



# CHEER

Thông tin

## Nghiên cứu & Đánh giá

Giáo dục Đại học

Trung tâm Nghiên cứu & Đánh giá GDDH, Trường ĐH Nguyễn Tất Thành, TP. HCM, Việt Nam

Số 4-2015

# ĐÁNH GIÁ VÀ ĐỐI SÁNH NĂNG LỰC NGHIÊN CỨU

Giá trị, hay tầm quan trọng của nghiên cứu sẽ gắn với những yếu tố như là tính thiết yếu trong việc giải quyết những vấn đề mà cuộc sống đang đặt ra tại thời điểm đó, và những vấn đề có khả năng sẽ xảy ra trong tương lai.

V. Lynn Meek và Jeannet J. van der Lee

RESEARCH EVALUATION

# Lời nói đầu

**T**rong số trước, chúng tôi đã giới thiệu bản Tuyên ngôn San Francisco như là sự bày tỏ một thái độ của giới học thuật chống lại những định kiến máy móc trong việc dùng chỉ số tác động của tập san như một thước đo chính yếu cho giá trị của nghiên cứu.

Câu hỏi đặt ra là, vậy có những phương pháp nào ta có thể dùng để đánh giá chất lượng của một công trình nghiên cứu, năng lực nghiên cứu của một cá nhân hay đơn vị?

Bản tin Nghiên cứu và Đánh giá GDĐH số 4 xin giới thiệu bài lược dịch một báo cáo do hai tác giả V. Lynn Meek và Jeannet J. van der Lee thực hiện có nhan đề “Về các chỉ số hoạt động nhằm đánh giá và đối sánh năng lực nghiên cứu trong trường đại học”. Bài này đã làm rõ những khái niệm cơ bản như chỉ số hoạt động (performance indicator) hay đối sánh (benchmark), từ đó trình bày các phương pháp định lượng và định tính đang được thực hiện để đánh giá chất lượng nghiên cứu khoa học trong thực tiễn quốc tế, với những điểm mạnh và yếu của nó. Như các tác giả đã kết luận: “Có nhiều vấn đề khó khăn liên quan tới việc sử dụng tất cả các chỉ báo hoạt động. Công nhận điều này không có nghĩa là phản đối việc sử dụng nó mà là lưu ý phải thận trọng trong việc sử dụng nó theo cách nào, ở đâu, vì sao, và diễn giải nó như thế nào”. Bài viết có thể không đem lại cho chúng ta một công thức có thể dễ dàng áp dụng, nhưng làm phong phú thêm hiểu biết của chúng ta, để có thể tiến gần hơn đến một thực tiễn hợp lý nhất trong công tác quản lý khoa học.

Ban biên tập Bản tin xin trân trọng cảm ơn GS. V.Meek đã cho phép sử dụng bản dịch tiếng Việt này cho Bản tin. Vì tài liệu rất dài, chúng tôi chỉ có thể tóm tắt và lược dịch những nội dung quan trọng nhất.

Trân trọng

**BAN BIÊN TẬP**



# Về các chỉ số hoạt động nhằm đánh giá và đối sánh năng lực nghiên cứu trong trường đại học

Tác giả: **V. Lynn Meek và Jeannet J. van der Lee**  
Lược dịch: **Phạm Thị Ly**



## 1. Tổng quan về Đối sánh (benchmarking) và Chỉ số hoạt động (performance indicators)

Garlick và Pryor (2004:28) qua công trình nghiên cứu của họ đã cho thấy rằng “có một mức độ không chắc chắn rất đáng kể trong nhận thức của mọi người về vấn đề đối sánh thực chất là cái gì và liệu nó giúp ích gì cho nhà trường, cho giảng viên và sinh viên. Có một sự lẫn lộn giữa đối sánh, kiểm định chất lượng và nhu cầu về những chỉ số hoạt động định lượng chính yếu nhất.” Bởi vậy, trước hết ta nên xem xét vài định nghĩa và ví dụ nhằm minh họa sự khác biệt giữa những khái niệm này.

### 1.1. “Đối sánh” (*Benchmarking*) có nghĩa là gì?

Jackson and Lund (2000:6) định nghĩa: “Đối sánh là, trước hết và trên hết, một quá trình học hỏi được kiến tạo để tạo điều kiện cho những người tham gia quá trình này có thể so sánh những hoạt động, hay dịch vụ, sản phẩm của họ với người khác hay tổ chức khác, nhằm tìm hiểu chỗ mạnh chỗ yếu trong tương quan so sánh với nhau, để tự cải thiện hoặc điều chỉnh hoạt động của mình.

Garlick và Pryor (2004:19) xây dựng công trình của họ dựa trên khái niệm này để tìm hiểu đặc điểm đối sánh trong bối cảnh đại học những thuật ngữ như hợp tác, lãnh đạo, sự phản ánh, sự bình duyệt, v.v. Điều này nhằm nối kết các bên liên quan trong và ngoài trường theo một cách thức giúp họ thực hiện trao đổi kiến thức về vấn đề tại sao, bằng cách nào, ở đâu và với nội dung gì, sự cải thiện có thể diễn ra.

Việc thực hiện đối sánh sẽ đòi hỏi:

1. một sự cam kết mạnh mẽ, một quyết tâm của lãnh đạo nhà trường (bao gồm cả cam kết về nguồn lực) để đạt đến kết quả và thực hiện những cải thiện cần thiết.
2. một quá trình hợp tác và cam kết của các nhóm liên quan

Về điểm thứ hai, Garlick and Pryor (2004:3) thấy rằng có một khuynh hướng áp dụng các tiêu chuẩn này với thái độ đối phó thay vì là một quá trình tham gia, đối thoại và học tập với mục đích tìm hiểu những gì cần cải thiện. Martin (2003) cũng nói rằng khó mà khép kín cái vòng tròn này và tích hợp những phản hồi có được do đối sánh vào việc cải thiện hoạt động. Chúng tôi cho rằng phản ánh là chức năng quan trọng bậc nhất của quá trình đối sánh, vì nó là những phản hồi chứa đựng nhiều thông tin quý báu cho việc cải thiện hoạt động và tăng cường chất lượng nhà trường.

Đối sánh đã trở thành rường cột chính cho quá trình cải thiện không ngừng (Charles & Benneworth 2002). Về điều này, Garlick and Pryor (2004:9) miêu tả kỹ hơn hai mục tiêu của việc đối sánh:

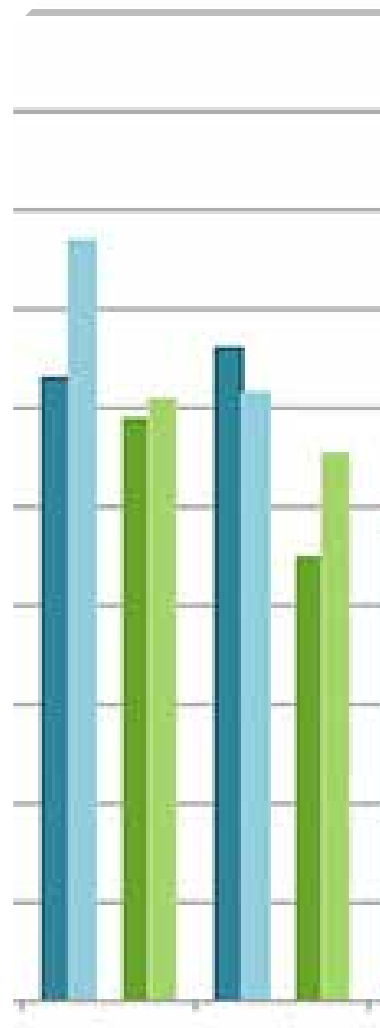
... trước hết, đó là phương tiện đánh giá chất lượng và chi phí hoạt động của quá trình vận hành một tổ chức trong bối cảnh tổ chức đó được xem xét dưới góc độ một doanh nghiệp, hay trong sự so sánh với khuôn mẫu được xem là tốt nhất. Điều này thường được dùng như là một phần trong trách nhiệm giải trình của nhà trường đối với các tổ chức kiểm định, hay với các cơ quan có thẩm quyền quản lý và tài trợ.

Hai là, và có tính chất cơ bản hơn, đối sánh có thể được dùng như một công cụ quản lý thường xuyên nhằm xem xét việc dạy và học, việc hợp tác và lãnh đạo nhằm liên tục cải thiện hoạt động của trường.

