

Deutsche Ausgabe: DOI: 10.1002/ange.201606591
Internationale Ausgabe: DOI: 10.1002/anie.201606591

Reproduzierbarkeit in der chemischen Forschung

Robert G. Bergman* und Rick L. Danheiser*



Robert G. Bergman
Gerald E.K. Branch
Distinguished Professor
University of California,
Berkeley



Rick L. Danheiser
A.C. Cope Professor
of Chemistry
Massachusetts Institute
of Technology

Reproduzierbarkeit ist ein zentrales Merkmal aller Wissenschaft. In letzter Zeit aber wurde immer wieder besorgt gefragt, inwieweit Forschungsergebnisse – vor allem in der Biomedizin – noch einfach reproduziert werden können. In welchem Ausmaß ist Reproduzierbarkeit ein Thema in der *chemischen Forschung*? Wie kann man Probleme mit der Reproduzierbarkeit in der Chemie minimieren?

Nichtreproduzierbarkeit

Probleme mit der Reproduzierbarkeit können sehr schwerwiegend, aber auch eher geringfügig sein. In der Chemie kommt es sehr selten vor, dass veröffentlichte Ergebnisse überhaupt nicht reproduzierbar sind, dagegen ist es keineswegs unüblich für Chemiker, Schwierigkeiten dabei zu haben, Reaktionsausbeuten und -selektivitäten sowie andere in Veröffentlichungen beschriebene Daten genau zu reproduzieren. Solche Nichtreproduzierbarkeiten sind nur allzu verbreitet, vor allem in der Synthesechemie.

Wir sind davon überzeugt, dass eine der Quellen für Nichtreproduzierbarkeit – das bewusste Fälschen von Daten – in der Chemie zwar nicht völlig unbekannt, aber doch selten ist. Häufiger auftretende Beispiele für Probleme mit

der Reproduzierbarkeit sind Fälle, in denen Forscher fehlerhafte Daten für richtig hielten oder sie bewusst oder unbewusst ihren Erwartungen anpassen. Eine dritte Ursache für Nichtreproduzierbarkeit sind in gutem Glauben veröffentlichte Ergebnisse („Bona-fide-Ergebnisse“), bei denen es einfach nicht gelingt, sie auch in anderen Labors zu erhalten.

Das bewusste Fälschen von Daten kann auf mehrere Arten passieren, nicht nur durch komplettes Erfinden. Wie erheblich ist das Problem, dass Daten modifiziert werden, um die Interpretation der Ergebnisse durch den Forscher zu stützen? Etwas Material dazu lieferte die Redaktion der Zeitschrift *Organic Letters*, die festgestellt hat, dass 2–3% der eingereichten Manuskripte Indizien aufweisen, die für das manuelle Entfernen von Signalen aus NMR-Spektren sprechen; diese Zahl ist in den wenigen Jahren, seit nach solchen Eingriffen in Spektren gesucht wird, nicht gesunken. Verwandte Verstöße gegen die gute wissenschaftliche Praxis sind das Verwerfen von Datenpunkten, die nicht zu den erwünschten Ergebnissen passen, und die ausschließliche Angabe der besten Ausbeute oder Selektivität einer wiederholt durchgeföhrten Synthese.

[*] Prof. R. G. Bergman
Department of Chemistry
University of California
Berkeley, CA 94720-1460 (USA)
E-Mail: rbergman@berkeley.edu
Prof. R. L. Danheiser
Department of Chemistry
Massachusetts Institute of Technology
Cambridge MA 02139 (USA)
E-Mail: danheisr@mit.edu

Unbewusste Befangenheit

Bei den am breitesten diskutierten Fällen nicht reproduzierbarer Ergebnisse in der Chemie geht es um Forscher, die wirklich glaubten, ihre Ergebnisse seien richtig. Berühmte Beispiele aus

der Chemie und anderen Disziplinen sind die gefeierte Behauptung der „kalten Fusion“ durch Pons und Fleischmann und die Homöopathiestudie von Benveniste, nach der eine Lösung, die im Mittel weniger als ein Molekül einer bestimmten Antikörper-Präparation enthielt, die Degranulierung basophiler Zellen auslösen sollte.

