Supporting Materials: Part 2 of 3 Scholarship

Shawn P. Gallagher
Department of Psychology

This binder includes:

p. ii-v. Overview of each item.

p. 1-18. Copy of a successful grant application as well as documentation describing the grant and notice of our award for $24,860

p. 19-91. Copies of seven peer-reviewed publications

p. 92-98. Copies of three editor-reviewed publications

Link to Original Film:
sites.millersville.edu/sgallagh/Application/Supporting/Digital_Files/CattellFilm.wmv
Overview of Scholarship since last Promotion
(from the application)

My scholarly endeavors address one or more of four broad domains of expertise: (1) Pedagogy and Curricular Development, (2) Sensation and Perception, (3) History of Psychology, and (4) Clinical Eye Research.

Ongoing Funded Research – Award $24,860

Halpern, B.L., Gallagher, S.P. & Andrews, H. (2014-Present). The Impact of Chemotherapy on Corneal Endothelial Cell Density. This award is a direct result of my involvement in the Cornea Donor Study, a national, long-term investigation of the effectiveness of cornea transplant surgery (see my publication, Sugar et al., 2009). As a collaborating researcher in the Cornea Donor Study, I was alerted to a possible link between cancer and the success of corneal transplants. This led me and my longtime collaborator, Barton L. Halpern, M.D. (see Halpern’s letter of support), to contact a local oncologist, Randall Oyer, M.D., and propose a collaborative research project with physicians at the Ann B. Barshinger Cancer Institute. We made a formal presentation to the oncology staff and they welcomed us into the institute and encouraged us to apply for financial support through the Louise von Hess Medical Research Fund at Lancaster General Hospital. I have included a copy of the proposal as well as a description of the grant and a notice of our award.

Peer-Reviewed Publications (*Student co-authors)


This article is an example of a student realizing the full potential of an honors project. By having my advisees follow a schedule that requires them to defend their thesis one full semester before graduation, they can present their work or, in exceptional cases, publish before leaving Millersville. The journal serves “educators and industry personnel associated with design, presentation, management, and reproduction of graphic forms of communication” and this article alerts them about an unusual perceptual phenomenon that could affect how some people see printed text.


I wrote this piece as part of a series of biographies published for the 130th Anniversary of the American Journal of Psychology. I was the only non-historian invited to contribute and, as you can see, it covers the three topics that I teach most: History, Statistics, and Perception. Edwin G. Boring (1886-1968) shaped our field
as “Psychology’s Historian” as well as a former editor of the American Journal of Psychology, the first journal dedicated to experimental psychology. I was honored to have been trusted with this assignment.


These two papers were published in a journal that serves neuroscience educators. The first (Hallock et al., 2017) is another example of how a well-designed undergraduate project can lead to publication (see Hallock’s letter of support). This simple and inexpensive lab exercise shows how memory experiments can go beyond the memorization of pictures and word lists. The second, highly collaborative, paper (Harrington et al., 2015) aims to help educators incorporate primary source literature in their psychology and neuroscience classes.


This paper is a result of my effort in the domain of departmental assessment. It explains how an expensive assessment test - which is administered by dozens of psychology departments each year – is not living up to its promises (see Carlson’s letter of support).


This paper describes one of the many original experiments that I have developed for my Sensation and Perception course. (see DeRose’s letter of support)


For the past 27 years, I have been collaborating with ophthalmologists to help them improve eye surgery and, from 1999-2013, I was the local coordinator for a nationwide study that evaluated the effectiveness of corneal transplants. In that role, I was responsible for examining cornea transplant recipients who were seen at a local office and sending their data back to the JAEB Center for Health Research.
in Tampa, Florida. My work in that role and my background in statistics led them to trust me as a contributing data analyst and writer for a few of their manuscripts. It also laid the groundwork for my ongoing funded study on the impact of chemotherapy on the cornea.

**Editor-Reviewed Publications**


These contributions are relatively small but significant. The first one is a news item that I published to fulfill a promise that I made to a librarian at the Cajal Institute in Madrid who lamented that Americans psychologists were unaware of the treasures that Spain’s greatest neuroanatomist, Santiago Ramón y Cajal (1852-1934), had left behind.

I must admit that I never thought much of book reviews but, thanks to Research Gate, a social networking website that allows me and other scientists to share our work, I am reminded that my review of *Forty Studies that Changed Psychology* is, by far, my most popular publication on the site.

**Original Documentary Film:** DVDs are included in *Supporting Materials: Scholarship* but the film can also be accessed at: sites.millersville.edu/sgallagh/Application/Digital_Files/CattellFilm.wmv

**Insert.** Gallagher, S.P. (2010). *Memories of the Cattell School.* Presented at the Annual meeting of the Eastern Psychological Association, New York, New York. The film has been adopted by the *Psychology’s Feminist Voices* website, a resource hosted by York University, Toronto, Canada the home of the premier graduate program in the History of Psychology. It has also been accepted into the digital archives of the Cummings Center for the History of Psychology at the University of Akron.

*I did all the interviews, recording, and audio/video editing for this film.*

Every community has a history and, shortly after I started to teach History and Systems of Psychology, I started exploring the local history of Lancaster, Pennsylvania, in hopes of finding something that I could bring back to the classroom. I hit the proverbial jackpot when I discovered that one of the giants of American psychology, James McKeen Cattell (1860-1944), spent the last years of his life in Lancaster where he managed the Science Press from 1939 until his death in 1944. While exploring the lives of his children, who also moved to the area in the 1930s, I was astonished by the accomplishments of his daughter, Psyche (1893-1989), a pioneer and little-known hero in the world of early childhood education.
As I learned more about Dr. Cattell, I felt an urgency to find and interview people who knew her.

Dr. Cattell battled institutionalized sexism to become one of the first women to get a Master’s degree from both Cornell and Harvard and the first woman to earn a doctorate in education from Harvard in 1927. Her academic accomplishments are rivaled by her personal accomplishments and, although I cannot be certain if she was the first, she was certainly among the few unmarried women permitted to adopt children in the 1940s.

After her formal education, she conducted research through the 1930s and, as the decade ended, she set her sights on joining her father and brother in Lancaster and establishing a school dedicated to the education of preschool children. This concept was radical, especially in the conservative climate of prewar Lancaster, Pennsylvania.

This film is about the nursery school that Dr. Cattell established in her home in 1941. I was delighted and honored to find and interview her son, Hudson, who presents his mother’s vision and offered a first-hand account of living in a house that doubled as a school. Former students were also helpful, although most admitted that their memories constituted little more than “cookies and naptime.” Susanna Benson, whose mother struggled to provide for her daughter and sustain an academic career while her husband was at sea, was one dramatic exception. Her stories are detailed and complimented by a wonderful collection of photographs from the earliest days of the school. Two former teachers presented me with very different, but equally flattering, perspectives of Dr. Cattell’s educational approach. One teacher, Janet Strube, provided me with a priceless audiotape of Dr. Cattell addressing a group of parents in the early 1970s. After extensive audio filtering, her messages shine through in the clear, purposeful cadence of an accent that is one part New England and, carrying a trace of her mother’s homeland, one part Old England.

When the school opened, in 1941, Dr. Cattell was proposing a radical new approach to the early childhood experience; when the school closed in 1973, preschool education in the United States had embraced and, in the eyes of one young teacher, perhaps gone far beyond Dr. Cattell’s expectations of what a nursery school could be.
Proposal for funding from the Louise von Hess Medical Research Institute for
The Impact of Chemotherapy on Corneal Endothelial Cell Density

Barton L. Halpern, MD: Principal  
bhalpern@aol.com

Shawn Gallagher, PhD: Co-Investigator  
sgallagher@milersville.edu

Hans Andrews: Co-Investigator

A. Background
We propose the use of specular microscopy (photography of the corneal endothelial cell layer) to test whether chemotherapy damages the endothelial cell layer of the cornea. This study has the great advantages of involving minimal costs and patient time, and no more risk than a photograph.