Hơn thế nữa, điều rất quan trọng là đối sánh phải là một quá trình của tất cả giảng viên và nhân viên chứ không phải chỉ là việc của một đơn vị phụ trách việc đó trong trường (chẳng hạn như Phòng/Ban Bảo đảm Chất lượng) hoặc là việc của cấp lãnh đạo (Garlick & Pryor 2004:29).

Theo Charles and Benneworth (2002:4) đối sánh có nhiều yếu tố rất có giá trị có thể nêu ra như:

- a. Phạm vi đối sánh bao gồm nhiều mảng hoạt động khác nhau. Bất cứ trường nào cũng có thể tìm ra ít nhất một vài lãnh vực mà mình có thể mạnh và có thể thành công, thay vì bị đánh giá dựa trên một nhóm các tiêu chí chọn lọc từ bên ngoài.
- b. Việc đối sánh có thể hỗ trợ cho quá trình ra quyết định chẳng hạn liệu nhà trường có nên dành nhiều nỗ lực hơn cho việc phục vụ khu vực. Những nỗ lực đó có thể được kết nối với những hoạt động song song khác ở cấp khu vực nhằm xác định đâu là những lĩnh vực ưu tiên.
- c. Cách tiếp cận của việc đối sánh được xây dựng dựa trên các chỉ báo định lượng và định tính, các thước đo quá trình, các chỉ báo cho thấy vị trí dẫn đầu hay tụt hậu. Những thứ đó có thể giúp chúng ta biết rằng liệu những kinh nghiệm hay có đang được áp dụng hay không mà không cần phải chờ đo lường thành quả mới biết.







d. Đối sánh cho phép kết hợp những hình thức đo lường khác nhau và các mô hình hoạt động.

### Các kiểu đối sánh

McKinnon, Walker và Davis (2000:7) đưa ra một tóm tắt rất có ích về những cách tiếp cận chính với việc thiết lập hoạt động đối sánh. Họ phân biệt hai kiểu đối sánh: dựa vào tiêu chí tham chiếu và dựa vào định lượng. Cách tiếp cận thứ nhất xác định đặc điểm thế nào là tốt trong một lĩnh vực cụ thể, nhờ đó tạo điều kiện cho nhà trường đối sánh những thành công của mình trong lĩnh vực đó thông qua so sánh trực tiếp với các tiêu chí. Ngược lại, cách tiếp cận thứ hai phân biệt mức bình thường và mức cạnh tranh trong thành tích đạt được, tạo điều kiện cho việc đánh giá những khác biệt giữa các trường. Việc đặt trọng tâm vào những con số kiểu này bị Garlick và Pryor (2004:19) phê phán mạnh mẽ bởi lẽ “nó nhấn mạnh đánh giá hoạt động thay cho sự cải thiện, và nó phân chia các lãnh vực chức năng thay vì phải nối kết nó lại, bởi vậy đã làm hạn chế quá trình học hỏi để biết rằng liệu có thể cải thiện những chỗ nào và hạn chế những cam kết dài hạn với việc thực hiện những cải thiện ấy”.

McKinnon, Walker and Davis (2000:7-9) cũng nêu ra những vấn đề trọng yếu cần xem xét trong khi thực thi đối sánh trong phạm vi từng trường và giữa trường này với trường khác. Đó là:

- ▶ nhấn mạnh kết quả thay vì quá trình,
- ▶ nhấn mạnh những kinh nghiệm tốt so với tốt nhất, khi “tốt nhất” được định nghĩa là mức độ cao nhất trong đối sánh;
- ▶ thử nghiệm cho việc cải thiện không ngừng;
- ▶ xác định điểm mốc đối sánh nhằm đo lường hiệu quả của các chức năng chứ không chỉ là đo đếm các con số;
- ▶ điều chỉnh những điểm không ngang bằng trong đặc điểm của các trường để kết quả đối sánh có thể diễn đạt dưới hình thức tỉ lệ, thời gian, và tương quan tương đối.
- ▶ việc lựa chọn và sử dụng các mốc đối sánh cần phản ánh sự đa dạng ở những trường có nhiều cơ sở đào tạo khác nhau.
- ▶ cụ thể hóa quy mô đối sánh: cấp trường, cấp khoa hay bộ môn, v.v.;
- ▶ tái xác định thường xuyên những công cụ, phương tiện, thước đo dùng cho việc đối sánh để phản ánh những cải thiện trong định nghĩa về dữ liệu và quá trình thu thập dữ liệu;
- ▶ chọn lựa những mốc đối sánh có hiệu lực và sức thuyết phục;
- ▶ tổ chức các chuẩn đối sánh thành những nhóm chức năng;

- ▶ thực hiện đánh giá qua ý kiến chuyên gia nhằm đạt được sự khách quan; và
- ▶ số lượng các chuẩn đối sánh cần nằm trong phạm vi quản lý được, vì nó cần được giám sát thường kỳ.

### Việc đối sánh đang diễn ra như thế nào trong thực tế?

Ở Úc đối sánh là một thực tế tương đối mới nhưng đến nay vẫn chỉ giới hạn trong việc đánh giá các chức năng quản lý chẳng hạn hoạt động của thư viện và phòng thí nghiệm, quản trị thiết bị, v.v.) hơn là nhấn mạnh vào hoạt động giảng dạy và nghiên cứu. Thái độ chính với đối sánh là coi nó như phương tiện củng cố hoặc nâng cao uy tín tối đa hơn là coi nó như một công cụ quản lý nhằm cải thiện chất lượng hoạt động.

Garlick and Pryor (2004:vii) thấy rằng việc đối sánh đã trở thành phổ biến hơn ở các trường ĐH Úc trong 15 năm qua. Tài liệu hướng dẫn của McKinnon đưa ra năm 2000 (McKinnon, Walker, & Davis 2000) đề xuất 67 chỉ báo hoạt động trong đó có 9 chỉ báo cho hoạt động nghiên cứu. Tuy vậy, nghiên cứu trên đây của Garlick cũng cho thấy việc sử dụng đối sánh chủ yếu là hời hợt và mang tính ngoại vi. Nó chưa được dùng để cải thiện tổ chức trong những lĩnh vực cốt yếu. Sự phát triển của nó không theo kịp đà tiến của những thay đổi gần đây trong kiểm định chất lượng và báo cáo kết quả. Một khó khăn chính yếu khi đối sánh trường này với trường khác là tình trạng thiếu nhất quán trong phương pháp đo lường và các chuẩn đối sánh được sử dụng.

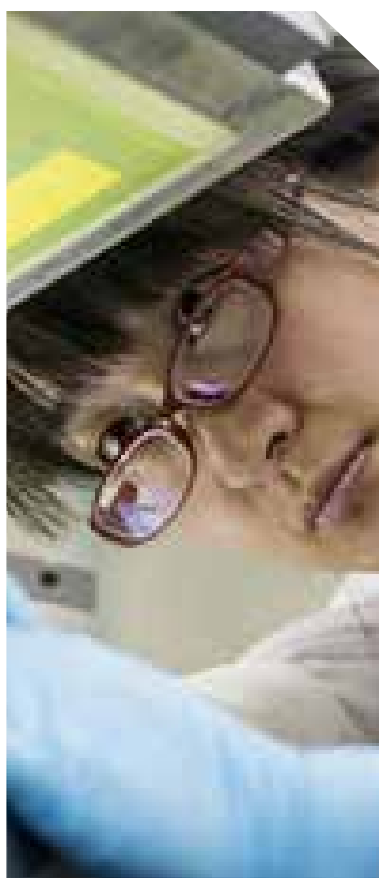
### 1.2. “Chỉ báo hoạt động” (Performance Indicators) nghĩa là gì?

Goedegebuure, Maassen and Westerheijen (1990:29) lưu ý rằng “hiện không có một định nghĩa có thẩm quyền nào về thuật ngữ: “chỉ báo hoạt động”, và tuy nó được hiểu thuần túy là định lượng, vẫn có những chỉ báo định tính đang tồn tại”. Chẳng hạn, Cuenin (1987) công nhận rằng ở mức thấp nhất, chỉ báo hoạt động là một giá trị bằng số đem lại cho ta một thước đo nhằm đánh giá hoạt động của một hệ thống qua định lượng.

