Appendix B. Supplementary Materials for the paper: 
Is Science Becoming more Interdisciplinary? 
Measuring and Mapping Six Research Fields over Time

These materials support the analyses of the paper. They are not included there so as to minimize 
distraction from the main points of the paper and to save space.

SCs Similarity matrix

Both the Integration metric and the overlay science maps draw upon a co-citation matrix of 244 Web of 
Science Subject Categories (“SCs”). As described in the paper, these are derived from a sample of 30261 
articles indexed by Web of Science. They reflect six separate one-week samples of US-authored articles 
chosen in 2005-07. VantagePoint software was used to extract the source (journal) from the cited 
references, and then to match journals to SCs. A 244 SC x 244 SC co-occurrence matrix was then 
computed, reflecting the number of times each pair of SCs was cited by a given article. From that, we 
have calculated the cosine values among the SCs. This cosine matrix is provided in the auxiliary MS Excel file (“21 Macro-disciplines – 244 SCs”) on the worksheet called 244x244 cos matrix jun12 (2008).

From that cosine matrix, we then used Principal Components Analysis (PCA) to extract factors (more 
properly, principal components). We settled on a 21-factor solution, as described in the paper. That matrix 
of SCs by Factors appears as worksheet: 21-solution-IR. We name each factor by perusing the SCs which 
load on it, taking note of those most highly related. These are the “macro-disciplines” that appear as 
Appendix A. They are also in the Excel file on the worksheet: Macro-disciplines SCs list.

This document contains the following:

- Tables B1-B6 that show the “Top 10” Cited SCs for each of the six benchmark Subject Category 
paper sets in 2005 and in 1975.

Citation Characteristics of the Six Subject Categories

Some detail may enrich our understanding of the citation behavior. The Neurosciences, 2005, sample of 
1768 articles cite 121,176 distinct SC instances. That averages to almost 69 cited SCs per article.¹ Those 
1768 articles cite 205 of the 244 SCs at least once. Researchers reach out widely! Table B1 provides 
citation detail for Neurosciences.

Table B1. Top 10 Subject Categories Cited by Neurosciences Articles: 2005 and 1975

<table>
<thead>
<tr>
<th>Cited Subject Categories</th>
<th>2005 % of Cited Instances</th>
<th>2005 % of Cited Instances</th>
<th>1975 % of Cited Instances</th>
<th>1975 % of Cited Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEUROSCIENCES</td>
<td>30.3%</td>
<td>24.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLINICAL NEUROLOGY</td>
<td>5.9%</td>
<td>5.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOCHEMISTRY &amp; MOLECULAR BIOLOGY</td>
<td>5.7%</td>
<td>6.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MULTIDISCIPLINARY SCIENCES²</td>
<td>5.2%</td>
<td>5.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHARMACOLOGY &amp; PHARMACY</td>
<td>4.6%</td>
<td>7.1%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

¹ Those articles average 47 references each (see Table 2). The higher number of cited SCs reflects that many 
journals are associated with more than one SC.
² The Multidisciplinary Sciences SC includes several highly prominent journals – Science, Nature, and PNAS, 
among others.
Table B1 lists the “Top 10” most cited Subject Categories, of the 205 SCs cited by the sample of Neurosciences articles for 2005. The percentages are based on citation instances. That is, if one article cites 10 different “Clinical Neurology” articles, that counts as 10 instances, for this one record. The “Records Rank” reflects how many of the 1768 articles cited that SC at least once. Note that articles cite Neurosciences about 5X as often as any other SC. This suggests a core of shared knowledge. On the other hand, the articles cite a lot of other material as well – in 2005, about 70% of their references are to other SCs. As per a previous footnote, some journals are associated with more than one SC – we don’t differentiate whether certain references are to journals that link to Neurosciences AND to another SC.

Table B2 provides similar information for Biotech. Some interesting features:

- These articles cite another SC (Biochem) more often than they cite their own SC, Biotech
- That said, Biotech is cited by a higher percentage of the articles (Records Rank is #1; 895 of 990 records) than is Biochem (817 of 990 records citing this SC), but just not as many citation instances.
- Note a dramatic shift over 3 decades – Biotech articles draw much more upon Genetics research than their 1975 predecessor articles did.
- Likewise, Cell Biology is cited considerably more frequently than previously.
- Conversely, Biotech in 2005 is proportionately less inclined to reference Microbiology or Immunology than previously.

Table B3 does likewise for EE. Unlike the others, EE and Neurosciences show a small increase in concentration of citation in their own SC from 1975 to 2005. In other words, they have become a bit less integrative in this regard. Here too, there are notable shifts in the degree to which EE research draws upon other SCs – note in particular the marked jumps for “AI” and Optics from 1975 to 2005.