The corneal endothelium is a cellular monolayer that covers the posterior surface of the cornea. The endothelial cells control fluid transport between the aqueous and cornea, and are responsible for maintaining the slightly dehydrated state that enables corneal transparency. Endothelial cell loss or dysfunction is the primary cause of corneal opacification and the subsequent need for corneal transplantation.

Corneal endothelial cell damage can be directly assessed using a specular microscope which can photograph the corneal endothelium with a quick and painless procedure. The shape and density of the photographed cells are an excellent indicator of corneal function (i.e. of the endothelial cell layers ability to maintain the slightly dehydrated state that enables corneal transparency).

As a general statement, many chemotherapeutic agents work by destroying rapidly dividing cells. The fact that in a clinical (in vivo) setting, corneal endothelial cells do not divide, suggests that chemotherapy is likely to have little to no effect on corneal endothelial cell counts (the null hypothesis). However, indirect evidence suggests that systemic chemotherapy may compromise the function of these mitochondrion-rich cells.

- Many chemotherapy agents cause dose-dependent polyneuropathies that are typically assessed in skin biopsies. Recent evidence (Ferrari et al., 2013) shows similar dose-dependent nerve fiber damage in the small-fiber nociceptive neurons of the cornea. Another case study described limbal stem cell deficiency in a patient who underwent systemic hydroxyurea treatment for sickle cell disease (Ding et al., 2009). This is direct evidence that systemic chemotherapeutic agents can affect corneal function.
- Case studies of cornea transplant patients have shown rapid opacification in corneas from donors who received a full course of systemic antimetabolite therapy (Papathanassious et al., 2010; Van Meter, 2007). Also, a large-scale study of donor corneas points to an increased risk of failure in grafts that came from
cancer patients, compared to those from donors without cancer (Sugar et al., 2009). This is indirect evidence that chemotherapy affects the corneal endothelium.

We believe that we may provide important information to oncologists about the long-term effects of chemotherapy on a non-dividing tissue (corneal endothelial cells), which may have relevance for other non-dividing cells that are inaccessible without biopsy.

This study should provide useful clinical information for tissue eye banks and corneal transplant surgeons as they consider the potential long-term consequences of accepting cornea tissue from patients who have had chemotherapy. Currently, some eye banks leave the decision up to the surgeon as to whether to accept, for purposes of transplantation, a corneal from a deceased chemotherapy patient. At this point, a surgeon’s decision is largely uninformed. A null result would be clinically beneficial, by suggesting to cornea surgeons that corneas donated from patients who have had chemotherapy may be safely used for corneal transplantation, thereby increasing the pool of donor corneas. On the other hand, finding that chemotherapy has a deleterious effect on corneal endothelial cells might cause surgeons to reject some donor corneas, and therefore reduce the pool of available corneas. This would benefit patients receiving donor corneas, as it might improve the survival rate of those corneas that are used for the purpose of transplantation.

To our knowledge, no-one has measured endothelial cell density before and after a chemotherapy regimen to explore the effects of chemotherapy on in vivo endothelial cells. The measurements and statistical analyses are simple and straightforward. Regardless of whether or not the study finds chemotherapy to have a deleterious effect on corneal endothelial cells, the results would be valuable to clinicians.

**B. Research Questions/Aims**

1. Is there a difference in the density of corneal endothelial cells pre and post chemotherapy administration within subjects?
   - We will measure corneal endothelial cell density with a Konan SP 9000 specular microscope. The device is an industry standard that can photograph and digitally store images of the cornea. Software algorithms measure endothelial cell density in cells/mm². Unlike first-generation specular microscopes, the Konan microscope does not come into contact with the eye and multiple clear images of both corneas can be obtained in less than five minutes.

2. Is there a difference in the shape of corneal endothelial cells pre and post chemotherapy administration within subjects?
   - If endothelial cell density changes over time, a secondary concern is the morphological changes in the corneal endothelial cells. These cells, when healthy, have an optimal hexagonal shape. As endothelial cells are lost, the surviving cells will morph to “fill in the gaps.” (Figure 1)
Figure 1. Left: Normal corneal endothelium showing uniformity in endothelial cell size and shape. Right: Corneal endothelial cells showing subjectively sub-clinical pleomorphic changes with greater variability in size and shape. The hexagonal nature of the endothelial cells can be objectively quantified with an algorithm that measures the variability in the distances between the centers of adjacent cells.

Methods

Study Procedures
Patients undergoing cytotoxic chemotherapy will be recruited on a voluntary basis, with the approval of referring oncologists. If the patients are willing to participate, we will obtain informed consent and take a baseline measurement of the central cornea of both eyes. For the benefit of participating patients, we will obtain the photo at the Ann B. Barshinger Cancer Institute

The basic protocol will involve:
Initial Baseline Visit – Before the Initiation of Chemotherapy
1. Excluding patients who have received radiation therapy to the head or neck.
2. Excluding patients receiving non-cytotoxic chemotherapy.
3. Completing a brief medical history form for each new patient. (Attachment 1)
4. Completing an ocular history questionnaire. (Attachment 2)
5. Photographing the corneal endothelium with the specular microscope

One Year and two-Year Visits
6. Completing follow-up medical information form. (Attachment 3)
7. Photographing the corneal endothelium with the specular microscope

Other than the necessary interaction at the time of examination, the only patient contact necessary will be in the scheduling of the three visits. Patients will be notified of the general findings of the study and also will able to discuss their personal results with the primary investigator.

Data Collection
Data collection will involve a medical history (Attachment 1, Attachment 2, Attachment 3), endothelial cell photographs taken at the baseline visit, at one and two yearly intervals. Oncology nurses will be trained by the study team to obtain these pictures.
Endothelial cell count density and cell morphology regularity (the variability of the cell-to-cell distances) will be measured using Konan algorithms. The Konan specular microscope and its associated computer are not accessible through the internet or any wireless devices. Measurements derived from the Konan software will be stored in an Excel file along with patient data.

The hypotheses will be tested using related-measures $t$-tests.

**Inclusion and Exclusion Criteria**

Provided the patient is not inconvenienced by the visit, we invite participation of any adult (over 18 years) patient with cancer beginning a new regimen of cytotoxic chemotherapy. The only ocular exclusion criteria would involve eye trauma (including a history of eye surgery) that are known to influence endothelial cell density. Patients with an ENT or brain related malignancy or patients with metastases involving the brain or head and neck requiring a radiation or surgical intervention will be excluded from the study. The specific treatments (radiation/surgery) typically given to these patients may bias the data.

**Sample Size Calculations and Justification**

We aim to have a minimum of 50 patients complete the pretest and posttest measurements. Since we have no specific time constraint, we are willing to continue the project and recruit more patients to gain a larger sample size, as long as we have the support of the participating oncologists. We welcome their input in helping us to determine the willingness of patients to participate as well as information about the expected drop-out rates of particular patient subgroups.

In the population of healthy adults, the mean endothelial density is 2700 cells/mm$^2$ with a standard deviation of 250 cells/mm$^2$. Adults experience an estimated 0.5% annual decrease in cell density. Furthermore, patients with identified corneal disease can lose up to 10% before experiencing subjective changes in vision. Therefore, if we conservatively predict a 3% decrease in pretest-posttest endothelial cell density (Cohen’s $d = .33$), our study would have $\beta = .24$, using one-tailed $t$-tests with $\alpha = .05$. This is, of course, only an estimate and because few have explored this topic, it is difficult to anticipate a specific decrease in endothelial cell density. Furthermore, different drugs may have different effects on endothelial cell density and our preliminary data might help us identify drugs that may have larger effects than others.
Statistical Analyses
We will employ repeated-measures t-tests (one-tailed, α = .05) or repeated-measures analyses of variance for patients who are available for multiple (more than two) measurements.

