debatte darüber, dass zu viel „Forschungsmüll“ produziert wird [2]. Tatsächlich lassen sich auch hier viele, selbst hochrangig publizierte Erkenntnisse nicht reproduzieren: Die Rate der Nicht-Replizierbarkeit im biomedizinischen Bereich liegt Schätzungen zufolge sogar noch höher als in der Psychologie, nämlich bei 75–90% und auch vielzitierte Studien bilden hiervon keine Ausnahme [3].

Replikation in der Bildungsforschung

Vor diesem Hintergrund ist stark zu vermuten, dass auch die Medizinische Ausbildungsforschung von diesem Problem betroffen ist. Genaue Erkenntnisse liegen in Ermangelung entsprechender Studien dazu nicht vor, allerdings gibt es einige Indizien, die dafür sprechen, dass auch hier dringender Handlungsbedarf besteht. So erschien kürzlich im Educational Researcher, dem Organ der American Educational Research Association (AERA) eine Arbeit, die untersuchte, wie häufig und mit welchen Ergebnissen Replikationsstudien im Bereich der Educational Sciences veröffentlicht werden [4]. Eine Analyse aller Artikel, die in einem 5-Jahreszeitraum in den 100 am höchsten gerankten Zeitschriften dieses Bereichs erschienen waren ergab, dass nur 0,13% (221 von 164 589) der Publikationen Replikationsstudien sind. Immerhin zwei Drittel dieser Studien erbrachten dasselbe Ergebnis wie die ursprüng-

lichen Arbeiten. Dieser vergleichsweise hohe Anteil relativiert sich allerdings insofern, als die Hälfte der replizierten Studien von denselben Autoren veröffentlicht worden waren, die auch die Erstpublikation verantworteten. Wurden nur die Studien analysiert, in denen es keine Überlappung der Autorenschaft gab, dann reduzierte sich der Anteil erfolgreicher Replikationen auf etwa die Hälfte. Berücksichtigt man das jeweils unterschiedliche Vorgehen dann zeigt sich hier somit in etwa dasselbe Bild wie auch in der aktuellen Veröffentlichung zu den psychologischen Studien, was angesichts der nahen Verwandtschaft der beiden Bereiche nicht weiter verwunderlich ist. Darüber hinaus liefert auch die psychologische Replikationsstudie noch weitere Hinweise darauf, dass in der medizinischen Ausbildungsforschung vermutlich auch nicht alles Gold ist, was glänzt. Dort zeigte sich nämlich, dass erfolgreiche Replikationen wahrscheinlicher waren, wenn das Signifikanzniveau (p) kleiner war (also strenger als .05) und die Effektstärken größer. Beides sind Bedingungen, die wir in Studien der medizinischen Ausbildungsforschung leider eher selten finden [5]. Weniger erfolgreich replizieren ließen sich außerdem Studien, in denen komplexe Prozeduren verwendet wurden, die wiederum in der Ausbildungsforschung häufiger anzutreffen sind [6]. Angesichts dieser Befunde liegt die beunruhigende Vermutung nahe, dass viele Erkenntnisse der medizinischen Ausbildungsforschung einer genaueren Überprüfung nicht standhalten würden. Brauchen wir also mehr Replikationsstudien in der Medizinischen Ausbildungsforschung? Warum gibt es davon bisher so wenige und was muss sich ändern [7]?

Replikation ist mehr als die Wiederholung von bereits Bekanntem

Manch einer mag sich hier vielleicht an die letzten Review-Kommentare erinnern, in denen das eingereichte Manuskript dafür kritisiert wurde, dass es leider zu wenig interessant sei, weil es lediglich bereits Bekanntes wiederhole. Wie verträgt sich eine solche Kritik mit der Forderung nach mehr Replikationsstudien? Um diese Fragen zu beantworten ist es notwendig, sich genauer mit Funktion und Charakteristika von Replikationen auseinanderzusetzen.

Ganz allgemein dienen Replikationsstudien dazu, wissenschaftliche Erkenntnisse zu verifizieren. Wie Schmidt [8] in einem lesenswerten Übersichtsartikel zu Replikationen in den Sozialwissenschaften dargestellt hat, lassen sich dabei im Detail die folgenden Funktionen unterscheiden: Die Kontrolle von Stichproben- und Zufallsfehlern (z.B. durch eine verzerrte Probandenauswahl), von schlechter interner Validität (z.B. durch Störvariablen, Regression zur Mitte, Testing-Effekte etc.) sowie von Fälschungen. Außerdem können Replikationsstudien dazu dienen, die Generalisierbarkeit der Ergebnisse einer Studie auf andere bzw. größere Populationen zu überprüfen oder die der Ursprungsstudie zugrundeliegenden Hypothesen zu testen. Solche spezifisch funktionellen Überlegungen sind