<table>
<thead>
<tr>
<th>Cited Subject Categories</th>
<th>2005 Records Rank</th>
<th>2005 % of Cited Instances</th>
<th>1975 Records Rank</th>
<th>1975 % of Cited Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSYCHIATRY</td>
<td>9</td>
<td>4.4%</td>
<td>9</td>
<td>3.6%</td>
</tr>
<tr>
<td>PHYSIOLOGY</td>
<td>5</td>
<td>3.8%</td>
<td>3</td>
<td>8.7%</td>
</tr>
<tr>
<td>BEHAVIORAL SCIENCES</td>
<td>8</td>
<td>3.4%</td>
<td>11</td>
<td>3.1%</td>
</tr>
<tr>
<td>CELL BIOLOGY</td>
<td>7</td>
<td>2.8%</td>
<td>15</td>
<td>2.4%</td>
</tr>
<tr>
<td>PSYCHOLOGY</td>
<td>10</td>
<td>2.6%</td>
<td>12</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cited Subject Categories</th>
<th>2005 Records Rank</th>
<th>2005 % of Cited Instances</th>
<th>1975 Records Rank</th>
<th>1975 % of Cited Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOCHEMISTRY &amp; MOLECULAR BIOLOGY</td>
<td>2</td>
<td>13.1%</td>
<td>2</td>
<td>15.2%</td>
</tr>
<tr>
<td>BIOTECHNOLOGY &amp; APPLIED MICROBIOLOGY</td>
<td>1</td>
<td>11.0%</td>
<td>1</td>
<td>14.0%</td>
</tr>
<tr>
<td>GENETICS &amp; HEREDITY</td>
<td>4</td>
<td>7.9%</td>
<td>35</td>
<td>0.4%</td>
</tr>
<tr>
<td>MICROBIOLOGY</td>
<td>7</td>
<td>6.6%</td>
<td>3</td>
<td>14.8%</td>
</tr>
<tr>
<td>MULTIDISCIPLINARY SCIENCES</td>
<td>3</td>
<td>6.2%</td>
<td>4</td>
<td>5.8%</td>
</tr>
<tr>
<td>CELL BIOLOGY</td>
<td>5</td>
<td>4.8%</td>
<td>17</td>
<td>1.1%</td>
</tr>
<tr>
<td>MEDICINE, RESEARCH &amp; EXPERIMENTAL</td>
<td>10</td>
<td>3.0%</td>
<td>8</td>
<td>2.3%</td>
</tr>
<tr>
<td>ONCOLOGY</td>
<td>11</td>
<td>2.7%</td>
<td>15</td>
<td>1.9%</td>
</tr>
<tr>
<td>BIOCHEMICAL RESEARCH METHODS</td>
<td>6</td>
<td>2.6%</td>
<td>9</td>
<td>1.4%</td>
</tr>
<tr>
<td>IMMUNOLOGY</td>
<td>9</td>
<td>2.5%</td>
<td>5</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

Table B3. Citing by Engineering, Electrical & Electronic Articles: 2005 and 1975
Table B4 presents the corresponding breakout for Math. The striking concentration of citation within Math is apparent. While occurrences are only around the 1% level of total citations, the range of other “multidisciplinary” SCs being cited by Math articles is intriguing (e.g., Chemistry, Multidisciplinary).

Table B4. Citing by Mathematics Articles: 2005 and 1975

<table>
<thead>
<tr>
<th>Cited Subject Categories</th>
<th>2005 Records Rank</th>
<th>2005 % of Cited Instances</th>
<th>1975 Records Rank</th>
<th>1975 % of Cited Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATHEMATICS</td>
<td>1</td>
<td>69.2%</td>
<td>1</td>
<td>79.1%</td>
</tr>
<tr>
<td>MATHEMATICS, APPLIED</td>
<td>2</td>
<td>12.0%</td>
<td>2</td>
<td>10.3%</td>
</tr>
<tr>
<td>PHYSICS, MATHEMATICAL</td>
<td>3</td>
<td>3.2%</td>
<td>10</td>
<td>2.6%</td>
</tr>
<tr>
<td>COMPUTER SCIENCE, THEORY &amp; METHODS</td>
<td>8</td>
<td>1.7%</td>
<td>12</td>
<td>0.8%</td>
</tr>
<tr>
<td>MULTIDISCIPLINARY SCIENCES</td>
<td>4</td>
<td>1.2%</td>
<td>3</td>
<td>1.0%</td>
</tr>
<tr>
<td>CHEMISTRY, MULTIDISCIINARY</td>
<td>5</td>
<td>1.1%</td>
<td>4</td>
<td>0.5%</td>
</tr>
<tr>
<td>MATERIALS SCIENCE, MULTIDISCIINARY</td>
<td>6</td>
<td>1.1%</td>
<td>6</td>
<td>0.5%</td>
</tr>
<tr>
<td>CHEMISTRY, PHYSICAL</td>
<td>7</td>
<td>1.0%</td>
<td>7</td>
<td>0.7%</td>
</tr>
<tr>
<td>PHYSICS, MULTIDISCIINARY</td>
<td>10</td>
<td>0.9%</td>
<td>13</td>
<td>0.6%</td>
</tr>
<tr>
<td>MATHEMATICS, INTERDISCIINARY APPLICATIONS</td>
<td>13</td>
<td>0.8%</td>
<td>9</td>
<td>0.7%</td>
</tr>
</tbody>
</table>