D. Human Subjects Research

- **Patient privacy and confidentiality** will be maintained through the secure filing of physical records as well as digital records. Images of the corneal endothelium will be stored on a computer with no internet or wireless communication access and the computer will remain in a medical office setting. Final master data files will be organized according to a patient code number, without names.
- The procedure uses a non-contact, no-flash medical camera, and there is no risk of participating in the testing. The procedure requires the patient to rest his or her chin on a stabilizing shelf while looking into an opening of a table-top box measuring approximately 12” wide, 18” deep, and 18” tall. If the patient can remain stationary, multiple images from each eye can be collected in less than 5 minutes.
- Stored information that includes patient information will be secured on the premises of a medical office (physical records) or a computer without internet access. Master data files and any published or otherwise publicized results will be coded with no identifying patient information.
- Recruiting and consenting patients
- Patients may discontinue participation and/or withdraw consent to use their data at any time. Their physical records will be destroyed and all digital information will be deleted.

E. References


Attachment 1
Medical History for Specular Microscopy Study prior to Chemotherapy
Initial Baseline Visit

Exclusion criteria include (1) non-cytotoxic chemotherapy, (2) an ENT or brain related malignancy or metastases involving the brain or head and neck requiring a radiation or surgical intervention will be excluded from the study.

Patient Name:
LGH Patient Code:
Date :
Patient Phone:
Patient Address:
Patient e-mail address:
Preferred Mode of Contact:
DOB:
Gender:
Eye History:
   Medical
   Surgical
   Eye medications
Medical History:
Primary Cancer Diagnosis and stage (warranting chemotherapy):
Medication History (other than eye):
Chemotherapy Regimen:
Attachment 2

OCULAR HISTORY QUESTIONNAIRE
Initial baseline Visit

Please answer the following questions about your eyes by circling the appropriate Yes or No answer. Use the space at the bottom if you need additional space to explain your responses.

1. The cornea is the clear part of the front of your eye. Do you have a corneal disease? Yes / No
   If you circled “Yes” please explain:

2. Have you ever had surgery on your eyes? Yes / No
   If you circled “Yes” please explain:

3. Have you ever had an injury to your eye that required medical attention? Yes / No
   If you circled “Yes” please explain:

4. Do you have an eye problem or disease that requires regular checkups with an eye doctor? Yes / No
   If you circled “Yes” please explain:

5. Are you taking any eye medications? Yes/No
   If you circled “Yes” please explain. ________________________


Attachment 3: 
Medical History for Specular Microscopy Study Post Chemotherapy 
“Repeat” at One and Two-Year Visits

Patient Name:
LGH Patient Code:
Phone:
Address:
E-mail address:
Preferred mode of contact:
DOB:
Gender:
Date:

Has the patient had any ocular injury or surgery since the beginning of chemotherapy (Surgery, Disease, Eye medication, Trauma)?

Has the patient had any significant changes in medical history since starting chemotherapy?

Primary Cancer Diagnosis Stage/Status:

Did the patient receive the prescribed chemotherapy treatments?

If no, how many months of treatment did they receive?

Is the patient on any maintenance chemotherapy? (Yes/No)

If YES what is the patient receiving?
Background
The Lancaster General Research Institute was established during the early part of 2004. It was established to conduct and provide support for research across the Lancaster General system. Lancaster General advances the health and quality of life of the individuals and communities we serve by providing superior health care that demonstrates community benefit.

A large part of the mission of the Lancaster General Research Institute is providing support for physician initiated and directed research. The institute supports investigators with procurement of grants, research design, research implementation, statistical analysis and preparation for dissemination for their research endeavors. These projects typically involve community physicians, current faculty, residents, medical students, and undergraduate pre-medical students in various roles. Research projects provide significant learning opportunities for those involved and have provided knowledge that directly leads to improvements and increased level of care for the Lancaster Community.

In April of 2006, the Louise von Hess Medical Research Institute was formed as a named entity within the Lancaster General Research Institute to specifically facilitate and support physician directed research. The focus of physician directed research within the Louise von Hess Medical Research Institute is upon education, advancement of the practice of medicine, application of research to medicine, and ultimately improving the health of the Lancaster and surrounding community. An endowment was established to help support future research involving community physicians, current faculty, residents, medical students, and undergraduate pre-medical students. The endowment draws only upon interest/dividends and the principal remains untouched in-perpetuity.

Application Deadline
No deadline. We operate on a rolling schedule and will schedule meetings to review proposals as they are submitted.

Purpose
The Louise von Hess Medical Research Institute focused upon supporting and encouraging physician learning through research. Its primary emphasis is upon education, advancement of the practice of medicine, application of research to medicine, and ultimately improving the health of the Lancaster and surrounding community. This endowment is
to help support research involving community physicians, current faculty, residents, medical students, and undergraduate pre-medical students. Funded projects must show evidence of IRB approval and meet applicable legal and ethical research practice guidelines.

How to Apply
Applications must be submitted by the listed deadline in electronic MS Word format. No hard copy applications will be accepted. Applicants are required to submit proposals of no more than 10 double-spaced pages for review. This page limit includes any letters of support, references, appendices or other supporting documents.

The proposal should contain the following information about the proposed project:
- A statement of the project’s principal objectives.
- A description of the research methodology to be used.
- A description of how the project’s findings would complement related work already completed or currently under way, as well as how it would contribute to the advancement of medicine and improving health in Lancaster and the surrounding community.
- Dissemination plans, including target audiences.
- An estimated timetable and budget for completion of the project.
- The qualifications of the applicant and key project staff members.
- The name of the primary contact person.

Eligibility
Oversight and project funding support determinations are made by a research oversight committee. Funded projects must:
- Demonstrate a benefit to the Lancaster Community
- Focus upon providing improved and high levels of care and outcomes
- Demonstrate a mechanism of education through dissemination of results, involvement of pre-medical students, medical students, residents, faculty or community physicians or as a part of the purpose of the study
- Involve in the context of the project of pre-medical students, medical students, residents, faculty, or community physicians

Research Oversight Committee
The research oversight committee of the Louise von Hess Medical Research Institute will be made up of the following five members:
1. Chief Medical Officer/Sr. VP of Medical Affairs of Lancaster General
2. Vice President of Development at Lancaster General
3. Director of Research at Lancaster General
4. Former Board Member, Louise von Hess Foundation
5. Physician on the Medical Staff at Lancaster General appointed by the President of the Medical Staff

Lancaster General Research Institute
Lancaster General Hospital • 555 North Duke Street • Lancaster, PA 17604
(717) 544-5999 • Fax (717) 544-7960 • www.LancasterGeneral.org
At the present, the research oversight committee of the Louise von Hess Medical Research Institute consists of the following members:

1. Monty Duke, MD (Chief Medical Officer/Sr. VP of Medical Affairs of Lancaster General)
2. Jay Bucher (Vice President of Development at Lancaster General)
3. Michael Horst, PhD (Director of Research at Lancaster General)
4. Samuel Rice, MD (Former Board Member, Louise von Hess Foundation)
5. Nikitas Zervanos, MD (Physician on the Medical Staff at Lancaster General)

**Selection Criteria**

Selection criteria are based upon the four eligibility indicators listed above and the overall quality of the proposal. Each will be assessed on a 5-point Likert-type scale by members of the research oversight committee. Scores for proposals can range from 5 to 25 and results will not be shared outside of the research oversight committee. The following table highlights the selection criteria scoring system.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Exceeds Criteria</th>
<th>Meets Criteria</th>
<th>Does Not Meet Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposal – quality, clarity and completeness of proposal</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Demonstrate a benefit to the Lancaster Community</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Focus upon providing improved and high levels of care and outcomes</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Demonstrate a mechanism of education through dissemination of results, involvement of pre-medical students, medical students, residents, faculty or community physicians or as a part of the purpose of the study</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Involve in the context of the project of pre-medical students, medical students, residents, faculty, or community physicians</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Upon initial review by the oversight committee, proposals may be returned to the applicant with questions for additional consideration with a request for re-submission. Responses to questions or applicable changes to the proposal will be re-considered by the oversight committee.