Tuy nhiên Dochy và Segers (1990 in Ball & Wilkinson 1994:418) mở rộng sự miêu tả này vào lãnh vực dữ liệu định lượng, bằng cách tuyên bố rằng yêu cầu trước tiên là nó phải có liên quan rõ ràng tới một lĩnh vực chức năng đã được xác định của trường; và hai là nó chỉ diễn đạt những nội dung đúng như cái tên của nó, tức là dấu hiệu cho thấy mức độ đạt được mục tiêu của nhà trường. Yêu cầu thứ ba là nó phải có hiệu lực trong bối cảnh của cái mà nó định diễn đạt và có thể đo lường hay diễn giải một cách đúng đắn.

Kells (1992:133) cũng cho rằng có ít nhất ba loại chỉ báo: (1) những





chỉ báo nhằm phản ánh mức độ đáp ứng của nhà trường đối với chính sách và mục tiêu của nhà nước; (2) chỉ báo về hoạt động dạy và học, nghiên cứu và phục vụ cộng đồng; và (3) chỉ báo cần cho công tác quản lý nội bộ của nhà trường.

### **Vấn đề chỉ báo hoạt động trong thực tế**

Mối quan tâm cao độ đối với vấn đề chỉ báo hoạt động bắt nguồn từ nhiều lý do, ít nhất là từ mong muốn của các trường thực hiện tốt hơn việc hoạch định chiến lược và đòi hỏi của nhà nước về hiệu quả; và mặt khác là tăng cường chất lượng GDĐH. Việc sử dụng chỉ báo hoạt động ngày càng nhiều cần được xem như là một phần khảng khít của một phong trào rộng lớn hơn đối với văn hóa đánh giá (Neave 1988) trong các trường ĐH và trong cả hệ thống, dưới áp lực đòi hỏi của cả hai: trách nhiệm giải trình và bảo đảm chất lượng.

Việc sử dụng các chỉ báo trong đo lường hoạt động của GDĐH đã tiến triển qua nhiều giai đoạn (Doyle 1995). Trước tiên là các chỉ báo được dùng để đánh giá và đối sánh năng lực nghiên cứu của các trường, một cơ chế để phân bổ nguồn lực trong thập niên 60. Sau đó trong những năm 80 nó được dùng như một nguồn thông tin cho các quyết định về tài chính và đánh giá mức độ thành tựu đạt được và xác định những ưu tiên của quốc gia; và rồi đến thập kỷ 90 nó được xem là công cụ để kiểm soát và bảo đảm chất lượng.

Sizer, Spee và Bormans (1992:137) liệt kê năm mục tiêu của việc sử dụng các chỉ báo hoạt động: quản lý, đánh giá, đối thoại, biện luận và phân bổ nguồn lực. Harris (1993:22) thử tiến tới việc định nghĩa chỉ báo hoạt động bằng cách xác định năm nhân tố quan trọng:

1. Nó được diễn đạt dưới hình thức những con số;
2. Nó tìm sự liên đới giữa đầu vào và đầu ra (khi đo lường tính hiệu quả của việc sử dụng nguồn lực);
3. Nó nối kết với các mục tiêu tổng quát của nhà trường (khi liên quan tới hiệu quả của việc đáp ứng các mục tiêu mong muốn);
4. Nó cho phép ta xác định được hoạt động của một người, một đơn vị, tổ chức đã thay đổi như thế nào qua thời gian và so với những người khác, đơn vị khác;
5. Nó có thể được dùng như một thứ kích lệ nhằm ảnh hưởng tới các hoạt động theo cách xã hội mong muốn;

Nhiều yếu tố Harris liệt kê trên đây rất dễ gây tranh luận, nhất là vấn đề chỉ báo hoạt động được coi như thuần túy định lượng, hay quan hệ giữa đầu vào và đầu ra được xem là thước đo của hiệu quả, hoặc cách dùng chỉ báo hoạt động để thực hiện kích lệ gắn với tài trợ.

Bởi thế có ngày càng nhiều tài liệu bàn về việc sử dụng và lạm dụng

các chỉ báo hoạt động trong GDDH (xem Cave, Hanney, & Kogan 1991; Doyle 1995; Goedegebuure, Maassen, & esterheijen 1990; Johnes 1988; Johnes & Taylor 1990; Kells 1990; 1992; 1993; Linke 1992; Politt 1990; Sizer 1992; Stolte-Heiskanen 1992; Yorke 1995). Tuy một vài nhà bình luận ngụ ý rằng việc áp dụng các chỉ báo hoạt động vào GDDH ngày càng trở nên một “khoa học chính xác” với định nghĩa và phương pháp của riêng nó, những người khác cực lực phản đối (Ashworth & Harvey 1994; và Cave, Hanney, & Kogan 1991; hay Politt 1987; và Sizer 1988).

Chỉ báo hoạt động thường được dùng cho một trong ba mục đích: cải thiện, lập kế hoạch, và giải trình trách nhiệm. Ba mục đích này không loại trừ nhau, nhưng được dẫn dắt bởi những động lực khác nhau. Cải thiện hoạt động và lập kế hoạch là vấn đề của từng trường, và có thể cho rằng “chỉ báo hoạt động có giá trị nhất trong bối cảnh cụ thể của từng trường” (Findlay 1990:125). Tuy nhiên, vấn đề trách nhiệm giải trình chủ yếu là do nhà nước đặt ra, do vậy nó đã đặt chỉ báo hoạt động vào lãnh địa chính trị, gây ra tranh luận và phê phán dữ dội (...).

Tổng thuật tài liệu về đề tài này cho thấy có nhiều tranh luận và ý kiến mâu thuẫn nhau về việc sử dụng các chỉ báo hoạt động. Dù vậy ta có thể phát hiện một số điểm được đồng thuận.

Thứ nhất, “ta không được để mình bị cám dỗ bởi việc sử dụng vô số dữ liệu đánh giá. Dùng một số có giới hạn các chỉ báo có hiệu lực cao, là điều rất cần thiết...” (Dochy, Segers, & Wijnen 1990b:136). Thu thập dữ liệu chỉ để có dữ liệu là một việc làm vô nghĩa.

Hai, Jongbloed và Westerheijden (1994:48) lưu ý rằng “những mối nghi ngờ về hiệu lực chính đáng của những gì có thể đo lường được, đặc biệt khi mục tiêu lại là chất lượng, đã dẫn đến tan tành ảo tưởng với chỉ báo hoạt động. Hơn bao giờ hết, rõ ràng các chỉ báo định lượng là những dấu hiệu đòi hỏi sự diễn giải thận trọng trước khi nó được dùng làm cơ sở để ra quyết định”.

Ba, nhiều nhà bình luận tỏ ra nghi ngờ tính đúng đắn của việc áp dụng vào GDDH lối phân tích hiệu quả dựa trên đầu vào đầu ra tương tự như các nhà máy xí nghiệp. Dù vậy, rõ ràng là chính phủ nhiều nước đang áp dụng hoặc dự định áp dụng mô hình này ít ra là một phần cho việc cung cấp ngân sách cho các trường.

Bốn, chỉ báo hoạt động là phương tiện hỗ trợ cho việc phán đoán, chứ không thay thế được cho sự phán đoán và nhận định.

Năm, chỉ báo hoạt động có thể là định tính, mà cũng có thể là định lượng, và cố vạch ra một ranh giới phân biệt tuyệt đối giữa chỉ báo hoạt động và một quá trình chủ quan hơn là đánh giá đồng nghiệp, là một việc ít nhiều có thể xem là vô nghĩa. Thực ra, nhiều chỉ báo diễn đạt bằng con số như tỉ lệ công bố khoa học, rút cục lại là những thứ dựa trên đánh giá đồng nghiệp.







Sáu, một thống kê hay một thước đo trở thành một chỉ báo hoạt động chỉ khi nó được thắm vào một giá trị nào đó và đáng để dùng cho việc đánh giá xem một mục tiêu cụ thể nào đó đã đạt được hay chưa—“việc xác định giá trị và trách nhiệm giải trình trở thành có ý nghĩa chỉ trong những điều kiện đã tuyên bố và những mục tiêu đã được đồng thuận” (Elton 1988:207).

Cuối cùng, rõ ràng là việc sử dụng chỉ báo hoạt động đang được gộp vào một quá trình rộng lớn hơn, đó là đảm bảo chất lượng.

