

KẾT QUẢ TỔNG QUAN CÁC NGHIÊN CỨU QUỐC TẾ VỀ ĐÁNH GIÁ NĂNG LỰC THỰC NGHIỆM CỦA HỌC SINH

Lê Thị Xuyến¹, Phạm Xuân Quế²

Tóm tắt: Tổng quan tài liệu là bước làm rất quan trọng trong một quá trình nghiên cứu nói chung. Để đề xuất các biện pháp phát triển và đánh giá năng lực thực nghiệm của học sinh chúng tôi tiến hành việc tổng quan các tài liệu liên quan. Bài báo trình bày những nội dung tổng quan được liên quan đến các nghiên cứu về đánh giá năng lực thực nghiệm của học sinh qua các công bố quốc tế.

Từ khóa: tổng quan, năng lực thực nghiệm

1. MỞ ĐẦU

Đánh giá tổng quan là bước quan trọng để hỗ trợ người nghiên cứu xác định các vấn đề tiếp theo cần nghiên cứu và đưa ra các quyết định phù hợp [13, 18]. Đây gần như là bước phải được làm đầu tiên khi nhà nghiên cứu muốn bắt tay vào quá trình nghiên cứu của mình. Năng lực thực nghiệm là một năng lực đặc thù cần thiết được hình thành trong dạy học vật lý nói riêng và dạy học các môn Khoa học nói chung [2]. Vì vậy việc nghiên cứu các biện pháp phát triển hay đánh giá năng lực này mang ý nghĩa thực tiễn. Để có những định hướng tốt cho nghiên cứu của mình về đánh giá năng lực thực nghiệm ở học sinh lớp 10, chúng tôi tiến hành nghiên cứu tổng quan các tài liệu liên quan đến đánh giá năng lực này trên các kết quả đã được công bố thế giới. Nội dung trình bày bên dưới là cụ thể những kết quả mà chúng tôi đạt được.

2. NỘI DUNG NGHIÊN CỨU

2.1. Mục tiêu và phương pháp tổng quan tài liệu

- Mục tiêu tổng quan tài liệu để trả lời các câu hỏi: Bằng cách nào, công cụ nào có thể đánh giá năng lực thực nghiệm của học sinh? Những kết quả đánh giá năng lực thực nghiệm của học sinh cho thấy điều gì?

- Công cụ tìm kiếm các tài liệu tổng quan mà chúng tôi sử dụng chủ yếu trang web sau: <https://www.researchgate.net/>

- Tiêu chí: ưu tiên lựa chọn báo, sách, luận án, luận văn. Lựa chọn những ấn phẩm từ 1990 đến nay.

¹ Trường Đại học Sư phạm Hà Nội 2

² Trường Trường Đại học Sư phạm Hà Nội

2.2. Các nghiên cứu về các loại hình đánh giá

Khi tìm hiểu về các loại hình đánh giá được sử dụng trong đánh giá năng lực thực nghiệm của học sinh. Chúng tôi nhận thấy có hai loại hình được sử dụng chính là đánh giá quá trình và đánh giá dựa trên sản phẩm. Ở quy mô lớn, các đánh giá năng lực thực nghiệm thường chỉ dựa trên các sản phẩm của thực nghiệm, chủ yếu là ghi chép trong phòng thí nghiệm (ví dụ như Stebler, Reusser, & Ramseier 1998, Ramseier, Labudde & Adamina 2011, Gut 2012, Shavelson 1999) [9, 11, 23, 27]. Theo phân tích của Shavelson thì những ghi chép của học sinh trong phòng thí nghiệm có thể giúp phản ánh năng lực thực nghiệm của học sinh [23]. Chúng tôi cho rằng chỉ đánh giá dựa trên sản phẩm sẽ không cho chúng ta thấy được người học đã nỗ lực thế nào trong các hoạt động thuộc quá trình thực hiện thí nghiệm như: lắp ráp các thiết bị thí nghiệm, hiệu chỉnh những dụng cụ thí nghiệm, thậm chí sửa chữa các dụng cụ thí nghiệm và thu thập các dữ liệu thực nghiệm. Khi tiến hành so sánh giữa đánh giá quá trình và đánh giá kết quả với mỗi thành tố trong năng lực thực nghiệm thì nhóm tác giả *Nico Schreiber*¹, *Heike Theyßen*¹ and *Horst Schecker* [20, 26] chứng minh được rằng đánh giá sản phẩm (đánh giá chỉ dựa vào ghi chép) có thể giúp đánh giá kỹ năng chuẩn bị và kỹ năng đánh giá dữ liệu thực nghiệm chứ không giúp đánh giá khả năng thực hiện thí nghiệm. Như vậy để đánh giá năng lực thực nghiệm nên có những cân nhắc trong việc lựa chọn đánh giá quá trình hay đánh giá kết quả hay kết hợp cả hai để có kết quả chính xác.