**Available Funding**

Total available funding may vary due to financial performance of the endowment. The goal is to provide funding to more than one project, but exceptions can be made as deemed appropriate by the oversight committee. Funding requests must not exceed $25,000 and may be as small as $1,000. Only one award will be made to each medical specialty area with the exception of projects submitted by residents or fellows. Funding
would be based upon a calendar year from the time the award is approved.

Unused funds would be returned to the endowment principal or used to improve research infrastructure within the Louise von Hess Medical Research Institute. These improvements to research infrastructure may include purchase/upgrade of software to support research projects, upgraded/replacement/new computers for the research institute, training for research institute staff or summer research internships for medical students.

**Use of Grant Funds**
Grant funds may be used to support project staff salaries, consultant fees, data processing, supplies and other direct expenses, including a limited amount of equipment essential to the proposed project. Grant funds may not be used to subsidize individuals for the cost of health care, to construct or renovate facilities, for lobbying, for indirect expenses/overhead (typically a percentage of the grant at academic centers) or as a substitute for funds currently being used to support similar activities.

**Completion of the Project**
Upon completion of the project or at the end of the fiscal year of funding (June 30), a final report is required to document the outcomes or progress of the funded project. The report must be submitted by October 1 in electronic MS Word format. No hard copy reports will be accepted. The final report should be no more than 5 double-spaced pages. The Louise von Hess Medical Research Institute may wish to publicize funded projects and/or outcomes and will work with investigators to outline appropriate parameters.

The final report should contain the following information:
- A statement of the project’s principal objectives.
- A description of the research methodology used with a particular emphasis upon any changes from the original proposal.
- Discussion of the outcomes, findings or progress in the research project.
- A discussion of how the project’s findings complement related work already completed or currently under way, as well as how it would contribute to the advancement of medicine and improving health in Lancaster and the surrounding community.
- Fiscal report on how the grant funds were utilized.
- Evidence of IRB approval.
- The name of the primary contact person and other key persons involved in the project.
Questions
Michael A. Horst, PhD, Director of Research
Lancaster General Research Institute
Lancaster General Hospital
717-544-5999
mahorst@lancastergeneral.org
2007

The Role of the Family Physician in the Development of Healthy Eating Habits
Michelle A. Cardona, MD, MPH: Lancaster General Hospital, Resident Physician - Department of Family and Community Medicine

Purpose: To investigate the relationship of physician and family influences on child eating behaviors and determine if differences exist based on physicians' and patients' race, age, or gender, or physicians' practice location.

Award Amount: $15,000

The Family and Community Medicine Residency Lactation Curriculum project (RLCP): Year 1 of Project
Gladys M. Frye, MD: Associate Director Family Medicine Residency Program at LGH

Purpose: Training our new physicians to effectively manage breastfeeding issues and encourage mothers to breastfeed. Show the effectiveness of a breastfeeding curriculum implemented in the Family Practice Residency Program at Lancaster General Hospital in terms of the knowledge base, attitudes, and skills of the residents and explore if it has any effect on the breastfeeding initiation rates and 6 month continuation rates of our Family Health and Quarryville patients delivered at Women and Babies Hospital.

Award Amount: $13,340

2008

The Family and Community Medicine Residency Lactation Curriculum project (RLCP): Year 2 of Project
Gladys M. Frye, MD: Associate Director Family Medicine Residency Program at LGH

Purpose: Training our new physicians to effectively manage breastfeeding issues and encourage mothers to breastfeed. Show the effectiveness of a breastfeeding curriculum implemented in the Family Practice Residency Program at Lancaster General Hospital in terms of the knowledge base, attitudes, and skills of the residents and explore if it has any effect on the breastfeeding initiation rates and 6 month continuation rates of our Family Health and Quarryville patients delivered at Women and Babies Hospital.

Award Amount: $16,850

Hospice Use in End-Stage Renal Disease
Laurence E. Carroll, MD: Hypertension and Kidney Specialists
Joan K. Harrold, MD: Hospice of Lancaster County
Michael A. Horst, PhD: Lancaster General Research Institute

Purpose: Define overall hospice use in ESRD (Dialysis) patients in Lancaster County. Determine the hospice use before and after withdrawal from ESRD Rx (Dialysis & Transplant) in Lancaster County. Identify how many patients defer any ESRD Rx in Lancaster County and do they use hospice services. Use
data gathered from this study in Lancaster County to compare our rates of hospice use and dialysis discontinuation to identify strategies for improvement, potential interventions, outcome measures and a study design to test the proposed interventions which would be the next phase of this project.

Award Amount: $10,500

2009
Lancaster County Colorectal Cancer Outcomes Study
Randall A. Oyer, MD: Hematology Oncology Associates of Lancaster

Purpose: The purpose of this initial phase of the project is to conduct a study in conjunction with Franklin and Marshall College to create a project baseline and estimate the current colorectal cancer screening rate for Lancaster County since the last data available for this area was 2002. These data will be used in conjunction with a larger project to improve colorectal cancer screening rates in Lancaster County.

Award Amount: $7,900

2010
The Effect of a Rapid In-office Test on Antibiotic Prescribing Practices for Acute Infectious Conjunctivitis
Amanda S. Davis, MD: Lancaster General Hospital, Resident Physician - Department of Family and Community Medicine

Purpose: To evaluate how providers’ antibiotic prescribing rates for acute conjunctivitis are affected by the availability of a rapid, in-office test for adenovirus, the most common viral pathogen.

Award Amount: $24,850

2011
INDUCE-it Trial: The Effect of Inducing Cytochrome P450 on the Efficacy of Clopidogrel in Patients with Reduced-Function CYP2C19 Polymorphisms Undergoing Percutaneous Coronary Intervention
Kathy Makkar, PharmD1, Rupal Dumasia, MD2, Jill Rebuck, PharmD1, Michael Horst, PhD3, Yee Lee, PharmD1, Richard Paoletti, RPh1, Roy Small, MD2
1Department of Pharmacy, 2Division of Cardiology, 3Research Institute; Lancaster General Hospital, Lancaster, PA

Purpose: myocardial infarction (MI), stroke, or death from cardiovascular causes in patients with acute coronary syndrome (ACS). Nearly 30% of the population are genetically unable to adequately metabolize clopidogrel via the CYP2C19 hepatic enzyme into its active form and have a 32.4% reduction in plasma exposure to the active metabolite and diminished platelet inhibition. These patients may be at increased risk of death from any cause, nonfatal MI, and stroke. St. Johns wort is a known inducer of the CYP2C19 enzyme. Our hypothesis will be rigorously studied by evaluating the following specific study aims in patients that are classified as clopidogrel reduced-metabolizers:
1. To identify the difference in platelet reactivity in patients receiving St. John’s wort or placebo
2. To characterize the difference in platelet inhibition in patients receiving St. John’s wort or placebo
The study is a prospective, randomized, double-blind, placebo-controlled, cross-over study of patients post percutaneous coronary intervention who require dual-antiplatelet therapy with aspirin and clopidogrel.