## 2. Đánh giá và so sánh chất lượng trong hoạt động nghiên cứu khoa học

Mục tiêu của bài này là trình bày những cách đánh giá chất lượng của các chương trình nghiên cứu trong một trường ĐH nhằm xác định những chương trình mạnh để ưu tiên kinh phí. Tiêu chí chất lượng bao gồm:

- tác động của kết quả nghiên cứu
- tính bền vững của nghiên cứu
- tầm quan trọng của nghiên cứu
- tiềm năng của nghiên cứu

Phần thảo luận dưới đây sẽ đi theo cấu trúc này.

Tập trung và chọn lọc đã trở thành chuẩn mực toàn cầu trong việc tài trợ nghiên cứu khoa học (NCKH), do chi phí cao ngất của nhiều lĩnh vực NCKH và mong muốn của các trường, của chính phủ các nước đều giống nhau trong việc “chọn kẻ chiến thắng”. Không một nước nào đủ sức tài trợ cho tất cả các trường trở thành đẳng cấp quốc tế, và ngay cả các trường ĐH đẳng cấp quốc tế cũng chẳng đủ sức chi trả cho mọi ngành khoa học của mình. Điều đó có nghĩa là, ta cần nhận ra rằng việc tìm kiếm chỉ những chương trình mạnh mà ưu tiên tài trợ sẽ xói mòn nghiêm trọng sự phát triển của hoạt động NCKH trong phạm vi nhà trường. Bởi vì những lĩnh vực NCKH mới hình thành sẽ không được nhận ra và tài trợ kinh phí thích đáng để khởi sự. Tương tự, những lĩnh vực tụt hậu trong trường có thể cần thêm một liều kinh phí mạnh để tái đầu tư cho những nỗ lực NCKH của họ, hoặc cần một sự hỗ trợ nhất định mà việc đối sánh sẽ giúp tìm ra. Chúng tôi ủng hộ quan điểm cho rằng dùng chỉ báo hoạt động cho việc phân bổ kinh phí là không thích hợp, trừ khi nó gắn với bối cảnh rộng lớn hơn của những mục tiêu và sứ mạng nghiên cứu nói chung của nhà trường và làm giảm nhẹ vai trò của hoàn cảnh trong đó ta xác định và xử lý các hoạt động.

### 2.1 Tác động của nghiên cứu

Trong bài này chúng tôi định nghĩa “tác động của nghiên cứu” là những ảnh hưởng mà nghiên cứu đã, đang, hoặc sẽ tạo ra trong một bối cảnh văn hóa- xã hội, kinh tế và môi trường rộng lớn hơn.

Cho dù những tác động này diễn ra trong ngắn hạn hay dài hạn, nó vẫn là một yếu tố quan trọng phải xem xét, dù rằng hiển nhiên là đơn giản hơn nếu chỉ tính tới những tác động tức thời, bởi lẽ những tác động dài hạn thường không biết được do những khiếm khuyết trong hiểu biết hiện tại của chúng ta. Người ta thường nêu ra trường hợp Rutherford chẳng hạn, ông tin rằng nghiên cứu của mình chẳng bao giờ có ứng dụng cụ thể gì. Hơn nữa, tác động của nghiên cứu có thể được cảm nhận khác nhau trong những khu vực khác nhau. Ví dụ, một kết quả nghiên cứu nào đó có thể có tác động tích cực lên nền kinh tế trong ngắn hạn nhưng có thể để lại hậu quả tiêu cực về môi trường trong dài hạn và thỏa hiệp giữa những kết quả mâu thuẫn như thế với các thước đo tác động quả là không dễ dàng. Sornette và Zajdenweber (1999) nhấn mạnh rằng xác định hiệu quả kinh tế của nghiên cứu trong một thời kỳ cụ thể là một việc rất không đáng tin cậy. Bởi vậy tìm kiếm những chỉ báo cho tác động của nghiên cứu là một khó khăn thách thức đáng kể. Nghiên cứu nhằm xác định quá trình các tác động của hoạt động NCKH nảy sinh cũng là điều rất quan trọng (Allen Consulting Group 2005).

Hiện nay đang có một sự thúc đẩy của chính phủ các nước đòi hỏi NCKH phải mang lại hiệu quả kinh tế tức thời (xem Slaughter & Rhodes 2003), tuy nhiên, cái điều không được quan tâm xem xét là một chính sách như thế tiềm ẩn nhiều hiểm nguy bởi những tác động khả dĩ trong tương lai đã không được tính đến.

Những kiểu tác động trực tiếp hơn và nói chung ít được đo lường, là những tác động đối với cộng đồng hàn lâm trong việc truyền thông và phổ biến kết quả nghiên cứu. Khi tác động được đo lường bên ngoài cộng đồng hàn lâm, thì người ta chú ý nhiều đến tác động kinh tế (bao gồm cả thương mại hóa những kết quả nghiên cứu được tài trợ từ ngân sách công, việc các doanh nghiệp áp dụng kiến thức và xây dựng kỹ năng làm tăng năng suất lao động) (Allen Consulting Group 2005:vii).

“Một khía cạnh của chất lượng trong nghiên cứu là nó đáp ứng nhu cầu của khoảnh khắc hiện tại, hay là “đúng lúc”, “kịp thời”. Những nghiên cứu như thế sẽ có nhiều khả năng được sử dụng và trích dẫn” (Rudd 1988:49). Bởi vậy, một nhân tố giúp xác định tác động của nghiên cứu là nhu cầu về nghiên cứu ấy tại thời điểm đánh giá. Một phương pháp phổ biến là dùng các danh mục liệt kê số lượng trích dẫn, bài báo khoa học và bình duyệt. Ví dụ, các chỉ báo được ĐH Quốc gia Australia sử dụng để xác định tác động của nghiên cứu là “số lượng trích dẫn, bằng sáng chế, cấp phép chuyển giao công nghệ, và những hình thức công nhận chủ yếu khác như mời giảng, mời báo cáo trong các hội thảo quốc gia hay quốc tế” (ANU 2003:3). “ĐH Quốc gia Australia (ANU) tin





rằng điều cốt lõi là nhìn xa hơn những thước đo về thu nhập và bài báo khoa học nếu như ta có thể đạt được những chỉ báo thực sự cho chất lượng của kết quả nghiên cứu”.

Một phương pháp phổ biến khác để đo lường tác động của nghiên cứu là bình duyệt đồng nghiệp. Tuy vậy điều quan trọng nhất là bình duyệt đồng nghiệp phải được bổ sung bằng những chỉ báo hoạt động khác. The Allen Consulting Group (2005:vii) lưu ý rằng để có thể nắm bắt đầy đủ tác động của nghiên cứu (ngoài những tác động kinh tế và khoa học đã nêu trên) của các nghiên cứu được tài trợ bằng tiền ngân sách cần phải đặt câu hỏi “giá trị xã hội của nó là gì?” và “nghiên cứu này liệu có đóng góp gì cho những khía cạnh mà xã hội coi là có giá trị?”. Họ đề xuất bốn nhóm giá trị xã hội:

- ▶ vật chất – bao gồm hàng hóa và dịch vụ sẵn có cho xã hội;
- ▶ con người – bao gồm sức khỏe thể chất, sức khỏe tinh thần, chất lượng của cuộc sống tinh thần, nội tâm, và mức độ con người có thể có những trải nghiệm thú vị;
- ▶ môi trường – bao gồm sự đa dạng sinh học, chất lượng không khí, đất, nước sạch, biển; và
- ▶ xã hội – bao gồm sự gắn kết với xã hội, không bị tổn hại bởi tội phạm, mức độ được bảo đảm các quyền chính trị và mức độ dân chúng tham dự vào các quá trình chính trị.