deshalb wichtig, weil sich das Forschungsdesign nach dem Zweck der Replikationsstudie bestimmt [8]. Dazu muss man sich vor Augen führen, dass eine erfolgreiche Replikationsstudie nicht einfach ein „Klon“ der Originalstudie sein kann. Zum einen wäre das gerade in personenbezogenen Studien, wie sie für die medizinische Ausbildungsforschung (oder psychologische und sozialwissenschaftliche Studien ganz allgemein) typisch sind, überhaupt nicht möglich, da weder Probanden noch Versuchsleiter zu zwei verschiedenen Zeitpunkten vollkommen identisch sein können. Zum anderen wäre die identische Kopie einer Studie aber auch gar nicht sinnvoll, da ein wesentlicher Aspekt der Verifikation von Forschungsergebnissen ja häufig gerade darin besteht, sie zu einem anderen Zeitpunkt, an einem anderen Ort von einer anderen Person wiederholen zu können, um Belege für die Übertragbarkeit und Verallgemeinerbarkeit der gefundenen Effekte zu liefern. Daher ist es entscheidend, genau zu reflektieren, welche Aspekte bei einer Replikationsstudie identisch bleiben und welche verändert werden müssen, um aussagekräftige und interpretierbare Ergebnisse zu erhalten. Geht es beispielsweise darum, auf Stichproben- und Zufallseffekte zu kontrollieren, dann müssen möglichst viele Variablen und Kontextbedingungen mit denen der ursprünglichen Studie übereinstimmen, während die Stichprobe verändert wird. In der Praxis ist das typischerweise dann der Fall, wenn derselbe Wissenschaftler eine Befragung oder ein Experiment an einer anderen oder größeren Stichprobe wiederholt. Wird eine Replikationsstudie dagegen durchgeführt, weil Zweifel an der internen Validität der Ursprungsstudie bestehen, dann muss die eigentliche Intervention bzw. das Messverfahren bei der Replikation möglichst identisch sein während alle anderen Kontextvariablen verändert werden können. Das geschieht in der Regel dann, wenn eine Studie von einem anderen Wissenschaftler mit anderen Probanden an einem anderen Ort unter anderen Umständen wiederholt wird. Alle Studien, bei denen entweder die Überprüfung von Zufallseffekten, interner Validität, Fälschung oder die Frage der Generalisierbarkeit im Vordergrund stehen, kann man daher auch als *direkte Replikationen* bezeichnen, weil sie jeweils bestimmte Aspekte der ursprünglichen Versuchs- bzw. Studienbedingungen möglichst genau wiederholen [8].

Anders dagegen verhält es sich, wenn die Überprüfung von Hypothesen Ziel der Replikation ist. Hier geht es nämlich gerade darum, eine alternative experimentelle Umsetzung oder ein anderes methodisches Vorgehen zu finden, um zusätzliche Evidenz zu dem in Frage stehenden Konstrukt zu produzieren. Solche Studien kann man daher als *konzeptuelle Replikationen* bezeichnen [8]. Sie sind deshalb besonders relevant, weil sie zur Theoriebildung und zum besseren Verständnis von Konstrukten und Konzepten beitragen [9]. Ein Nachteil dieser Studien besteht allerdings darin, dass bei einem Scheitern der Replikation keine Rückschlüsse darüber möglich sind, ob dies auf Schwächen oder Verzerrungen in der ursprünglichen Studie zurückgeführt werden kann, denn diese

lassen sich nur mittels einer direkten Replikation aufdecken [8].

Eine sinnvolle Wiederholung von bereits anderswo beschriebenen Erkenntnissen zeichnet sich somit vor allem dadurch aus, dass die Replikation bereits a priori das Ziel der Studie ist und damit auch genaue Überlegungen angestellt worden sind, welche Funktion vor dem Hintergrund der jeweiligen konzeptuellen Aspekte und der bereits vorliegenden Daten sinnvoll ist.

Entlein oder Schwan?

Ist die Replikation also das hässliche Entlein der Wissenschaft, das neben Studien, die gänzlich neue Erkenntnisse versprechen, ziemlich grau aussieht und grau bleibt? Aus Sicht des einzelnen Forschers muss man diese Frage wohl bejahen, jedenfalls sind Replikationsstudien im Hinblick auf Publikationsmöglichkeiten und Zitierhäufigkeit maximal unattraktiv. So zeigte sich in der bereits zitierten Übersichtsarbeit zu den Educational Sciences, das 43 der 100 untersuchten Top-Journals keine einzige Replikationsstudie veröffentlicht hatten. Außerdem wurden die ursprünglichen Studien durchschnittlich 31 Mal, die Replikationsstudien dagegen nur fünf Mal zitiert (4). Selbst wenn man berücksichtigt, dass dabei auch eine Rolle spielen mag, dass die Replikationsstudien später als die Primärarbeiten erschienen sind, bleibt dennoch ein großer Unterschied bestehen.