Award Amount: $7,832

2012
The Study of Gene Signatures in Melanoma
Shanthi Sivendran, MD, MSCR
1Hematology/Oncology Specialists

Purpose: The objective of this research project was to establish an immune signature for Stage II and II cutaneous melanoma at high risk of recurrence. This is to be accomplished through four aims: (1) screen dermatology databases at Mount Sinai and Geisinger to identify melanoma specimens from patients who recurred/not recur; (2) establish a protocol for extraction of RNA; (3) establish an inflammatory signature for early stage melanoma at high risk of recurrence and (4) validate the gene signature from the discovery set in a validation set of recurrent and nonrecurrent stage II/III cutaneous melanoma specimens.

Award Amount: $9,000

2013
Incidence of Breast Cancer Relumpectomies
Daleela Dodge, MD, Hannah Groff, Michael Horst, PhD
1LG Health Physicians Surgical Group, 2Franklin and Marshall College, 3LG Research Institute

Purpose: To determine the rate of breast cancer relumpectomies at Lancaster General Health during a recent and historical period in order to demonstrate the hospital’s changes in surgical technique and imaging during diagnosis and in the operating room. Additionally, we will examine the range of surgical tumor margins used as an indication for re-excision.

Award Amount: $4,000

2014
The Impact of Chemotherapy on Corneal Endothelial Cell Density
Barton Halpern, MD, Shawn Gallagher, PhD, Hans Andrews
1Eye Doctors of Lancaster, 2Millersville University, 3Penn State College of Medicine

Purpose: We propose the use of specular microscopy (photography of the corneal endothelial cell layer) to test whether chemotherapy damages the endothelial cell layer of the cornea. We believe that we may provide important information to oncologists about the long-term effects of chemotherapy on a non-dividing tissue (corneal endothelial cells), which may have relevance for other non-dividing cells that are inaccessible without biopsy. This study should provide useful clinical information for tissue eye banks and corneal transplant surgeons as they consider the potential long-term consequences of accepting cornea tissue from patients who have had chemotherapy. Regardless of whether or not the study finds
chemotherapy to have a deleterious effect on corneal endothelial cells, the results would be valuable to clinicians.

Award Amount: $24,860

**The Effect of Time to INR Reversal on TBI Outcome**
Fred Rogers, MD\(^1\), Hans Andrews\(^2\), Katelyn Rittenhouse\(^3\)
\(^1\)Lancaster General Trauma, \(^2\)Penn State College of Medicine, \(^3\)University of North Carolina School of Medicine

Purpose: To determine the effect of time to reversal of elevated international normalized ratio (INR) on outcome in anticoagulated trauma patients. Further, we will conduct subanalyses of patients with therapeutic (INR2-3) and supratherapeutic (INR>3) INRs to determine if time to INR reversal affects these populations differently.

Award Amount: $3,600

**Support of Internship to Help Family Medicine Residents and Faculty Complete Several Research Projects**
Peggy Nepps, PsyD\(^1\)
\(^1\)Lancaster General Family Medicine Residency

Purpose: To provide an opportunity for a pre-medical college student to assist family medicine residents and residency faculty at Lancaster General Hospital in the implementation of their research projects as well as an opportunity to learn more about careers in medicine by means of interaction with residents and shadowing with Lancaster General health professionals. The intern will likewise have the opportunity to participate in a number of daily educational sessions held for the residents and medical students at LGH. The program would involve having the student assist with several specific research projects:

- Comparison of Smoking Cessation Rates Before and After Implementation of the Patient Centered Medical Home
- The Relationship between Obesity and Food Security
- Does Resident Education on Childhood Obesity Affect Recommended Time to Follow-up of These Patients in a Family Medicine Office?
- Assessment of the Breastfeeding Self-Efficacy Scale Short Form (BSES-SF) in a Family Medicine Residency Program Patient Population
- Mindfulness Meditation as Adjunctive Treatment for Anxiety Disorders.

Award Amount: $3,500

**Symptom Management, Quality of Life and Satisfaction with Care for Advanced Stage Cancers**
Shanthi Sivendran, MD\(^1\), Randy Oyer, MD\(^1\), Nikolas Buescher\(^1\), Joan Harrold, MD\(^2\), Michael Horst, PhD\(^3\), Barbara Martin, PhD\(^3\), Kristina Newport, MD\(^2\)
\(^1\)Lancaster General Ann B. Barshinger Cancer Institute, \(^2\)Palliative Medicine Consultants, \(^3\)Lancaster General Research Institute
Purpose: Symptom Management, Quality of Life and Satisfaction with Care for Advanced Stage Cancers is the first part of a two part study (The Lancaster Cancer Care Model (LCCM) – Non-Concurrent Control Study) that was submitted to the National Cancer Institute (NCI) for funding. A decision on funding the full study should be forthcoming in the early part of 2015. The primary aim of the study is to compare the proportion of advanced cancer patients who have a hospitalization or emergency department visit in the last 6 months of life before and after implementation of a new care model that provides more comprehensive symptom management and supportive care, including earlier referral to palliative care. The secondary comparative aim is to assess measures of quality of life and satisfaction in both groups.

Award Amount: $93,120 over 2 years

**Fitbit/Healthy Weight Management Study**
James Ku, MD\(^1\), Joseph McPhee, MD\(^1\), Lawrence Wieger, DO\(^1\), Steven Johnson\(^1\), Michael Horst, PhD\(^2\), James Polinsky\(^3\)
\(^1\)Lancaster General Healthy Weight Management Center, \(^2\)Lancaster General Research Institute, \(^3\)Lancaster General Virtual and Mobile Health

Purpose: Specifically with our Healthy Weight Management patients that are pursuing sleeve gastrectomy surgery, we would like to determine if the Fitbit is associated with health outcomes in this population. The primary outcome is the percentage of excess weight loss with secondary outcomes of lifestyle/exercise pattern change and resolution of comorbidities.

Award Amount: $99,399 over 2 years
Performance of Grapheme-Color Synesthetes on a Color Sorting Task that Employ Graphemes

Faith McConnell¹
Shawn P. Gallagher, Ph.D.¹
Mark Snyder, Ed.D.²

Millersville University of Pennsylvania

¹Department of Psychology, Millersville University of Pennsylvania, Millersville, PA 17551
²Department of Applied Engineering, Safety, & Technology, Millersville University of Pennsylvania, Millersville, PA 17551

Author Notes

Correspondence should be addressed to Mark Snyder, Applied Engineering, Safety & Technology, 202-D Osburn Hall, Millersville University, PO Box 1002, Millersville, PA 17551. e-mail: Mark.Snyder@Millersville.edu.

This study was presented in part at the 87th annual meeting of the Eastern Psychological Association, New York, 2016.
Performance of Grapheme-Color Synesthetes on a Color Sorting Task that Employs Graphemes

Introduction

Color perception is not universal and, for the people who do have color vision, it still may vary. As many as 8% of men have an inherited deficiency in color perception and a small fraction of women (less than 1%) have the ability to see an unusually wide variety of colors (Gegenfurtner & Sharpe, 1999). This variation is often due to differences in the anatomy of the human eye (e.g. photoreceptor deficiencies) but some of it is due to how the brain processes color. Some people even experience specific, vivid colors when viewing particular printed letters, regardless of the text’s color. These people have one form of an unusual condition called synesthesia. Brain scientists believe that these strange experiences are due to “crossed wires” in the brain that activate color detecting parts of the brain whenever specific letters or numbers are viewed (Ramachandran & Hubbard, 2001). Although unusual, synesthesia is not rare and may affect as many as 1 in 200 people (Ramachandran & Hubbard, 2001). Few of these people (called synesthetes), however, realize that their experiences are atypical, but some know that they are unusually particular about color choices and select font colors that “fit” the characters and symbols in their creative works (unpublished observations). For example, one of our student synesthetes printed his resume using only dark green and brown text because, according to him, these colors “matched” his two initials which appeared in bold capitalized text at the top of the document.