Về vấn đề này, Hội đồng Khoa học Xã hội và Nhân văn thực hiện một nghiên cứu về cách đo lường giá trị của một kết quả nghiên cứu. Họ lưu ý rằng đo lường chất lượng của nghiên cứu trong lĩnh vực KHXH và nhân văn còn khó hơn, do thiếu những thước đo kết quả NCKH được định nghĩa rõ ràng như trong khoa học tự nhiên. Việc đánh giá nghiên cứu đòi hỏi phân tích ba nhân tố. Trước hết, phải có một phương pháp toàn diện và thích hợp. Giới hàn lâm nói chung đánh giá điều này rất tốt. Hai là, có một giá trị ngầm định của nghiên cứu (không biết được vào lúc bắt đầu). Ba là, kết quả nghiên cứu có tác động (giá trị) khi các khuyến nghị được sử dụng. Đây là một lĩnh vực có xu hướng bị bỏ qua khi người ta chỉ nhấn mạnh tới sản phẩm đầu ra. Tuy vậy, có thể nói rằng, cũng không phải tất cả mọi nghiên cứu đều cần phải có tác động tức thời. Việc phân tích có thể đơn giản là nhằm đạt đến một hiểu biết nào đó mà ta cũng chưa biết là nó sẽ dẫn đến đâu, cũng là một việc rất quan trọng. Ví dụ, với nhóm các khoa về khoa học tự nhiên, mục tiêu (vấn đề cần giải quyết) là làm sao có thể đưa các viện nghiên cứu và các doanh nghiệp lại với nhau. Đếm số lượng báo cáo bởi thế là một điều vô nghĩa. Ta có thể đánh giá giá trị của một nghiên cứu như thế bằng mức độ hữu ích của những ý tưởng và bằng sự đón nhận của giới doanh nghiệp.

## **2.2 Sự bền vững của nghiên cứu**

Sự bền vững của nghiên cứu, trong bối cảnh bài viết này, là nói về khả năng đứng vững được trong trung hạn của một công trình nghiên cứu. Những nhân tố ảnh hưởng tới tính bền vững của nghiên cứu có thể bao gồm những yếu tố dễ thấy như tài chính, hạ tầng, và những dịch vụ hỗ trợ. Tuy nhiên, có những nhân tố khác chưa được xác định sẵn sàng, đó có thể là nhu cầu của các tổ chức tài trợ nghiên cứu, hoặc nhu cầu đào tạo năng lực nghiên cứu trong một lĩnh vực cụ thể (một thước đo khả dĩ là số nghiên cứu sinh tiến sĩ đào tạo được). Sự bền vững của nghiên cứu có vẻ như liên quan tới tác động của nghiên cứu trong đó có năng lực của nhà nghiên cứu tìm được nguồn tài trợ tiếp, điều này gắn với tác động của những nghiên cứu trước đây họ đã thực hiện.

Việc quản lý chương trình nghiên cứu cũng là một yếu tố quyết định tính bền vững của nghiên cứu, bởi lẽ thiếu một sự quản lý, lãnh đạo và tiếp thị phù hợp, một chương trình nghiên cứu lẽ ra có thể phát triển bền vững sẽ có thể trở thành không bền vững do quản lý kém. Mặt khác, để nghiên cứu có được sự bền vững, nó phải gắn với ưu tiên của nhà trường. Ví dụ, một nhà khoa học lỗi lạc có thể duy trì một lãnh vực nghiên cứu chỉ nhờ vị trí của ông hay bà ấy trong cộng đồng học thuật, nhưng nếu người này rời trường, thì việc nghiên cứu này sẽ sụp đổ nếu nó không được xây dựng trong cơ cấu của trường. Hơn thế nữa, vấn đề bền vững phải được làm nhẹ đi với ý thức rằng khoa học vốn rất năng động và thay đổi thường xuyên. Những chương trình nghiên cứu hiện tại cần bền vững, nhưng không đến mức khiến những lĩnh vực nghiên cứu mới có tính chất sáng tạo phải bị làm ngưng.

### **2.3 Tầm quan trọng của nghiên cứu**

Theo Allen Consulting Group (2005: vi): Tuy rằng khi đánh giá chất lượng của những dự án nghiên cứu được tài trợ bằng ngân sách công, người ta phải xem xét vấn đề giá trị của nó; nhưng cũng có thể xem là thích hợp khi người ta phân bổ nguồn lực cho một vài loại nghiên cứu chỉ dựa trên sự ưu tú của nó, dù rằng nó không trực tiếp trả lời câu hỏi cơ bản của các nhà làm chính sách là liệu nghiên cứu này mang lại giá trị gì nếu ta tài trợ cho nó. Để trả lời câu hỏi này cần nhấn mạnh lợi ích xã hội chứ không chỉ là phẩm chất học thuật của nghiên cứu.

Giá trị, hay tầm quan trọng của nghiên cứu sẽ gắn với những yếu tố như là tính thiết yếu trong việc giải quyết những vấn đề mà cuộc sống đang đặt ra tại thời điểm đó, và những vấn đề có khả năng sẽ xảy ra trong tương lai. Ví dụ, hiện nay người ta mong đợi các trường ĐH nghiên cứu những vấn đề có tầm quan trọng đối với quốc gia. Những chỉ báo có thể đoán chừng được để đo tầm quan trọng này là mức độ mà nghiên cứu đáp ứng được nhu cầu ấy. Bởi vậy ngay cả khi một nghiên cứu nào đó không đáp ứng một nhu cầu cụ thể của hiện tại, nó vẫn có thể có một ý nghĩa rất quan trọng trong một tương lai dài hạn hơn.







Tầm quan trọng của nghiên cứu gắn với tính bền vững của nó, tức là nếu nó quan trọng thì nó có nhiều cơ hội để được duy trì lâu dài hơn. Tuy vậy, có một vấn đề là một nghiên cứu có thể bị coi là không quan trọng là do hiểu biết hiện nay của chúng ta bị hạn chế về những vấn đề nảy sinh trong tương lai mà chúng ta không nhìn thấy trước được. Bởi vậy cần có ý thức về những giới hạn trong nhận thức của chúng ta về tầm quan trọng của nghiên cứu và cần cân nhắc để đánh giá xác đáng về mức độ quan trọng của một dự án nghiên cứu.

## **2.4 Tiềm năng của nghiên cứu**

Sornettee và Zajdenweber (1999:1) đặt câu hỏi về việc liệu chính phủ các nước và các công ty nên hỗ trợ hoạt động NCKH ở mức độ như thế nào. Câu trả lời rất đáng được dẫn ra đây.

Câu hỏi đa diện và phức tạp này bao hàm những lợi ích có ý nghĩa chất lượng, chẳng hạn thỏa mãn sự ham hiểu biết tự nhiên của con người, đáp ứng nhu cầu kiến thức và tác động đến văn hóa, giáo dục, nhưng một trong những phần được xem xét kỹ lưỡng nhất của câu hỏi ấy là đánh giá sự giàu mạnh và tăng trưởng kinh tế mà các kết quả nghiên cứu đã làm nảy sinh. Đã có nhiều bản báo cáo nêu ra những chứng cứ cho thấy các lợi ích kinh tế tích cực đến từ nghiên cứu cơ bản. Trong một số lĩnh vực nhất định như công nghệ sinh học, vật lý bán dẫn, truyền thông quang học, tác động của nghiên cứu có thể thấy trực tiếp, trong lúc ở những ngành khác thì con đường từ khám phá tới ứng dụng chứa đầy ngạc nhiên bất ngờ. Hệ quả là, có nhiều thứ không chắc chắn khi ta muốn lượng hóa hiệu quả kinh tế của đầu tư công cho nghiên cứu cơ bản. Quan điểm này có thể giúp ích ít nhiều cho các nhà làm chính sách khi họ cố gắng xác định nên tài trợ ở mức độ nào.

Ở đây, chúng tôi khuyến nghị rằng (..) đánh giá về tiềm năng của nghiên cứu là một “khoa học không chính xác”. Trong việc phán đoán hay nhận định về tiềm năng phải giữ sự quân bình giữa ngắn hạn và dài hạn, giữa những lợi ích đời thường và những biến động lớn.

## **2.5 So sánh quốc tế**

Đánh giá về năng lực nghiên cứu của các trường ĐH do tập đoàn GUNI-AP tiến hành dĩ nhiên sẽ liên quan tới so sánh quốc tế bởi các cơ sở đào tạo của họ nằm ở Trung Quốc, Úc, Ấn độ, Indonesia, Nhật, Hàn Quốc, Malaysia và Thái Lan. Điều quan trọng là cần hiểu rõ chỗ mạnh chỗ yếu của cách tiếp cận này và phương pháp xử lý những vấn đề cụ thể, chẳng hạn sự khác nhau trong cơ cấu tổ chức nhà trường.

Động lực phổ biến nhất để thực hiện so sánh quốc tế là “cải thiện chất lượng/ trọng tâm của nghiên cứu và tăng cường trách nhiệm giải trình đối với việc sử dụng ngân sách công” (DEST 2005a:45). Một trong những vấn đề nan giải với việc đối sánh giữa các trường ĐH là tình trạng thiếu nhất quán trong các mốc đối sánh được dùng, và phương

pháp đo lường, khiến việc so sánh thành ra rất khó khăn. Tài liệu Hướng dẫn Đối sánh của McKinnon, Walker and Davis (2000) đem lại ít nhiều gợi ý, tuy nhiên kiểu tiếp cận “một kích cỡ vừa cho tất cả” này đã bị phê phán là nó không nhận ra sự đa dạng trong vòng đời, vị trí, lối quản trị, quy mô, hay những tham tố khác nữa... giữa các trường.