2.3. Các nghiên cứu về công cụ đánh giá năng lực thực nghiệm

- Có ba công cụ thường được sử dụng bao gồm: bài kiểm tra viết, bài kiểm tra dựa trên mô phỏng, bài kiểm tra thực hiện thí nghiệm thực. Thông thường trong những đánh giá quy mô lớn người ta hay sử dụng bài kiểm tra viết để đánh giá năng lực thực nghiệm của học sinh. Nhưng ở các nghiên cứu chuyên sâu, các chuyên gia giáo dục đã cho thấy bài kiểm tra viết không thể giúp đánh giá chính xác năng lực thực nghiệm của học sinh. Ví dụ như nghiên cứu của Shavelson [23], nhóm của ông đã thu thập ba đầu điểm của mỗi học sinh bằng ba bài kiểm tra: kiểm tra viết, kiểm tra qua phần mềm mô phỏng, kiểm tra qua tiến hành thí nghiệm thực (chỉ đánh giá thông qua sản phẩm ghi lại trong phòng thí nghiệm chứ không đánh giá quá trình). Sau đó so sánh chúng với điểm chuẩn là điểm thực hành thí nghiệm được đánh giá bởi các quan sát trực tiếp của các chuyên gia. Kết quả cho thấy điểm đánh giá những ghi chép trong phòng thí nghiệm của học sinh đạt tương quan cao ($r = 0.7$) với điểm chuẩn, còn điểm đánh giá qua bài viết và điểm chuẩn đạt tương quan thấp ($r = 0.46$), tương quan giữa điểm đánh giá qua phần mềm mô phỏng và điểm chuẩn chỉ đạt $r = 0.5$. Nhóm nghiên cứu của Nico schreiber [5, 26] cũng sử dụng ba công cụ trên, người học sẽ hoạt động cùng với hai trong ba công cụ đó nhưng bắt buộc một trong số đó sẽ là tiến hành thí nghiệm thực. Giống như Shavelson điểm chuẩn được

chọn cũng là điểm tiến hành thí nghiệm thực. Tuy nhiên không giống Shavelson các tác giả này lại quan tâm đến quá trình tiến hành thí nghiệm. Do đó các thao tác tiến hành thí nghiệm của sinh viên được quay video. Còn với bài kiểm tra qua mô phỏng, những thao tác diễn ra trên màn hình cũng được ghi lại, họ thiết lập một thuật toán và thang điểm tương ứng để mã hóa các hành động của sinh viên. Ban đầu họ kì vọng rằng bài kiểm tra viết sẽ có tương quan thấp với bài kiểm tra tiến hành thí nghiệm thực, ngược lại bài kiểm tra bằng mô phỏng sẽ có tương quan cao với bài kiểm tra tiến hành thí nghiệm thực. Tuy nhiên kết quả phân tích lại cho thấy cả hai bài kiểm tra viết và mô phỏng đều có tương quan thấp với bài kiểm tra thực hành ($r < 0.47$; $r < 0.55$).

- Các nghiên cứu đó cho thấy rằng mặc dù có mối tương quan khác nhau giữa bài kiểm tra viết và mô phỏng với bài tiến hành thí nghiệm thực nhưng cả bài kiểm tra viết và mô phỏng đều không thể thay thế thí nghiệm thực trong việc đánh giá năng lực thực nghiệm. Theo chúng tôi, bài kiểm tra viết có thể giúp đánh giá các kiến thức về lý thuyết sai số, nêu tác dụng của các dụng cụ, nêu ý tưởng thiết kế phương án thí nghiệm nhưng không thể phản ánh được khả năng thao tác với các dụng cụ thực. Tuy nhiên nếu sử dụng bài kiểm tra viết để đánh giá một vài thành tố nào đó của năng lực thực nghiệm thì cũng nên sử dụng các bài tập, câu hỏi dưới dạng tự luận, giải quyết nhiệm vụ, hạn chế sử dụng các câu hỏi trắc nghiệm khách quan. Với bài kiểm tra bằng mô phỏng có thể phù hợp trong việc đánh giá ý tưởng thiết kế phương án thí nghiệm, lựa chọn và bố trí thí nghiệm, đánh giá kiến thức về xử lý sai số, nêu tác dụng của dụng cụ, giải thích kết quả nhưng những đòi hỏi về hiệu chỉnh dụng cụ thí nghiệm, kĩ thuật tiến hành các thao tác/các bước, cách đọc số liệu, cải biến, sửa chữa các dụng cụ thí nghiệm trong giai đoạn thực hiện thí nghiệm thì chỉ thí nghiệm thực mới giúp đánh giá chính xác. Cũng cần phải thấy rằng việc đánh giá năng lực thực nghiệm bằng thí nghiệm thật, đặc biệt là với số lượng học sinh đông thì khá tốn kém (Neumann 2004) [21] và có thể gây tâm lí mệt mỏi cho cả người thi và người đánh giá. Vậy câu hỏi đặt ra là ta sẽ kết hợp sử dụng các công cụ đánh giá trên như thế nào để tiết kiệm được chi phí và đánh giá đúng năng lực thực nghiệm của học sinh Việt Nam?