Our aim is to draw attention to this fascinating phenomenon and determine if synesthesia shapes the perception of text, or if it is an intermittent experience that can be “tuned out” when,
for example, the demands of a printing project contradict a synesthete’s perception of congruence. We believe that our findings shed light on a little-known phenomenon that may be affecting the subjective nature of design and experience.

**Testing Color Vision**

The human eye contains two kinds of light-sensitive cells that convert light to a neural message that the brain can process. Rods are cells that operate in low light conditions and provide no color information; cones function when light is abundant and create the foundation for color vision. Color perception, therefore, starts with the cones in our eyes which respond best to short (usually seen as blue), medium (usually seen as green or orange), and long-wavelength (usually seen as red) light. Fewer than five percent of people have atypical color vision due to non-functioning cones and, at least in the United States, these people are often diagnosed as children during routine eye exams. Other differences in color perception are more difficult to identify. Although most people see color using the same three cones, the cortical (or brain-based) part of color perception is equally important. Cortical processing, like image processing software, may compromise or enhance color vision in ways that are more difficult to detect.

The Farnsworth-Munsell 100 Hue Test (100 Hue Test, X-Rite, Grand Rapids MI) uses color-sorting to measure color discrimination ability (Figure 1). The test can, of course, diagnose typical forms of color blindness, but it can also be used to monitor slow changes in color perception caused by eye diseases. The test consists of four trays of 85 sortable colored caps that span the visible spectrum and the objective is to arrange the caps into a spectrum of hues that varies progressively from one color to the next. Interested readers can explore a publicly available version of the task by Daniel Flück ([www.color-blindness.com](http://www.color-blindness.com)). This computer-based
test is automatically administered and scored but is otherwise like the original. Instead of manually sorting disks on a table, participants use a computer mouse to sort colored tiles displayed on a monitor (Figure 1).

**Synesthesia**

Synesthesia is a perceptual phenomenon in which stimuli in one sensory modality evoke experiences in another. For example, some people experience specific colors when they hear specific notes played on a piano; others associate tastes with shapes. Grapheme-color synesthesia may be the most common and easiest to objectively verify and, like color blindness, is often heritable (Ramachandran & Hubbard, 2001). When an individual has grapheme-color synesthesia, they experience color when viewing certain, but not necessarily all, graphemes (letters, numbers, or other printed symbols) even when they are printed in black on white paper. For example, if a grapheme-color synesthete is shown the letter “A” they may experience a red glow, or photism, around the letter but fully realize that the color they “see” is in their mind, not on the paper. Synesthesia has been recognized for more than a century (Galton, 1880), but only recently have scientists been able to validate these experiences by showing synesthetes perform exceptionally well on tasks that involve visually searching for specific letters that “pop-out in color” when printed in black and white (Ramachandran & Hubbard, 2001). For example, finding a letter “F” in a sheet full of “Es” is difficult for most people, but if a synesthete associates “F” with red and “E” with green, the task is as easy as spotting the only ripe apple in a tree.

Eagleman, Kagan, Nelson and Sarma (2007) have demonstrated that the strength of synesthetic experiences can be measured with computerized tests that repeatedly present participants with graphemes while asking them to choose the associated color from an enormous
array of hues. The strength of the synesthetic association is quantified from the consistency in color choices across multiple presentations (Figure 2). Synesthetes find such tasks easy and are reliable in the way they match specific letters with specific colors; non-synesthetes find these tasks impossible.

One proposed reason as to why synesthetes see colors when looking at graphemes might be that stray neural connections are linking and accidentally activating neighboring brain regions (Ramachandran & Hubbard, 2001). Two brain regions believed to play a role in synesthesia are the V4 color center (so called because it is the *fourth* in a group of visual processing areas) and the nearby posterior temporal grapheme area (PTGA), which is active when people view numbers and text. Neurons in the V4 color center are commonly activated in response to color but stray connections from the PTGA might reach V4 and trigger the sensation of color when colorless letters are viewed (Nunn, Gregory, Brammer, Williams, Parslow, Morgan, & Gray, 2002).

Currently, researchers and vision scientists still do not know if this miswiring offers synesthetes an advantage or disadvantage. Bannissy et al. (2009) found that, when graphemes were not involved, synesthetes demonstrated enhanced color discrimination abilities as measured by the Farnsworth-Munsell 100 Hue Test. The researchers used the cap-sorting task in the typical fashion and found that grapheme-color synesthetes have unusually good color discrimination skills, perhaps as a result of having more brain regions involved in color processing.

Although synesthesia might improve color discrimination when graphemes are not involved, it might create a disadvantage when synesthetes must sort or arrange colored text. Smilek, Dixon, Cudahy, and Merikle (2001) showed how synesthesia can lead to confusion in some specific situations. They instructed synesthetes to find specific characters, or “targets,” presented
against backgrounds that were either congruent or incongruent with the color of the target’s photism. They found that, for example, if a synesthete experienced the digit "5" as red, it was easy for them to spot it among other digits that evoked the experience of green, provided the digits were displayed on a white background. However, they also found that the same task was difficult when the digits were presented on a background that was congruent with the color of the target’s photism; the same red photism that made the digit "pop out" on a white background could camouflage it on a red background.

**Purpose of the Study**

Color-grapheme synesthetes are good at color sorting tasks, but certain graphemes might confuse their perceptual abilities in specific situations. The purpose of this study was to determine if photisms can affect a synesthete’s performance on a color discrimination task when it involves sorting colored graphemes, rather than caps or blocks. Synesthesia may give graphic designers better color discrimination skills when they are sorting and arranging non-grapheme objects, however, graphemes might evoke photisms that can confuse a synesthete’s ability to discriminate or match colors. This effect could have tremendous implications each time a synesthetic designer choses a font color for either print or video display. The current study combined themes from the previously mentioned research and examined grapheme-color synesthetes’ performances on color hue sorting tests that used graphemes chosen to minimize and maximize the odds of color confusion.

Our first task was to develop a computerized Grapheme Hue Test so our participants would be able to sort colored “A”s and “B”s or “4”s and “5”s instead of colored disks (as in the original 100 Hue Test) or colored square tiles (as in the computer-based version of the 100 Hue Test). If
the synesthetes had great difficulty sorting, for example, an array of “A”s that were displayed in shades of green, it could be because the actual printed colors don’t match the color of the letter’s associated photism. If a particular grapheme was among those that did not generate a specific colored photism, we would expect the synesthete to have no difficulty in sorting colored letters. Such findings would support the theory that synesthesia can affect color perception and discrimination skills when the affected individual is manipulating colored graphemes. We predicted that synesthetes would commit more errors and require more time when sorting colored graphemes that generate photisms than when sorting colored graphemes that do not.

**Method**

**Participants**

This study was conducted with the approval of the Institutional Review Board. Candidate synesthetes were recruited from a population of undergraduates and classified using the system developed by Eagleman et al. (2007) through the Synesthesia Battery that is publicly available at [www.synesthete.org](http://www.synesthete.org) (Figure 2). Six confirmed synesthetes performed the full battery of tests and received $25 gift cards for their time.

**Procedures**

We used Adobe Illustrator (Adobe Systems, San Jose CA) to create the Grapheme Hue Test which generated grapheme arrays in colors that matched those used in the computer-based 100 Hue Test (Figure 3). We also made our test shorter than the original by increasing the step increments and employing 36 sortable elements, rather than 85. Finally, we validated our test by administering it to eight participants who had also taken the computer-based 100 Hue Test.
Participant scores on the original 100 Hue test were highly correlated with scores on our Grapheme Hue Test, $r(6) = .97$, $p < 0.01$.