Diese Beispiele unterstreichen einen der irritierendsten Aspekte der Nichtreproduzierbarkeit in den Naturwissenschaften, über den unserer Meinung nach zu wenige Wissenschaftler nachdenken: die unbewusste Befangenheit des Forschers. Eine besonders aufschlussreiche Studie zu diesem Thema führte vor vielen Jahren der Psychologe Robert Rosenthal durch. Einer Gruppe von Experimentatoren wurde erzählt, die von ihnen in einem Labyrinth-Experiment eingesetzten Ratten seien „Labyrinth-schlau“ (gut trainiert), einer zweiten Gruppe wurde gesagt, sie hätten „Labyrinth-dumme“ (untrainierte) Ratten erhalten. In Wirklichkeit waren die Ratten rein zufällig den beiden Gruppen zugeordnet worden. Dennoch erledigten die Ratten, von denen die Experimentatoren annahmen, sie seien schlauer, die Aufgabe besser. Die Ratten unterschieden sich nur hinsichtlich der Erwartungen der Experimentatoren an sie.

Die Forscherbefangenheit hat es inzwischen in psychologischen und soziologischen Kreisen zu einer Art Beirühmtheit gebracht, vor allem wenn eine schlechte statistische Datenanalyse das Problem erheblich verschärft. Unserer

Meinung nach gibt es dieses Problem aber auch in den „härteren“ Naturwissenschaften. Neben den beiden oben genannten Fällen kennen die meisten Chemiker auch andere, in denen Forscher hartnäckig experimentelle Befunde und/oder Interpretationen oder Theorien verteidigten, obwohl die Forscherwelt sie schon lange als falsch erkannt hatte.

Richtige Ergebnisse reproduzieren

Fälschung und Fehler sind zwei Gründe für Sorgen, doch wir sind der Meinung, dass die Mehrzahl der Reproduzierbarkeitsprobleme, auf die Chemiker stoßen, mit Bona-fide-Ergebnissen zu tun haben, d.h. mit Ergebnissen, die von den Autoren tatsächlich erhalten wurden. Ergebnisse in der organischen, metallorganischen und anorganischen Synthesechemie zu reproduzieren stellt selbst erfahrene und geschickte Forscher immer wieder vor Probleme. Wie schwierig es sein kann, Ergebnisse in der organischen Synthesechemie zu reproduzieren, zeigt die Erfahrung von einem von uns als Chefredakteur von *Organic Syntheses*. Diese Zeitschrift ist in einer Hinsicht einzigartig: Jedes experimentelle Ergebnis muss im Labor eines der angesehenen Mitglieder des Board of Editors vor der Veröffentlichung reproduziert werden. Zwischen 2010 und 2016 wurden 7.5% der bei *Organic Syntheses* eingereichten Manuskripte abgelehnt, weil die Ausbeute und/oder die Selektivität, die die Autoren angaben, nicht innerhalb einer akzeptablen Toleranz im Labor eines der Editoren reproduziert werden konnten. Da die Autoren wissen, dass ihre Angaben durch einen der Editoren von *Organic Syntheses* überprüft werden, kann man davon ausgehen, dass sie die beschriebenen Ergebnisse tatsächlich erhalten hatten und sicher waren, dass ihre Arbeit reproduzierbar ist. Zudem muss man wissen, dass *Organic Syntheses* außergewöhnlich umfangreiche experimentelle Details verlangt und dass bei während der Überprüfung auftretenden Problemen die Editoren die Autoren um Hilfe bitten. Die Tatsache, dass sich trotz all dieser Vorteile die Ergebnisse in einem von dreizehn Manuskripten als nicht reproduzierbar erwiesen, unter-

streicht die Herausforderungen, die mit der Reproduzierbarkeit in der Synthesechemie verbunden sind.

Empfehlungen

Welche Maßnahmen kann die Chemiekommende ergreifen, um die Reproduzierbarkeit veröffentlichter Forschung zu verbessern? Im Folgenden stellen wir Empfehlungen für die Leiter von Forschungsgruppen und deren Mitarbeiter, für Herausgeber und Redakteure von Zeitschriften sowie für Gutachter vor.