- Khi giải thích kết quả của mình, nhóm nghiên cứu của Nico schreiber [26] vẫn tin tưởng rằng bài kiểm tra mô phỏng có thể thay thế thí nghiệm thực nếu nó được thiết kế một cách tinh xảo. Họ hoài nghi rằng kết quả tương quan thấp giữa bài mô phỏng và thí nghiệm thực là do: năng lực thực nghiệm của học sinh không phải lúc nào cũng ổn định, nội dung kiến thức, nhiệm vụ đòi hỏi hoạt động thực nghiệm có ảnh hưởng đến mức độ thể hiện năng lực thực nghiệm của học sinh. Sự ảnh hưởng của nội dung vật lí đến kết quả đánh giá năng lực thực nghiệm cũng được các nhà nghiên cứu khác khẳng định như Gott và Duggan (2002) [22]. Điều đó có nghĩa là thiết kế công cụ đánh giá năng lực thực

nghiệm cần phải đảm bảo bám sát nội dung chương trình học tập của HS, cần thiết kế các câu hỏi phù hợp để đánh giá đúng mức độ của từng chỉ số hành vi.

2.4. Các nghiên cứu chỉ ra các khó khăn của người học khi biểu hiện các chỉ số hành vi của năng lực thực nghiệm

Các nghiên cứu ngoài xác định trình độ của học sinh còn đưa ra những khó khăn mà học sinh gặp phải khi biểu hiện các chỉ số hành vi. Mayer và cộng sự đã phát hiện ra rằng hầu hết học sinh (đến lớp 10) không đặt được câu hỏi cho mỗi quan hệ định lượng của các biến [19]. Hofstein và cộng sự [1], nhận thấy rằng những học sinh đặc biệt thiếu kinh nghiệm gặp khó khăn khi đặt ra các câu hỏi khoa học. Một thiếu sót đáng chú ý khác trong việc tạo câu hỏi là học sinh thường không tập trung vào một yếu tố duy nhất có thể dẫn đến các sai sót khi làm thí nghiệm [16]. Về xây dựng các dự đoán/giả thuyết/ hệ quả suy luận từ giả thuyết, de Jong & van Joolingen đã chỉ ra rằng học sinh chỉ đơn giản là “có thể không biết, giả thuyết sẽ trông như thế nào” [4]. Một quan niệm sai lầm phổ biến khác của học sinh là họ thường xem các dự đoán/giả thuyết như những tuyên bố yêu cầu xác nhận hơn là coi các giả thuyết là dự đoán cần được kiểm tra. Học sinh cũng thường không đưa ra được các dự đoán/giả thuyết thay thế [14, 15]. Vì thế học sinh thường giữ lại giả thuyết ban đầu của họ ngay cả khi đối mặt với dữ liệu mâu thuẫn [8].

Các nghiên cứu cũng đã chỉ ra những khó khăn của học sinh trong việc thiết kế và tiến hành các thí nghiệm. Schauble và cộng sự [25] nhận thấy rằng học sinh thường chỉ xem xét một biến và dường như không hiểu mối quan hệ của hai biến được khám phá trong thí nghiệm. Học sinh cũng gặp khó khăn trong việc hiểu hoạt động của các biến [25]. Một vấn đề phổ biến khác là học sinh thường thiết kế hoặc tiến hành các thí nghiệm chỉ với một biến thể của biến độc lập, tức là các nghiên cứu được thiết kế mà không có đối chứng [12]. Một vấn đề nữa liên quan là học sinh thay đổi nhiều hơn một biến độc lập tại một thời điểm. Kết quả là học sinh đề xuất một kế hoạch thí nghiệm thiếu căn cứ [3]. Trong nghiên cứu của Duggan, Johnson, và Gott [7] đã chỉ ra khó khăn của người học liên quan đến việc vận hành biến độc lập như các biến liên tục và kiểm soát các biến nhiễu. Sự thiếu hiểu biết của học sinh cũng liên quan đến việc không nhận thấy tầm quan trọng của các phép đo lặp lại [6, 17].