Table B5 shows the pattern for Medicine-R&E. Note another sharp rise for Genetics. Oncology and Cell Biology have become more salient as well. While Medicine-R&E articles are most apt to cite at least one Medicine-R&E article (#1 Records Rank), they cite a higher total number of Biochem and Immunology items in 2005. The citation instances distribution in 2005 is less concentrated in the top 3 cited SCs than it was in 1975 – another facet of increasing interdisciplinarity.

Table B5. Citing by Medicine – Research & Experimental Articles: 2005 and 1975

<table>
<thead>
<tr>
<th>Cited Subject Categories</th>
<th>2005 Records Rank</th>
<th>2005 % of Cited Instances</th>
<th>1975 Records Rank</th>
<th>1975 % of Cited Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOCHEMISTRY &amp; MOLECULAR BIOLOGY</td>
<td>2</td>
<td>10.2%</td>
<td>2</td>
<td>15.1%</td>
</tr>
<tr>
<td>IMMUNOLOGY</td>
<td>7</td>
<td>9.6%</td>
<td>9</td>
<td>10.0%</td>
</tr>
</tbody>
</table>
Table B6 presents the comparisons for Physics-AMC. This feels like a “regular” profile of intuitively related research areas centered on the SC in question. Physics-AMC’s citation pattern is less concentrated on itself in 2005 than it was in 1975. Association with Physical Chemistry has notably risen.

Table B6. Citing by Physics – Atomic, Molecular & Chemical: 2005 and 1975

<table>
<thead>
<tr>
<th>Cited Subject Categories</th>
<th>2005 Records Rank</th>
<th>2005 % of Cited Instances</th>
<th>1975 Records Rank</th>
<th>1975 % of Cited Instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHYSICS, ATOMIC, MOLECULAR &amp; CHEMICAL</td>
<td>1</td>
<td>28.0%</td>
<td>1</td>
<td>44.2%</td>
</tr>
<tr>
<td>CHEMISTRY, PHYSICAL</td>
<td>3</td>
<td>11.7%</td>
<td>3</td>
<td>5.5%</td>
</tr>
<tr>
<td>PHYSICS, MULTIDISCIPLINARY</td>
<td>2</td>
<td>9.8%</td>
<td>2</td>
<td>12.9%</td>
</tr>
<tr>
<td>OPTICS</td>
<td>6</td>
<td>8.1%</td>
<td>4</td>
<td>8.0%</td>
</tr>
<tr>
<td>CHEMISTRY, MULTIDISCIPLINARY</td>
<td>5</td>
<td>6.5%</td>
<td>5</td>
<td>5.8%</td>
</tr>
<tr>
<td>PHYSICS, CONDENSED MATTER</td>
<td>7</td>
<td>3.8%</td>
<td>10</td>
<td>1.8%</td>
</tr>
<tr>
<td>MULTIDISCIPLINARY SCIENCES</td>
<td>4</td>
<td>3.8%</td>
<td>6</td>
<td>2.4%</td>
</tr>
<tr>
<td>SPECTROSCOPY</td>
<td>9</td>
<td>2.6%</td>
<td>7</td>
<td>3.5%</td>
</tr>
<tr>
<td>PHYSICS, APPLIED</td>
<td>8</td>
<td>2.1%</td>
<td>8</td>
<td>2.1%</td>
</tr>
<tr>
<td>PHYSICS, MATHEMATICAL</td>
<td>10</td>
<td>1.8%</td>
<td>14</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