Figure 3 shows one unsorted set of graphemes in the Grapheme Hue Test. The synesthetes had four grapheme sets to sort; first, they sorted two scrambled sets of graphemes for which they had no photisms, these were control trials, and then two for which they did, these were photism trials.

To score performance on the Grapheme Hue Tests, we assigned each grapheme a number (not visible to the participant) that represented where it should fall in a perfectly sorted array. We then calculated a deviation scores for each trial per the methods described by Bannissy et al. (2009). Lower scores indicated fewer errors and a score of 0 indicated that all graphemes had been perfectly sorted. We recorded times and deviation scores for each trial.

**Results**

We hypothesized that synesthesia could create confusion in a color sorting task and that synesthetes would have higher deviation scores in the photism trials than the control trials. A one-way ANOVA demonstrated no significant differences in mean deviation scores across the four trials, $F(3,20)=.38$, $p > .05$ (Figure 4).

We also hypothesized that synesthetes would require more time to sort colors in the photism condition than they would for the control condition. This hypothesis could not be supported. A one-way ANOVA revealed a significant difference among the four conditions $F(3,20)=4.00$, $p<.05$, but a post hoc Tukey test revealed that the difference was only between the first control condition and each of the remaining three. This result was contrary to our expectations and most likely shows a practice effect; participants struggled a bit while learning the task during the first
trial, but improved on subsequent trials (Figure 5). Switching the participants from the control to the photism trials did not lead to a sudden decrease in performance.

**Discussion**

We developed a grapheme sorting task that effectively tested color discrimination ability but, contrary to our expectations, synesthetes did not differ in their ability to sort photism and non-photism graphemes. The results of this study suggest that synesthetes quickly adapt to their photisms or that they can ignore them when necessary. Although we cannot conclude that synesthesia has *no* effect on one’s ability to sort or discriminate printed graphemes, these effects are probably small and certainly would not preclude one from a career in graphic design. When one synesthete in this study was asked if our sorting tasks were difficult, he replied that he could “eventually ignore the shape of the letters and only focus on the colors.” Almost all of our participants reported a similar ability. Although synesthetes perceive photisms and associate specific graphemes with specific colors, our data give us no reason to suspect that synesthesia significantly compromises color discrimination skills.
References


Figure 1. The cap-sorting (top) and the computer-based version (bottom) of the Farnsworth-Munsell 100 Hue Test.
Figure 2. Screenshot from the Synesthesia Battery (Eagleman et al., 2007).
Figure 3. Grapheme Hue Test modelled on the computer-based Farnsworth-Munsell Color Hue Test.
Figure 4. Mean deviation scores for participants in each phase of the grapheme sorting task. A one-way ANOVA demonstrated no significant difference across conditions, $F(3, 20) = .38, p > 0.05.$
Figure 5. Mean time required for participants to complete each phase of the grapheme sorting tasks. Although a one-way ANOVA demonstrated a significant difference across the four conditions, $F(3, 20) = 4.00, p < 0.05$, a post-hoc Tukey test showed no significant differences among Control 2 and the Photism 1 and Photism 2 conditions ($HSD = 161.29$ sec).
Glossary of Terms

**Cortical** — involving or resulting from the action or condition of the cerebral cortex — that part of the brain that functions chiefly in the coordination of sensory and motor information.

**Grapheme** — a unit within a writing system—such as letters and numbers.

**Grapheme-color Synesthesia** — a person with grapheme-color synesthesia will associate colors with letters and numbers involuntarily. For example, when shown the letter “A” they may sense red.

**Photism** — a synesthetic visual sensation. To synesthetes, it is an involuntary, consistent and memorable response.

**Photoreceptors** — a receptor for light stimuli. There are two types of photoreceptors in the human eye: rods and cones.

**Synesthesia** — a neurological phenomenon in which stimulation of one sensory or cognitive pathway leads to automatic, involuntary experiences in a second sensory or cognitive pathway. For example, it can involve associations between letters, shapes, colors, tastes, smells, etc. People who experience these crossover associations are known as synesthetes.
Note: This statement by the editor of the *American Journal of Psychology* explains the series of articles that includes my contribution on E.G. Boring which begins on the following page.
Edwin G. Boring: The Historian’s Path in the Pages of The American Journal of Psychology

SHAWN P. GALLAGHER
Millersville University of Pennsylvania

Although he is best known for his classic textbook, A History of Experimental Psychology, Edwin Garrigues Boring published dozens of articles in The American Journal of Psychology and used its various formats to guide the discipline in the early 20th century. This report reviews a small sample of his publications, including obituaries, notes, and experimental articles, and presents them in historical and biographical context. A central objective is to show how Boring shared the values of his structuralist training with the emerging American schools and how time allowed him to reconsider his approach to history and the legacy of his iconic mentor, Edward Bradford Titchener.

KEYWORDS: Edwin Garrigues Boring, Edward Bradford Titchener, structuralism, American Journal of Psychology, history of psychology

To write about Edwin Garrigues Boring is to imagine him sitting in his cramped third-floor office in Harvard’s Emerson Hall, waiting to address your 4,000-word manuscript with 8,000 words of criticism. In a passionate homage to Boring, his protégé Stanley Smith Stevens (1968) described a multifaceted Boring who valued, above all else, expressive literacy. In the 18-page obituary, Stevens offered at least four pages that described Boring’s commitment to writing, editing, and criticism. The description was not disproportionate. Boring came of age when psychology seemed unmoored and fractured, and he wrote to provide cohesion and direction. For the first half of the 20th century, The American Journal of Psychology (AJP) carried his message.

This is not a comprehensive biography, and I write knowing that Boring had a specific definition of scientific psychology (Kelly, 1981), that he was a politically complex individual (Boring, 1951; Winston, 2002), and that he was an imperfect chronicler (Thomas, 2009, 2016). I will describe selected publications to show how Boring’s work in AJP reflected his journey from structuralist student to disciplinary historian. In fact, the historian’s voice was apparent in AJP long before Boring published any of his landmark books on the origins of experimental psychol-
ology (Boring, 1929, 1942, 1950). With room for not only experimental reports but also for commentary, obituaries, and book reviews, *AJP* allowed him to address his audience from different directions, always in an effort to explain why history matters and how it can and should direct the emerging schools. Even his biographical pieces, retrospective by definition, end with an eye to the future and a reminder that legacies are never complete. One legacy, that of his mentor, preoccupied Boring and gave his historical writing—a background against which it could become clear (Boring, 1961, p. 22).

**An Engineer With the Psychological Point of View**

In 1905, E. G. Boring was an engineering student at Cornell when he elected to take Edward B. Titchener’s course in elementary psychology. He was enthralled by the Englishman who lectured in a master’s gown on topics ranging from tonal beats to the English language. Indeed, it was Boring’s use of language that distinguished him in the undergraduate classroom, and a written response to a test question—not a dazzling display of technical skill—drew Titchener’s praise; Wilhelm Wundt’s former student singled out the aspiring engineer and told him, “You have the psychological point of view!” (Boring, 1961, p. 19).

Boring graduated in 1908 and carried the memory of Titchener’s accolade back to his home state of Pennsylvania, where he tried his hand as an electrical engineer at the Bethlehem Steel Company and as a teacher at a Moravian parochial school. He disliked engineering because, contrary to his expectations, it was more about profit than discovery; he disliked teaching because, among other things, his students once glued him to a chair. Still dependent on his father’s support, he had become the family failure. All the while, “the magic of Titchener’s lectures...was still working” (p. 21). In 1910, the magic lured him back to Cornell, and with a $500 annual assistantship, he had the financial independence to pursue a PhD under Titchener.