### Những khó khăn khi so sánh quốc tế

Allen Consulting Group (2004:21) khi so sánh chất lượng nghiên cứu các trường ở New Zealand và Anh đã thấy rằng hạn chế chủ yếu là phạm vi dữ liệu sẵn có về thành tích nghiên cứu của các trường qua các định chế pháp lý quốc tế khác nhau. Nếu tất cả các trường (và các khoa trong đó) ở tất cả các nước đều thu thập và công bố dữ liệu về thành tích nghiên cứu với cùng một tiêu chí về phương pháp, thì thực hiện một nghiên cứu so sánh đáng tin cậy không có gì khó. Khốn nỗi trong thực tế dữ liệu được thu thập và công bố với những cách làm khác nhau rất đáng kể giữa các tổ chức thực hiện. Bởi vậy trước khi tính tới việc dữ liệu phải được dùng như thế nào để so sánh giữa các nước, thì điều rất quan trọng là phải hiểu ta đang có dữ liệu kiểu gì. Có bảy loại dữ liệu về thành quả NCKH:

- ▶ dữ liệu trắc lượng thư mục (bibliometric data);
- ▶ các loại giải thưởng cho cá nhân nhà nghiên cứu;
- ▶ dữ liệu về nghiên cứu sinh;
- ▶ dữ liệu về giảng viên, cán bộ nghiên cứu;
- ▶ thu nhập từ nghiên cứu từ những nguồn ngoài trường;
- ▶ dữ liệu về các hoạt động thương mại hóa kết quả nghiên cứu; và
- ▶ kết quả từ bình duyệt đồng nghiệp.

Cần lưu ý là có những thước đo chất lượng nghiên cứu không được áp dụng phổ biến cho tất cả mọi lĩnh vực NCKH chuyên ngành. Ví dụ, ngay cả công cụ trắc lượng thư mục phổ biến nhất là số lượng trích dẫn cũng có mức độ được chấp nhận khác nhau trong các chuyên ngành khác nhau. Tuy nó được chấp nhận rộng rãi trong khoa học tự nhiên, nó không được chấp nhận như một chỉ báo thích đáng trong cộng đồng hàn lâm ngành khoa học xã hội và nhân văn.

(...)

### Ví dụ về đối sánh quốc tế

Williams và Van Dyke (2004), ở Melbourne Institute of Applied Economic and Social Research, gần đây thực hiện việc so sánh và xếp hạng các trường ĐH Australia so với các trường quốc tế cùng loại. Họ dùng cả những thước đo định tính và định lượng, nhằm mục đích cung





cấp thông tin so sánh cho sinh viên, giảng viên và nhân viên tương lai. Một ví dụ khác là Phụ trang GDĐH của Thời báo Times của UK đã so sánh các trường về vị trí chung. “Bảng xếp hạng vận dụng các thước đo định lượng và định tính với tỉ lệ 50-50. So sánh các trường và tính trọng số cho các đánh giá về uy tín, về khía cạnh giảng dạy và nghiên cứu”. Một ví dụ khác là bảng xếp hạng Giao thông Thượng hải tiến hành xếp hạng quốc tế với các trường chỉ dựa trên thước đo của trắc lượng thư mục để so sánh các trường về thành tích hoạt động học thuật và nghiên cứu” (DEST 2005a:43-44).

(...)

## 2.6 Tiêu chí của chất lượng trong nghiên cứu

“Hiện nay không có cách nào lành mạnh và nhất quán để đo lường chất lượng của các nghiên cứu được thực hiện ở các trường ĐH và các tổ chức nghiên cứu được ngân sách nhà nước tài trợ, cũng như những lợi ích mà những nghiên cứu này đem lại cho cộng đồng hàn lâm và cho xã hội” (Australian Government 2004). Tuy vậy, gần đây chính phủ Australia đã đưa ra một bản báo cáo có nhan đề *Bộ khung đánh giá cho chất lượng nghiên cứu: Một Đánh giá về chất lượng và tác động của nghiên cứu ở Australia*, (xem DEST 2005a), nhằm đưa ra một số hướng dẫn trong vấn đề này. Đây chỉ là một ví dụ về một giải pháp hình thành trong bối cảnh một nước, một tổ chức cụ thể chứ chưa phải là một giải pháp phổ quát. Trong thực tế, tìm xem liệu một giải pháp phổ quát như thế có thể tồn tại hay không là một việc thật vô ích. Trong thế giới ngày càng phức tạp ngày nay, chính bối cảnh sẽ mang lại cho ta một khuôn khổ trong đó giải pháp sẽ thành hình. Điều tốt nhất ta có thể làm là tiêu hóa vô số những công cụ và bộ khung ấy để áp dụng nó trong từng hoàn cảnh cụ thể mà ta thấy nó có thể mang lại thông tin giúp ích cho việc vận hành của mình.

(...)

### Đo lường chất lượng

Rudd (1988:48) trình bày những ý kiến hữu ích về việc này, cụ thể là trong bối cảnh việc đánh giá chất lượng nghiên cứu, tiêu chí về chất lượng, các thước đo khả dĩ có thể dùng, được xem là thích hợp. Ông miêu tả các tiêu chí chủ yếu dùng để đánh giá chất lượng nghiên cứu như sau:

- a. đánh giá chủ quan (bình duyệt đồng nghiệp);
- b. số lượng công bố khoa học;
- c. số lượng trích dẫn;
- d. số lượng và trị giá các hợp đồng nghiên cứu và tài trợ nghiên cứu giành được;

## **Đánh giá chủ quan (bình duyệt đồng nghiệp);**

Bình duyệt đồng nghiệp là phương pháp phổ biến rộng rãi nhất để nhận định về chất lượng của nghiên cứu (Rudd 1988) trong nhiều bối cảnh khác nhau chẳng hạn xin tài trợ, nộp bài báo khoa học hay tuyển người. "Tiêu chí được dùng có thể là khả năng đóng góp cho kho tàng tri thức, mức độ thích hợp của nó trong tổng thể một chương trình nghiên cứu, hay có liên đới với những lĩnh vực ưu tiên, chẳng hạn ở Hà Lan thì đó là "cho thấy giá trị xã hội" của nghiên cứu (Rudd 1988:49).

Điểm thuận lợi của cách tiếp cận này là một người trong cuộc có thông tin sẽ có thể phát hiện ra xu hướng sớm hơn nhiều so với một người ngoài cuộc hay một người không công bằng. Bởi vậy người ta buộc phải kết luận rằng bình duyệt đồng nghiệp là có giá trị, ý kiến của những người thực sự hiểu vấn đề phải được coi trọng (Rudd 1988:49). Có một số khó khăn với phương pháp bình duyệt là việc cộng cơ học điểm số đánh giá của những người bình duyệt có thể đem lại một câu trả lời sai. Dù cho hướng dẫn kỹ tới đâu, người bình duyệt vẫn có thể không tuân theo nhất là khi họ không tán thành với những hướng dẫn ấy khi đánh giá". Vấn đề quan trọng bậc nhất là làm sao tìm được những thành viên hội đồng thực sự khách quan, những người vừa có kiến thức trong lĩnh vực này vừa có khả năng nhận định sắc sảo và công bằng, không thiên vị. Người bình duyệt có thể coi ứng viên là một kẻ địch tiềm năng trong việc cạnh tranh nguồn tài trợ chẳng hạn, là điều khiến nhận định khó có thể khách quan, vô tư.

(...)

