

Rundgespräch „Die Gewährleistung von Replizierbarkeit und die Unterstützung von Replikation in den DFG-Förderformaten in der Biomedizin und Psychologie“

am 15.12.2015 in Bonn

Seitdem mehrere Studien auf eine geringe Replizierbarkeit von Forschungsergebnissen in der Biomedizin und Psychologie (z.B. Studie der Open Science Collaboration, 2015) hindeuten, findet in der Wissenschaft (z.B. in der 2014 erschienenen Lancet-Serie „increasing value, reducing waste“), aber auch in den Medien und der Gesellschaft, eine zunehmend international geführte Diskussion über mögliche Ursachen und Lösungsvorschläge statt. Im DFG-Rundgespräch sollten die Diskussionen in der Medizin und Psychologie zusammengeführt und Handlungsoptionen abgeleitet werden.

Die Teilnehmenden haben sich im Verlauf des Rundgesprächs auf folgende Definitionen der Begriffe Reproduktion, Replikation sowie Reproduzierbarkeit und Replizierbarkeit geeinigt

- *Reproduktion* meint eine Beobachtung, die konsistent mit Erwartungen ist, die aufgrund von empirischer Evidenz aufgestellt wurden.
- Bei einer *Replikation* wird eine Studie, die auf Reproduktion abzielte, wiederholt. Es wird überprüft, ob sich das Studienergebnis der Originalstudie in einer neu gezogenen Stichprobe wieder im statistischen Test zeigt. Diese Replikation kann entweder eine *direkte* (Ziel ist die möglichst genaue Wiederholung der Originalstudie) oder eine *konzeptuelle* (Ziel ist die Generalisierung des Effektes auf andere Konstrukte oder Modelle) sein.
- *Reproduzierbarkeit* bzw. *Replizierbarkeit* meint entsprechend die Möglichkeit bzw. Fähigkeit, empirische Ergebnisse wiederholend zu bestätigen.

Zentrale Ergebnisse der Diskussion:

Geringe Replizierbarkeit bedeutet nicht zwangsläufig schlechte Qualität der Forschungsergebnisse und umgekehrt ist Replizierbarkeit keine Garantie für gute Forschung. Auch explorative, riskante Forschung ist wichtig, bei welcher grundsätzlich noch nicht alle relevanten Einflussfaktoren bekannt sein können. Dennoch steht fest, dass eine intensivere Diskussion über notwendige Standards für qualitativ hochwertige Forschungsergebnisse eine inhärente und unabdingbar wichtige Selbstverpflichtung der Wissenschaft ist. Ein hinreichend selbstkritischer Blick ist eine Voraussetzung für das Vertrauen der Öffentlichkeit gegenüber der Wissenschaft.

Die Ursachen für eine unzureichende Replizierbarkeit von Ergebnissen sind vielfältig und stellen zweifellos ein Risiko für die hohen Qualitätsstandards der Wissenschaft dar. Besonders relevant erscheinen die Planungsphase von Projekten und deren Darstellung in Publikationen bzw. Forschungsanträgen. Fundierte Überlegungen zum Datenmanagement, verbunden mit der Nutzung anerkannter Datenrepositionen, hinreichend ausführliche Angaben zur Methodik und Statistik und die Reflektion der Validität von Vorarbeiten könnten die Replizierbarkeit von Ergebnissen erhöhen.

Geringe Replizierbarkeit wird aber auch von übergeordneten Faktoren verursacht. Zentral erscheint hier der vorherrschende Innovationsdruck, der begleitet ist von unzureichenden Förder- und Publikationsmöglichkeiten für Replikationsstudien und für Studien mit nicht signifikanten oder den Hypothesen widersprechenden Ergebnissen. In der Folge können sich Forschende dazu verleiten fühlen, nur selektiv von solchen Ergebnissen in Publikationen zu berichten, die die eigene Hypothese stützen (sog. p-hacking) oder ihre Hypothesen im Nachhinein an die Ergebnisse anzupassen (sog.

HARK-ing). Beide Praktiken führen zu einer geringen Replizierbarkeit und somit zu einem unvollständigen Blick auf den Erkenntnisstand. Zusätzlich wurde festgestellt, dass auch gesetzliche Regelungen die Durchführung von Replikationen erschweren. So stehen die aktuell geltenden Tierschutzrichtlinien einer Wiederholung von Ergebnissen mit tierexperimentellem Anteil im Wege.

Die Replizierbarkeit von Forschungsergebnissen wird von zahlreichen Akteuren beeinflusst und kann nur durch deren Zusammenarbeit wirkungsvoll verändert werden. Neben dem notwendigen Diskurs in der Wissenschaft, gibt es Wechselwirkungen mit der Forschungsförderung, den wissenschaftlichen Institutionen, Verlagen, Datenzentren/Repositorien oder anderen wissenschaftliche Infrastrukturen. Auch die Anforderungen von Industrie und Öffentlichkeit wirken auf die Replizierbarkeit von Forschungsergebnissen zurück.

Für die DFG als Förderorganisation ergeben sich unterschiedliche Ansatzpunkte für die Unterstützung und Begleitung des wissenschaftlichen Diskurses. Grundsätzlich sollte die DFG durch ihr Förderhandeln dazu beitragen, die Replizierbarkeit von Forschungsbefunden zu erhöhen. Replikationsstudien sollten grundsätzlich förderwürdig sein, um die Replizierbarkeit wichtiger Forschungsbefunde zu validieren. Während die Schaffung eines eigenständigen Förderbereichs für Replikationen nicht empfohlen wird, sollten die Rahmenbedingungen für bereits bestehende Fördermöglichkeiten ausformuliert und stärker bekannt gemacht werden. Darüber hinaus sollte im Begutachtungs- und Bewertungsprozess ein verstärktes Bewusstsein für die Themen Replizierbarkeit und Replikation geschaffen werden. Um über die Förderwürdigkeit von Replikationsstudien, insbesondere im Verhältnis zu Vorhaben, die neue Fragestellungen untersuchen, entscheiden zu können, sollten spezifische Kriterien definiert werden. Diese Kriterien werden in der Regel fachspezifisch zu entwickeln sein. Als Orientierung können hierbei die vom Fachkollegium „Psychologie“ zu diesem Zweck erarbeiteten Kriterien dienen. Im Abschlussbericht von Forschungsprojekten sollte verstärkt auch die Replizierbarkeit der Ergebnisse beleuchtet und bewertet werden. Die Präregistrierung von Klinischen Studien stellt einen wichtigen Schutz gegen die Inflation falsch-positiver Befunde dar und wäre auch für andere Projekte der Grundlagenforschung empfehlenswert. Grundsätzlich wirkt sich die Nutzung von existierenden Datenzentren/Repositorien positiv auf die Planungsphase und die Sorgfalt des Datenmanagements aus und sollte weiter durch die DFG unterstützt werden.

Der mit der Veranstaltung angestoßene wichtige Diskussionsprozess sollte in der DFG selbst und mit anderen relevanten Akteuren weitergeführt und vertieft werden.