Về khả năng “phân tích dữ liệu”, Germann và Abram [10] mô tả rằng học sinh thường không xem xét giả thuyết trong khi giải thích dữ liệu, nhưng họ vẫn đưa ra kết luận và vẫn đưa ra lý do cho kết luận của chính họ. Lubben & Millar [17] và Roberts & Gott [24] phát hiện ra rằng học sinh thường không xác định được kết quả bất thường và không đánh giá được kết quả bất thường một cách nghiêm túc trong khi đưa ra kết luận. Roberts & Gott [24] chỉ ra rằng học sinh gặp vấn đề khi xem xét kích thước mẫu, tính đại diện và tính hợp lệ của thiết kế khi đưa ra kết luận.

Các vấn đề và quan niệm sai lầm được nêu ở trên xảy ra ở nhiều học sinh với nhiều độ tuổi. Học sinh càng lớn tuổi thì các vấn đề giảm dần đi. Tuy nhiên, một số vấn đề vẫn còn tồn đọng ở lứa tuổi lớn hơn như: Các nhà nghiên cứu phát hiện ra rằng phần lớn học sinh lớp 12 theo một cách nào đó có thể giải quyết các vấn đề mới trong phòng thí nghiệm. Một số khó khăn của học sinh vẫn liên quan đến việc xây dựng chính xác các giả thuyết và xác định và định nghĩa các biến phụ thuộc và độc lập. Liên quan đến thiết kế thí nghiệm, học sinh bộc lộ những điểm yếu, chẳng hạn như không thể kiểm soát đầy đủ các biến. Các phát hiện bổ sung cho thấy lứa tuổi lớn vẫn khó khăn khi giải thích các kết quả nghiên cứu và thường đưa ra “giải thích trực quan” hoặc “giải thích dựa trên viễn tượng học và / hoặc nhân loại học” trong khi phân tích dữ liệu [28].

3. KẾT LUẬN

Như vậy từ các kết quả của việc tổng quan các công trình nghiên cứu quốc tế chúng tôi thấy rằng còn rất nhiều vấn đề có thể quan tâm nghiên cứu tiếp để đánh giá và phát triển năng lực thực nghiệm cho học sinh Việt Nam. Chẳng hạn học sinh Việt Nam sẽ gặp những khó khăn, sai lầm gì trong quá trình thực hiện các chỉ số hành vi của năng lực thực nghiệm, hay có những biện pháp nào để giúp học sinh Việt Nam khắc phục những lỗi đó, vượt qua những khó khăn đó; công cụ và phương pháp nào có thể dùng để đánh giá năng lực của học sinh Việt Nam. Trong các nghiên cứu tiếp theo chúng tôi sẽ tiếp tục tổng quan các nghiên cứu đánh giá năng lực này ở Việt Nam, tổng quan các nghiên cứu liên quan đến phát triển năng lực này ở trong và ngoài nước.

TÀI LIỆU THAM KHẢO

1. Hofstein et al (2005), *Developing students' ability to ask more and better questions resulting from inquiry-type chemistry laboratories*, J. Res. Sci. Teach. 42.
2. Beschlüsse der Kultusministerkonferenz (2005), *Bildungsstandards im Fach Physik für den Mittleren Schulabschluss: Beschluss vom 16.12.2004*.
3. Z.; Klahr Chen, D. (1999), *All Other Things Being Equal: Acquisition and Transfer of the Control of Variables Strategy*, Child Dev. 70.
4. T.; van Joolingen de Jong, W.R. (1998), *Scientific discovery learning with computer simulations of conceptual domains*, Rev. Educ. Res. 68.
5. Martin; Schreiber Dickmann, Nico; Theyßen, Heike (2012), *Vergleich prozessorientierter Auswertungsverfahren für Experimentaltests*, Jahrestagung in Oldenburg 2011 / Bernholt, Sascha (Hrsg.), tr. 449-451.