## **Số lượng công bố khoa học**

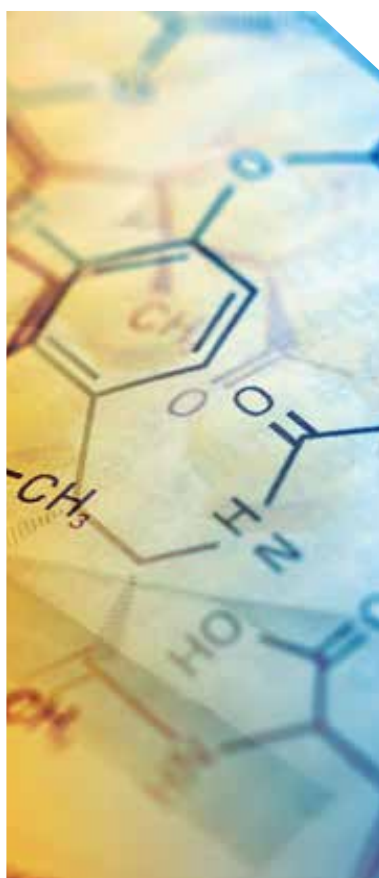
Khi các nước ngày càng thấy cần phải thiết lập chính sách nghiên cứu thích hợp và có hiệu quả, họ quay sang xây dựng những kỹ thuật nhằm quản lý kết quả nghiên cứu, cả trong việc tạo ra tri thức lẫn sử dụng tri thức cho việc phát triển kinh tế (Irvine & Martin 1985).

Năng suất trong nghiên cứu, tuy đo được, nhưng bị giới hạn bởi ảnh hưởng của mức độ cá nhân nhà nghiên cứu thành công trong việc xin tài trợ, và vì khả năng xin được tài trợ lại phụ thuộc thành tích trong quá khứ, thành ra có tình trạng nước chảy chỗ trũng: những người đã đạt được một sự công nhận nhất định sẽ dễ dàng được biết tới, được trích dẫn, và được cấp tài trợ; và được ghi nhận thành tích lớn hơn nữa, trong lúc những người không có tài trợ sẽ khó lòng có điều kiện nghiên cứu để có thể có số lượng công bố đáng kể. Bởi vậy, cái đang được đo khi ta đếm bài báo chỉ phần nào là khả năng kiếm tài trợ, là điều mà bản thân nó chỉ là một phần trong kết quả nhận định về chất lượng nghiên cứu của một cá nhân (Rudd 1988:50-51).

Rudd (1988:51) cũng nói rằng có một sự lầm tưởng rằng hễ bài báo có chất lượng ngang nhau thì cơ hội được tập san chấp nhận cũng







ngang nhau, nhưng có bằng chứng cho thấy không phải lúc nào cũng thế, mà mức độ nổi tiếng của người viết, hay của trường ĐH mà người ấy làm việc, cũng có ảnh hưởng tới khả năng bài báo được chấp nhận.

Bởi vì cách tính năng suất nghiên cứu nhấn mạnh số lượng bài báo, đã có hiện tượng gọi là công bố theo kiểu “salami”, tức một công trình được cắt ra thành nhiều bài báo khoa học hoặc nộp những bài báo có cùng nội dung cốt lõi chỉ khác nhau chút ít không đáng kể cho những tập san khác nhau.

Nhiều kỹ thuật tính số lượng bài báo khoa học dựa trên phân tích thống kê, được biết đến dưới tên gọi trắc lượng thư mục (bibliometrics). Nhiều nhà bình luận cho rằng rất cần không lẫn lộn giữa những yêu cầu kỹ thuật của cách tiếp cận mang tính trắc lượng thư mục với khái niệm tổng quát hơn về chỉ báo hoạt động. (Johnes 1988:56; Rudd 1988). Dù vậy, trắc lượng thư mục đang trở thành cơ chế phổ biến để đánh giá hoạt động NCKH, vì vậy phương pháp này và những hạn chế của nó rất đáng được thảo luận.

Alan Pritchard là người khai sinh ra thuật ngữ “trắc lượng thư mục” vào năm 1969 và định nghĩa nó là “ứng dụng các phương pháp toán và thống kê vào lĩnh vực sách và các phương tiện truyền thông khác” (ASTEC 1989:329). Một định nghĩa gần đây hơn là “một thước đo định lượng về đặc tính của một tư liệu thành văn, thường được dùng như cách để biết năng suất, mức phân phối, việc sử dụng nội dung của tư liệu ấy” (ASTEC 1989:329). Thống kê trắc lượng thư mục trở thành chỉ báo hoạt động khi nó được truyền vào một giá trị.

Tuy các chỉ báo hoạt động tạo ra một loạt dữ liệu thực tế có thể cung cấp thông tin cho các quyết định về chính sách và quản lý, những người làm việc trong lĩnh vực này thường cảnh báo rằng không có một chỉ báo nào, ví dụ như tỉ lệ công bố khoa học, có thể vẽ ra một bức tranh thực sự về hoạt động nghiên cứu của một khoa hay đơn vị cụ thể, và những dữ liệu này phải được tiếp cận với sự thận trọng và được bổ sung bằng những thông tin theo lối định tính (Tognolini, Adams, & Hattie 1994).

Họ cũng lưu ý rằng các chỉ báo hoạt động nghiên cứu bản thân nó không nên được dùng để xác định mức tài trợ. Tuy nhiên “các chỉ báo này rất được giới quan chức ưa dùng vì nó tiện dụng để so sánh và biện minh cho các quyết định về phân bổ kinh phí” (Brennan 1990:110).

Có một số khó khăn khi dùng số lượng trích dẫn để đo tác động của nghiên cứu, như Rudd (1988:52) đã nêu: “về mặt lý thuyết, đếm số trích dẫn để đo năng suất nghiên cứu thì có hiệu quả hơn là đếm số bài báo, ngay cả khi ta phân biệt giữa những tập san có mức độ chọn lọc cao thấp khác nhau.

Tuy nhiên, vấn đề là (a) ta thường lầm tưởng việc trích dẫn một

bài báo hay một cuốn sách cho thấy nó có đóng góp hữu ích cho tri thức chuyên ngành và công trình càng hữu ích thì càng được trích dẫn nhiều; (b) bằng cách trích dẫn những công trình khác, các tác giả đang xử sự như những mong đợi mà người khác đặt ra cho họ; và (c) việc công bố số lượng trích dẫn có nhiều hạn chế. Thực ra “mức độ thường xuyên được trích dẫn còn xa mới có thể coi là một thước đo xác đáng cho tầm quan trọng của nghiên cứu” (Rudd 1988:53). Cuối cùng, việc dựa vào số lượng trích dẫn để đánh giá chất lượng nghiên cứu sẽ thiên vị nặng nề khoa học cơ bản và bất lợi cho khoa học ứng dụng bởi lẽ các nhà nghiên cứu ứng dụng ít công bố bài báo khoa học hơn.

(...)

### **Danh mục các ấn phẩm khoa học**

Danh mục ấn phẩm khoa học được thiết lập nhằm công nhận những nghiên cứu có chi phí thấp, nhất là khoa học xã hội và nhân văn (Bond University 2003:2-3). Danh mục này được thiết lập với một công thức phức tạp kết hợp nhiều tham số: sách, chương sách, tập san có bình duyệt, báo cáo hội thảo có bình duyệt, tác phẩm sáng tạo, bằng sáng chế và các thiết kế có bình duyệt.

### **Số lượng và trị giá các hợp đồng nghiên cứu hoặc tài trợ nghiên cứu giành được**

Thước đo này được coi là tiêu biểu rất tốt cho chất lượng nghiên cứu, bởi có thể ước đoán rằng người ta phải có năng lực nghiên cứu tốt mới có thể giành được tài trợ hoặc hợp đồng. Tognolini, Adams and Hattie (1994:112) cho rằng “giá trị của các khoản tài trợ nghiên cứu phản ánh chất lượng của khoa, vì các tổ chức cấp tài trợ chỉ giao tiền cho những nhà nghiên cứu có phẩm chất cao, những người có hồ sơ thành tích xuất sắc trong việc tạo ra những công trình nghiên cứu quan trọng”.

Cũng như với những thước đo khác “thành công trong quá khứ sẽ quyết định thành công trong hiện tại”(Rudd 1988:54). Rõ ràng là kết quả đo lường của một thước đo như thế sẽ nghiêng lệch về những cá nhân hay nhóm nghiên cứu từng có thành công trước đây trong khi những nhóm nghiên cứu mới hình thành, những lĩnh vực nghiên cứu mới nảy sinh sẽ không được ghi nhận và vì vậy khó lòng mà xin được tài trợ. Hơn thế nữa, trong một số đơn vị hay lãnh vực, nhận được ít tiền tài trợ không nhất thiết có nghĩa là năng lực nghiên cứu của họ là kém cỏi (Tognolini, Adams, & Hattie 1994:113). Bởi vậy, không thể dùng giá trị của các khoản tài trợ nghiên cứu như một thước đo chất lượng trực tiếp như ta tưởng ban đầu. Nó phải được dùng với sự cẩn trọng, và người ta cần cân nhắc tất cả những đảo lộn mà dữ liệu có thể che giấu.