Correlation between Average Integration Score and Number of Journal per SC

In a prior study [PORTER et al., 2008], we had profiled 22 SCs for 2005 only, analyzing smaller samples of about 100 articles each. The 22 SCs showed a limited range of mean Integration scores, ranging only from 0.52 (EE) to 0.69 (Engineering, Biomedical). This pointed us toward the present sampling in more depth and over time. Combining the two samples, these 23 SCs (5 of the current sample of 6 SCs were included in the prior 22) are composed of anywhere from 14 (Materials Science, Biomaterials) to 262 journals (Biochemistry & Molecular Biology). Figure B1 consolidates Integration score results, using the most updated values. This shows average Integration score vs. # of journals for 23 SCs. The outlier on Integration is Math. The slope (regression between Integration and SC size for these results is slightly negative (-.064). [The correlation is -0.54; it changes only marginally to -0.51 if Math is removed]. Integration score does take into account similarity among cited SCs to reduce the influence of this SC size effect on I score. Nonetheless, this relationship between SC size and average Integration score is a consideration in cross-SC comparisons.
Figure B1. Average Integration Score (2005) vs. Number of Journals for 23 Subject Categories
Integration Profiles

Figure 5 finds that Integration score central tendencies for these large sets of articles tend to fall in a relatively narrow range. Excepting Math, the means and medians for the other 20 article sets (5 SCs at each of 4 years) range from 0.50 (EE, 1995) to ~0.67 (Medicine-R&E, 2005). Integration score for a given article can range from 0 (i.e., all references associated with the same SC) to an upper limit approaching 1 (i.e., every reference from distinct, unrelated SCs). Figure B2 presents the distributions of Integration scores for the Math articles for each of the 4 years. Note that many of these are zero – those articles cite only from a single SC (almost surely Math). The sampled articles were screened to limit inclusion to those with at least 4 cited references, that associated with at least 3 cited SC instances (i.e., each SC citation could be to the same SC). We did this to avoid characterizing how integrative an article is based upon tiny numbers (e.g., every article that cites only a single SC will have Integration = 0).

As per Table 2, the mean Math Integration score increases a lot from 0.211 in 1975 to 0.275 in 1985, and then only gradually to 0.283 in 1995 and 0.293 in 2005. Distributional profiles for the other 5 SCs are quite similar. Figure B3 illustrates the case of Biotech. [To shorten the article, the other 4 are included in a “For review only Appendix,” and will be made available on our websites.] This demonstrates that the increase in Integration is not restricted to changes in a limited range of articles (e.g., those scoring especially high), but is a general phenomenon. In other words, these are reasonably normal distributions where the mean and the median take similar values, as shown in Table 2.

Figure B2. Distribution of Integration scores for Mathematics articles: 1975 – 2005

Math
Figure B2. Distribution of Integration scores for Biotechnology & Applied Microbiology articles: 1975 - 2005

Biotech

Figures B3-B6 show the distributions for the other SCs.
Figure B3. Distribution of Integration scores for “Engineering, Electrical & Electronic” articles: 1975 - 2005

Figure B4. Distribution of Integration scores for “Medicine – Research & Experimental” articles: 1975 - 2005
Figure B5. Distribution of Integration scores for Neurosciences articles: 1975 - 2005

Figure B6. Distribution of Integration scores for “Physics – Atomic, Molecular & Chemical” articles: 1975 - 2005
Science Maps

Our science maps help one perceive how closely related the larger macro-disciplines are to each other. Interestingly, the labels such as physics, engineering or chemistry, as used in ISI SCs, do not generally all coalesce within a single macro-discipline. For example, the 8 SCs that are Physics sub-fields disperse into three macro-disciplines, and the 17 Engineering sub-fields disperse into 7 macro-disciplines. This runs counter to our casual intuition. It warns that the evolving research enterprise does not neatly conform to the traditional academic “disciplines” or organizational units. And, it cautions against assessing the degree of interdisciplinarity based on, say, academic department affiliations.

Sizing of nodes in these science overlay maps can be controlled in Pajek. One can apply different transforms (e.g., linear used herein or log), depending on whether one wants to emphasize larger or not-so-large nodal activity. We have opted to present these and the following overlays of various SCs to show the distribution of SCs being substantially cited by these sets of papers. We have not scaled so that proportions of citations to various nodes are normalized. One normalization option, for instance, would be to rescale the Biotech-2005 nodes so that the total citation activity were the same as for Biotech-1975. While advantageous in precise comparison of shifts in emphasis over particular SCs, this did not meet our main objective here to allow an overall visual perception of the diversity and emphases of the citations. [Tables B1-B6 indicate the proportional citation to the top SCs for each of the 6 focal Subject Categories.] Our science overlay mapping kit will allow users their choice to map as best fulfills their aims.