6. S.; Gott Duggan, R. Res. Sci. (2000), *Intermediate General National Vocational Qualification (GNVQ) Science: A Missed Opportunity for a Focus on Procedural Understanding?*, Technol. Educ. . 18.
7. S.; Johnson Duggan, P.; Gott, R (1996), *A critical point in investigative work: Defining variables*, J. Res. Sci. Teach. 33.
8. K. Dunbar (1993), *Concept Discovery in a Scientific Domain*, Cogn. Sci. 17.
9. P Labudde E Ramseier, M Adamina (2011), *Validierung des Kompetenzmodells HarmoS Naturwissenschaften: Fazite und Defizite*, Zeitschrift für Didaktik der Naturwissenschaften. 17, tr. 7-33.
10. P.J.; Aram Germann, R.J. (1996), *Student Performances on the Science Processes of Recording Data, Analyzing Data, \Drawing Conclusions, and Providing Evidence*, J. Res. Sci. Teach. 33
11. Christoph Gut-Glanzmann (2012), *Modellierung und Messung experimenteller Kompetenz: Analyse eines large-scale Experimentiertests*, Vol. 134, Logos Verlag Berlin GmbH.
12. M.; Hoi Phan Hammann, T.T.; Ehmer, M.; Grimm, T. (2008), *Assessing pupils' skills in experimentation*, J. Boil.Educ. 42.
13. et al Harriet Hunt và (2018), *An introduction to overviews of reviews: planning a relevant research question and objective for an overview*, Systematic Reviews. 39
14. D.; Dunbar Klahr, K. (1988), *Dual space search during scientific reasoning*, Cogn. Sci. 12.
15. D.; Fay Klahr, A.L.; Dunbar, K. (1993), *Heuristics for Scientific Experimentation: A Developmental Study*, Cogn.Psychol. Sci. 25.
16. D.; Dean Kuhn, D. (2005), *Is developing scientific thinking all about learning to control variables?*, Psychol. Sci. . 16.
17. F.; Millar Lubben, R. (1996), *Children's ideas about the reliability of experimental data*, Int. J. Sci. Educ. 18.
18. et al Lunny C (2017), *Toward a comprehensive evidence map of overview of systematic review methods: paper 1—purpose, eligibility, search and data extraction*, Systematic Reviews. 6.
19. J.; Grube Mayer, C.; Möller, A. (2008), *Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In Lehr-und Lernforschung in der Biologiedidaktik*, Harms, U., Sandmann, A., Eds.; Studienverlag: Innsbruck, Austria, .

20. H Theyßen N Schreiber, H Schecker (2009), *Experimentelle Kompetenz messen?!*, PhyDid A-Physik und Didaktik in Schule und Hochschule. 3(8), tr. 092-101.
21. Knut Neumann, D Schumacher và M Welzel (2004), *Didaktische Rekonstruktion eines physikalischen Praktikums für Physiker*, Logos-Verlag.
22. S Duggan R Gott (2002), *Problems with the assessment of performance in practical science: which way now?*, Cambridge Journal of Education. 32(2), tr. 183-201.
23. Maria Araceli Ruiz-Primo Richard J. Shavelson, Edward W. Wiley (1999), *Note on sources of sampling variability in science performance assessments*, Journal of Educational Measurement. 36(1), tr. 61-71.
24. R.; Gott Roberts, R. (2004,), *A written test for procedural understanding: A way forward for assessment in the UK science curriculum?*, Res. Sci. Technol. Educ. 22.
25. L.; Glaser Schauble, R.; Duschl, R.A.; Schulze, S.; John, J (1995), *Students' Understanding of the Objectives and Procedures of Experimentation in the Science Classroom*, J. Learn. Sci. . 4.
26. Nico Schreiber, Heike Theyßen và Horst Schecker (2012), *Experimental competencies in science: A comparison of assessment tools*, E-Book Proceedings of the ESERA 2011 Conference: Science learning and Citizenship. Part, tr. 66-72.
27. R. Stebler, Reusser, K. & Ramseier, E. (1998), *Praktische Anwendungsaufgaben zur integrierten Förderung formaler und materialer Kompetenzen: Erträge aus dem TIMSS-Experimentiertest*, Bildungsforschung und Bildungspraxis. 20(1), tr. 28-54.
28. P.; Nussinovitz Tamir, R.; Friedler, Y. (1982), *The design and use of a Practical Tests Assessment Inventory*, J. Boil.Educ. 16.

RESULT OF OVERVIEW INTERNATIONAL RESEARCH ON ASSESSMENT OF STUDENT'S EXPERIMENTAL CAPACITY

Le Thi Xuyen

Abstract: *The literature review is a very important step in the research process in general. In order to propose measures to develop and evaluate students' experimental competence, we conduct a review of related documents. This paper focuses on the contents we reviewed related to the research on experimental capacity assessment.*

Keywords: *overview, experimental capacity.*

(Ngày Tòa soạn nhận được bài: 16-9-2022; ngày phản biện đánh giá: 05-10-2022; ngày chấp nhận đăng: 24-10-2022)