Người ta tưởng rằng số lượng các khoản tài trợ thay cho giá trị của nó có thể là một thước đo “công bằng” hơn. Tuy vậy, dữ liệu cần phải





được chuẩn hóa theo giá trị trung bình của các khoản tài trợ nghiên cứu giữa các chuyên ngành khác nhau (xem Tognolini, Adams, & Hattie 1994:113). Điều này có thể làm nảy nở nhiều khoản tài trợ nhỏ.

### **Chỉ số định tính và định lượng**

Một điểm cần xem xét là các chỉ báo nên là định lượng (như trường hợp Australia hiện nay) hoặc về bản chất là định tính (như trường hợp UK).

Bảng 1 miêu tả một số thuận lợi và bất lợi của mô hình chỉ báo định lượng và mô hình bình duyệt định tính. Charles and Benneworth (2002:4) đã tóm tắt những bất lợi của cả hai loại chỉ báo định lượng và định tính. Trước hết, bất lợi của chỉ báo định lượng là:

- a. Nó phản ánh những hành động và chính sách trong quá khứ chứ không phải những chiến lược hiện tại.
- b. Nó bị ảnh hưởng mạnh bởi cấu trúc của nhà trường và yếu tố đầu vào ví dụ chất lượng của sinh viên.
- c. Nó thường thay thế một cách thô thiển cho những gì thực sự đáng phải đo, và có nguy cơ là giới hàn lâm có thể tìm cách hoàn thành các chỉ báo được yêu cầu thay vì phải đạt được kết quả mong muốn.
- d. Đo kết quả tuyệt đối thì dễ hơn nhiều so với đo giá trị gia tăng.
- e. Có thể cần có khoảng thời gian chờ lâu rất đáng kể để có thành công sau cùng.
- f. Những tác động lớn lao về mặt kinh tế có thể sẽ đòi hỏi phải dám chấp nhận rủi ro, vì vậy có khả năng sẽ có thất bại ngắn hạn và đạt kết quả thấp dù cho cách tiếp cận là đúng.

Thứ hai, những khiếm khuyết của các chỉ báo định tính là:

- a. Những bài học thành công có thể phụ thuộc vào bối cảnh, do đó có thể khó đánh giá những thành công tương đối của một cách tiếp cận nào đó.
- b. Hầu hết những đánh giá định tính đều gắn với sự miêu tả, và khái quát hóa nó là một việc khó.

Bảng 1: Thuận lợi và bất lợi của hai mô hình đánh giá

<b>Mô hình chỉ báo hoạt động</b>	
Thuận lợi	Bất lợi
<ul style="list-style-type: none"> <li>- Khuyến khích năng suất nghiên cứu</li> <li>- Khuyến khích tuyển dụng và giữ chân những người nghiên cứu giỏi</li> <li>- Khuyến khích hướng dẫn nghiên cứu sinh</li> <li>- Dễ thu thập dữ liệu</li> <li>- Có công thức vì vậy minh bạch</li> <li>- Tổn phí ít</li> </ul>	<ul style="list-style-type: none"> <li>- Các chỉ báo chỉ là sự thay thế thô thiển cho sự ưu tú trong nghiên cứu, và nhiều khi nhấn mạnh quá đáng vào số lượng thay vì chất lượng</li> <li>- Có thể tạo ra những mục tiêu khiến xử sự của giới hàn lâm bị lệch hướng</li> <li>- Một số lĩnh vực chuyên ngành hay một số trường (xưa nay nhận được nhiều tài trợ nghiên cứu từ bên ngoài hơn) sẽ được thiên vị với cái giá phải trả là sự thiệt thòi của ngành khác, trường khác.</li> </ul>
<b>Mô hình bình duyệt</b>	
Thuận lợi	Bất lợi, khó khăn
<ul style="list-style-type: none"> <li>- Trực tiếp hơn và đo lường chất lượng xác đáng hơn</li> <li>- Khuyến khích việc xác định các trung tâm xuất sắc</li> <li>- Là một phương pháp đánh giá cho phép cân nhắc những yếu tố khác ngoài bình duyệt</li> <li>- Một phương tiện hiệu quả để tập trung nguồn lực vào những lĩnh vực có thể mạnh</li> <li>- Có thể được dùng cho đối sánh quốc tế</li> <li>- Có thể thay đổi về quy mô, có tính đa chức năng và linh hoạt</li> </ul>	<ul style="list-style-type: none"> <li>- Tổn phí hơn trong việc quản lý và tuân thủ quy tắc</li> <li>- Không có mấy trọng lượng theo quan điểm của người sử dụng kết quả NCKH sau cùng</li> <li>- Kích thích những hiệu ứng mâu thuẫn với những chuẩn mực được chấp nhận trong chính sách tuyển giảng viên (ví dụ cạnh tranh giành giảng viên ngôi sao)</li> <li>- Có thể có định kiến (chẳng hạn kỳ thị các nhà nghiên cứu nữ)</li> <li>- Có thể gia cố thêm chủ nghĩa truyền thống trong học thuật khiến nó khó thay đổi và bất lợi cho những nghiên cứu không thuộc dòng chính thống</li> <li>- Khó mà tìm đủ người bình duyệt</li> </ul>



**Nguồn:** OECD (2004:13-14 in DEST 2005a:50-51)





Tác giả kết luận rằng:

Cả hai lối đánh giá định lượng và định tính đều có vấn đề về quy mô. Các thước đo này nên được thực hiện ở quy mô nào? Thước đo nào có thể áp dụng cho mọi khoa, mọi hoạt động? Nhiều hoạt động được thực hiện chỉ ở một số đơn vị trong trường, vì vậy có xu hướng ít được đo lường và báo cáo tới lãnh đạo. Có những người đã đóng góp nhiều nỗ lực đáng kể nhưng không được ghi nhận trong đơn vị của họ (Charles & Benneworth 2002:5).

Một cách logic ta cũng có thể thấy trước rằng thứ hạng của một trường có thể khác nhau tùy vào những tiêu chí được dùng để đánh giá: số lượng bài báo khoa học, trích dẫn, số lượng và giá trị của tài trợ nghiên cứu, các chỉ số định tính và định lượng khác, v.v. Tuy vậy Irvine và Martin (trong Rudd 1988) cho thấy rằng bất kể người ta dùng tiêu chí gì thì kết quả cũng khá nhất quán. Có lẽ phần nào là do có một xu hướng là chỉ những biến tố nào có thể đo được và do đó có thể đưa vào máy tính để xử lý là được tính tới (Rudd 1988:54).

Rõ ràng là có nhiều vấn đề liên đới với việc sử dụng tất cả các chỉ báo hoạt động. Công nhận điều này không có nghĩa là phản đối việc sử dụng nó mà là lưu ý phải thận trọng trong việc sử dụng nó theo lối nào, ở đâu, vì sao, và diễn giải nó như thế nào.

**Người dịch:** Phạm Thị Ly.

**Nguồn:** Meek, V. L., & van der Lee, J. J. (2005). Performance indicators for assessing and benchmarking research capacities in universities. In APEID, UNESCO Bangkok Occasional Paper Series Paper no. 2: United Nations Educational, Scientific and Cultural Organisation (UNESCO).

Chịu trách nhiệm xuất bản: TS. NGUYỄN MẠNH HÙNG

*Hiệu trưởng Trường ĐH Nguyễn Tất Thành*

Cố vấn khoa học: TS. Nguyễn Đức Hưng – GS. Nguyễn Lộc

Tổ chức bản thảo và biên tập: TS. Phạm Thị Ly

*Giám đốc Trung tâm Nghiên cứu & Đánh giá GDĐH*

Biên tập bản tiếng Anh: TS. Allen Heyd

Trình bày: Phạm Thanh Tâm

Mọi chi tiết xin liên hệ: Trung tâm Nghiên cứu và Đánh giá GDĐH Trường ĐH Nguyễn Tất Thành

Số 298A, Đường Nguyễn Tất Thành, Quận 4, TP. Hồ Chí Minh.

ĐT: (08) 83940 2810 - Fax: (08) 3940 4759 - Email: cheer@ntt.edu.vn