Völlig anders sieht die Bewertung von Replikationen dagegen aus Sicht der Scientific Community aus. Schon der Aufbau eines Artikels erinnert uns daran, dass die Wiederholbarkeit eines der definierenden Kernprinzipien von Wissenschaft ist. Gute wissenschaftliche Publikationen stellen Hintergrund, Methoden und Ergebnisse so transparent dar, dass andere Wissenschaftler sie nicht nur *in sensu* sondern – zumindest prinzipiell – auch *in vivo* nachvollziehen können (ob dieser Anspruch immer eingelöst wird, sei dahingestellt vgl. [10]). Aktuelle Entwicklungen in der Wissenschaft gehen noch darüber hinaus, insofern nicht mehr nur die Publikationen öffentlich zugänglich sind, sondern auch die Primärdaten, weil diese in entsprechenden Repositorien verfügbar gemacht werden. Damit sind sie jederzeit unabhängigen Überprüfungen zugänglich. Auch für Publikationen in der ZMA ist dies möglich, weil wir über unsere Plattform GMS Mitglied bei Dryad, einem internationalen Forschungsdatenspeicher sind (nähere Informationen dazu bei [11]). Vor diesem Hintergrund sind Replikationsstudien somit der Prüfstein, an dem sich entscheidet, ob die hehren Prinzipien der Wissenschaft einer konkreten Realitätsprüfung standhalten.

Während Innovationen zeigen, was möglich ist, zeigen Replikationen was wahrscheinlich oder gültig ist. Wissenschaftlicher Fortschritt kann nur aus beidem entstehen. Wir sollten also nicht mehr lange damit warten, aus dem Replikationsentlein einen Schwan zu machen – auch in der Medizinischen Ausbildungsforschung. Dafür muss sich allerdings auch das Anreizsystem zur Publikation von gut begründeten Replikationsstudien ändern. Die

GMS Zeitschrift für Medizinische Ausbildung wird dazu ihren Beitrag leisten und solchen Replikationsstudien ein Forum bieten.

Interessenkonflikt

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

Literatur

1. Open Science Collaboration. PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science*. 2015;349(6251):aac4716. DOI: 10.1126/science.aac4716
2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374(9683):86-89. DOI: 10.1016/S0140-6736(09)60329-9
3. Begley CG, Ioannidis JP. Reproducibility in science. Improving the standard for basic and preclinical research. *Circ Res*. 2015;116(1):116-126. DOI: 10.1161/CIRCRESAHA.114.303819
4. Makel MC, Plucker JA. Facts are more important than novelty: Replication in the educational sciences. *Educ Res*. 2014;20(10):1-13. DOI: 10.3102/0013189x14545513
5. Cook DA, Hatala R. Got power? A systematic review of sample size adequacy in health professions education research. *Adv Health Sci Educ Theory Pract*. 2015;20(1):73-83. DOI: 10.1007/s10459-014-9509-5
6. Norman G. RCT= Results confounded and trivial: the perils of grand educational experiments. *Med Educ*. 2003;37(7):582-584. DOI: 10.1046/j.1365-2923.2003.01586.x
7. Artino AR Jr. Why don't we conduct replication studies in medical education? *Med Educ*. 2013;47(7):746-747. DOI: 10.1111/medu.12204
8. Schmidt S. Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Rev Gen Psychol*. 2009;13(2):90-100. DOI: 10.1037/a0015108
9. Bordage G. Conceptual frameworks to illuminate and magnify. *Med Educ*. 2009;43(4):312-319. DOI: 10.1111/j.1365-2923.2009.03295.x
10. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: A systematic review. *Med Educ*. 2007;41(8):737-745. DOI: 10.1111/j.1365-2923.2007.02777.x
11. Arning U. GMS publishes your research findings – and makes the related research data available through Dryad. *GMS Z Med Ausbild*. 2015;32(3):Doc34. DOI: 10.3205/zma000976

Korrespondenzadresse:

Dr. med. Götz Fabry
 Albert-Ludwig-Universität Freiburg, Abt. für Med.
 Psychologie, Rheinstraße 12, 79107 Freiburg/Brg.,
 Deutschland, Tel.: +49 (0)761/203-5512, Fax: +49
 (0)761/203-5514
 goetz.fabry@mps.uni-freiburg.de

Bitte zitieren als

Fabry G, Fischer MR. Replication – The ugly duckling of science? GMS
Z Med Ausbild. 2015;32(5):Doc57.
DOI: 10.3205/zma000999, URN: urn:nbn:de:0183-zma0009993

Eingereicht: 28.10.2015

Überarbeitet: 30.10.2015

Angenommen: 30.10.2015

Veröffentlicht: 16.11.2015

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/zma/2015-32/zma000999.shtml>

Copyright

©2015 Fabry 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/>.