Forschungsgruppenleiter sollten in einem Austausch mit ihren Mitarbeitern sein – durch regelmäßige persönliche Gespräche über die Forschung, bei denen Primärdaten und Laborprotokolle durchgegangen werden. Auch bei Zusammenarbeiten, die einen breiten Bereich von Forschungsgebieten umfassen, sollte jeder einzelne Forschungsgruppenleiter sicherstellen, dass die Arbeit in seinem Bereich gewissenhaft durchgeführt wird. Auch wenn das Anwerben von Fördermitteln ein wesentlicher Beitrag zu einem Projekt ist, sollte Coautorenschaft zusätzlich die direkte Beteiligung an der Studie und einen signifikanten intellektuellen Beitrag zur Arbeit erfordern.

Kluge Skepsis gegenüber den eigenen Ergebnissen ist besonders wichtig. Forschungsgruppenleiter, die ihre Mitarbeiter (und sich selbst!) darin schulen, Ergebnissen, vor allem solchen, die sie für richtig halten wollen, mit Skepsis zu begegnen, sind gute Vorbilder. In manchen Forschungsdisziplinen ist es üblich, dass Gruppenleiter zentrale Ergebnisse intern prüfen lassen, bevor sie eine Veröffentlichung einreichen. So könnte bei einer neuen Synthesemethode ein Mitarbeiter ohne Erfahrung auf diesem Gebiet damit beauftragt werden, ein repräsentatives Beispiel nur anhand der zur Veröffentlichung vorgesehenen Versuchsbeschreibungen zu wiederholen. Damit könnte sichergestellt werden, dass diese Beschreibungen ausreichen, um Reproduzierbarkeit zu garantieren.

Wenn ein Mitarbeiter ein mögliches Problem mit der Reproduzierbarkeit

anspricht, sollte der Gruppenleiter der Versuchung widerstehen, den „Boten hinzurichten“; er sollte vielmehr die Bedenken derer, die es wagten, sich zu äußern, sorgfältig erwägen. Wurden falsche Ergebnisse veröffentlicht, sollte nicht alle Schuld auf den betroffenen Mitarbeiter geschoben werden – alle Coautoren, einschließlich des Forschungsgruppenleiters, sind für veröffentlichte Ergebnisse verantwortlich.

Wir möchten dazu auffordern, dass Hintergrundinformationen mit aussagekräftigen experimentellen Details für alle Veröffentlichungen verlangt werden. Die angesehenen Chemiezeitschriften haben genaue Vorgaben für die Charakterisierung neuer Verbindungen, und es würde ihnen gut zu Gesicht stehen, ihre Vorgaben im Hinblick auf experimentelle Details zu erweitern (oder zu ergänzen, sollte es derzeit hierzu noch keine spezifischen Vorgaben geben). Zudem sollten Gutachter vor allem auch die experimentellen Details bewerten. Alle Redaktionen sollten eingereichte Manuskripte auf Datenmanipulationen überprüfen.

Die Wissenschaftsgemeinschaft sollte zum Veröffentlichen in Zeitschriften ermutigen, die hohe experimentelle Standards pflegen, um so die Wahrscheinlichkeit für Reproduzierbarkeit zu erhöhen. Dafür könnte das übliche Belohnungssystem genutzt werden (Fördermittel, Beförderung, Preise).

Last but not least sollten Gutachter nicht wie römische Kaiser nur durch Heben oder Senken ihres Daumens anzeigen, ob ein Beitrag ausreichend neuartig und wichtig ist, um in einer bestimmten Zeitschrift veröffentlicht zu werden, oder nicht. Sie müssen bei Manuskripten gewissenhaft prüfen, ob die experimentellen Details adäquat sind und die Folgerungen stützen. Die Veröffentlichung mehrerer Arbeiten mit gefälschten Ergebnissen in letzter Zeit wäre zu vermeiden gewesen, hätten die Gutachter die Manuskripte genauer analysiert.

Zitierweise:

Angew. Chem. Int. Ed. **2016**, *55*, 12548–12549
Angew. Chem. **2016**, *128*, 12736–12737