We see this science overlay mapping as a powerful resource to facilitate further analyses. To illustrate -- as will be the case in our other five target SCs, there is not much cross-disciplinary referencing of the social sciences and humanities by Biotech. For instance, these 990 Biotech articles cite 42 Policy Sciences and 34 Ethical & Social Issues references. One wishing to probe further could scour the infrequently cited macro-discipline citations. For instance, here are titles of a few of the Biotech articles that cite “ethical & social issues” sources:

- A new era in the ethics of human embryonic stem cell research
- Life cycle assessment of various cropping systems utilized for producing biofuels: Bioethanol and biodiesel
- Risk and ethics in biological control

One could use the maps to trigger analyses to locate particular cross-disciplinary intersections. For instance, if our interests keyed on Biotech in the Ecological Sciences arena and attendant Policy issues, we could zoom in on select Biotech articles that cite work in both of these other macro-disciplines. Here, that would result in seven articles – one example:

- Economic framework for decision making in biological control

This notion could also be adapted to the level of specific SCs. For instance, 9 articles cite “Geriatrics & Gerontology” together with “Oncology” in the Biotech-2005 set, including:

- Integration of environment and disease into 'omics' analysis

Figure B7 shows the Medicine-Research & Experimental citation pattern for 2005; 1975 appears relatively similar so is not shown here (but is available on the website). This extremely broad ranging citation pattern presents a remarkable contrast to Math’s highly concentrated citations (Figure 10). Certainly not a shock, but it is compelling to see how widely Medicine-R&E research reaches for its knowledge base. Reexamining the tally of cited SCs finds that our set of 775 Medicine-R&E articles for 2005 cited 56 other SCs at least 100 times. [In contrast, the 684 Math-2005 articles cited only 7 other SCs that often.]
Physics—Atomic, Molecular & Chemical maps for 1975 and 2005 are generally similar in the diversity of SCs that are highly cited. The increase in Integration scores between these sets are due to the reduced concentration in the most highly cited SC (see Table B6). We show Physics-AMC for 2005 in Figure B8 (Physics-1975 will be posted on our website). It is interesting to note that this SC specialty is located by our statistical procedures in the Chemistry macro-discipline. The Figure nicely depicts that this research domain operates at the interface of Chemistry, Physics, and Materials Science. Table B6 reaffirms this. The Subject Category in the Biomed Sciences upon which it draws heavily is “Multidisciplinary Sciences.”

**Future developments in measurement and science mapping**

It is vital to get beyond IDR rhetoric to see which research areas do exchange knowledge to what degree. Measuring and mapping interdisciplinarity of more coherent bodies of research (c.f., RAFOLS and MEYER, forthcoming) will provide insights regarding how research communities intersect. That, in turn, can help elucidate means to enhance knowledge transfer – i.e., to convene a workshop that targets particular research communities who appear to have valuable knowledge worth sharing, but are not doing so to the extent that seems worthwhile. For the graduate student, we suggest that these tools can help identify valuable skill sets (e.g., to learn about a new tool or method from another discipline). Or, help orient a dissertation by mapping trajectories of related research to identify unexplored extensions. For instance, if no one has reported applying a given approach to a lingering issue, perhaps this is offers an opportunity for a significant scientific contribution.

We use the science maps to locate particular bodies of research among the macro-disciplines. That helps identify key cross-“disciplinary” relationships that might benefit from particular nurturing. The maps and supporting metrics also help perceive changes in degree of interrelationship over time. That could be useful in evaluating the effectiveness of particular R&D initiatives over the long term. Tables B1-B6 tally the degree to which the six SCs under scrutiny reference research published in other SCs. Increases in citation of certain research domains (e.g., Biochemistry, Genetics) suggest that students being trained in certain other fields might benefit from exposure to key knowledge and skills in these foundational fields. Figure B7 helps gauge the extent to which one research domain (as reflected by the “Medicine—Research & Experimental” Subject Category) draws upon other domains. Medicine-R&E research shows as reaching out very widely for its knowledge base. This offers empirical support for the notion that fundamental research contributes to medical advances.

We see promise in tracking research knowledge transfers at three tiers: publications (in terms of the audiences being addressed); the cited references upon which they draw; and the articles citing them. This holds promise to help assess the downstream value of fundamental research. We are generating visualizations that depict these knowledge exchanges in terms of the Subject Categories represented (e.g., for articles associated with the funding of a particular set of research projects aimed at engaging various science & engineering research communities with educational issues). Such research mapping could aid R&D policy makers in formulating support strategies. It could also help anticipate applications of ongoing research endeavors in different domains.
Figure B7. Medicine-Research & Experimental-2005 Citation Pattern

Figure B8. Physics—Atomic, Molecular & Chemical-2005 Citation Pattern