**An Experimental Investigation in “The German Tradition”**

Working with Titchener meant publishing in *AJP*. Granville Stanley Hall founded the journal in 1887 to promote the experimental psychology of Johns Hopkins and other American universities, and when Titchener joined as a cooperating editor in 1895, he made it the “organ for himself and the Cornell Laboratory” (Boring, 1961, p. 42). The lab published work fashioned in what Boring called the “German tradition” (p. 22), dedicated to exploring the contents of consciousness with a large effort to avoid applied intent. Boring’s earliest manuscripts were not exceptions. In late 1911, Titchener instructed his student to develop a thesis on visceral sensibility, and the results were published in *AJP* as “The Sensations of the Alimentary Canal” (Boring, 1915). The manuscript is a prime example of introspective psychology, and Titchener’s student dutifully avoided the temptation to address the clinical significance of what may be structuralism’s most physiological investigation. Boring (1915) reviewed work from those who believed that the internal organs were, in fact, sensitive and others who believed that visceral sensations arose not from the organs but from the body wall, pleura, or peritoneum. He aimed to resolve this (mostly German) debate by administering stimuli to the esophagus, stomach, rectum, and colon while noting “the dependence of sensation upon intensity of stimulus” and obtaining “a description of the psychological character of the experiences by taking full introspections upon all occasions” (p. 5).

Although one short paragraph gets tantalizingly close to addressing the functional significance of pain and hunger (p. 4), Boring had no interest in Why?; this was a psychophysical study that only asked What? and Where?. Most of the trials required “observers” to swallow one end of a rubber tube so the experimenters could thermally, electrically, mechanically, and chemically stimulate the esophagus and stomach. For example, the tubes and associated apparatus allowed Boring to place an inflatable bladder or deliver electric shocks at precise locations along the length of the esophagus (it seems the engineer’s training was not entirely wasted).

As was the custom in Titchener’s lab, the observers were departmental associates and students, presumably trained in introspective techniques and all identified by name in the report. Boring’s collaborators withstood the procedures to varying degrees; one frequently vomited the tube. Another was unable to talk (and therefore unable to report) with the tube in place, and another simply could not tolerate the tube without gagging. Boring, on the other hand,
spent two years practicing his technique, and in his biography he boasted that he could swallow one end of a tube and answer the telephone without removing it (Boring, 1961, pp. 27–28). He was the primary observer and the only one to undergo “experimental work” on the colon and intestine (Boring, 1915, p. 49). A photograph of his bare torso shows the coordinate system upon which all localization reports were mapped.

The 57-page article is mostly a single-subject study, and in the structuralist tradition, Boring presumed to speak for humanity when he reported his introspections. He concluded that “the esophagus is sensitive to warm and cold throughout its length,” although localization is not as precise as others had reported, that “electrical stimulation of the stomach gives rise to sensations characteristic of electric shock,” which may or may not be due to current spread to adjacent tissues, and that “distention of the rectum produces the call to defecation” (pp. 56–57). The article ends with a summative list of similar observations. There are no remarks about how the esophagus tolerates temperatures that are similarly tolerated by the mouth, for example, and there is no speculation about why hydrochloric acid in the stomach produces hunger pains. There is no discussion of how these results might vary from person to person. There is no speculation about how illness might affect sensations, and there is no mention of how the study might be expanded or improved. In the structuralist tradition, stimuli were presented, thresholds were documented, and introspections were reported.

Boring completed his graduate work, passed his examination, and earned his degree with very little trouble and very little direct consultation with Titchener. Indeed, he must have had the psychologist’s—or at least Titchener’s—point of view. He would soon find his own.

Mental Testing and the Allure of Functionalism

Boring remained at Cornell until 1917, when he took his first tortured steps away from Titchener and introspection. This split haunted him for the rest of his life. In his biography, time and again, Boring reflected on how Titchener regarded or, after his passing, would have regarded each professional decision. Titchener was Boring’s great man, the one who had shaped history, and his link to psychology’s past, but he did not want to follow the mentor he so admired. Therefore, Boring’s writing was dedicated to reconciling past and present; American psychology, he argued, could find its way by defining itself in contrast to tradition generally and Titchener specifically. Boring elevated Titchener by describing him as American psychology’s perfect opponent: “No intellectual movement can get moving without something to push against, and American psychology had Titchener” (Boring, 1961, p. 22). Boring’s passion for the history of psychology was born of the need to know “why American psychology, while attempting to copy German introspective psychology in the late nineteenth century, nevertheless went functional” (p. 69). The fault was partly his own.

Titchener was meeting with his exclusive group of experimentalists at Harvard on April 6, 1917, when the United States declared war. He had established the society in 1904 and annually met with these psychologists who shared, above all, his social proprieties. Although they were an eclectic group, functionalist pursuits such as mental testing were typically off limits. However, when the group’s attention turned to how American psychologists could contribute to the war effort, the Englishman tactfully recused himself, and Robert Yerkes—president of the American Psychological Association, comparative psychologist, and friend of behaviorist John Watson—took the chair. Yerkes invited Captain W. S. Bowen, Harvard instructor of military science, to advise the group on the army’s problem of eliminating the “feebleminded” from recruitment pools. “It was this small stone that began the avalanche of intelligence testing in the United States Army” (Boring, 1938, p. 415). After the war, Titchener continued his annual meetings, but in 1917 a large portion of American psychology had stormed away from him and taken E. G. Boring with it.

Boring followed Yerkes into military service and set to work testing recruits, scoring tests, and generating reports. Much to his own surprise, the structuralist graduate student found fulfillment as a practicing functionalist. He “saw clearly that good, honest, intelligent work in any field merits respect and that the testers closely resemble the pure experimentalists in habits of work, in enthusiasm, and in thoroughness” (Boring, 1961, p. 31). Boring had served for less than a year when the war ended, but before returning to
civillian life, he went to Washington to help Yerkes prepare and present the test results.

In 1919, Boring accepted G. Stanley Hall’s invitation to join the faculty at Clark University and fill the vacancy left by the untimely death of Titchener’s former student John Wallace Baird. Although he finally had his own lab, Boring postponed experiments to draft a significant historical manuscript. Through his military service, Boring saw utility in mental testing, but while studying the history of probability theory at the Library of Congress, he came to appreciate its limitations. In particular, he was troubled by the zeal with which the testers overapplied the normal law of error, and he aimed to warn them (Boring, 1920). He did not discourage mental testing, as Titchener might have, but instead dismantled the arguments of past researchers who had given the normal distribution its inappropriately exalted status. The article does not scold or dispute; it explains and warns. The mental testing “avalanche” was about to spread far beyond the military, and like a parent offering last-minute advice to departing offspring, Boring acknowledged the promise of a new frontier while he explained the dangers and the wisdom of the elders. He aimed to offer the testers “discriminating encouragement” (Boring, 1920, p. 1).

Directing Methodology With a Historical Perspective

Hall shared Titchener’s appreciation for history, and by 1920 the AJP had an established record of publishing historical pieces such as Freud’s (1910) landmark “The Origin and Development of Psychoanalysis.” The context Boring (1920) provided in “The Logic of the Normal Law of Error in Mental Measurement” is an early example of how well he learned the historian’s craft at Cornell and, more importantly, how well he could communicate to a broad audience. He did not write like Titchener, and comparing the two, S. S. Stevens noted, “The master wrote for himself, whereas the pupil wrote for the reader” (Stevens, 1968, p. 591). The normal law paper (Boring, 1920) invited a wide audience from “laboratories, school-systems, factories, or the army” (p. 1), and Boring wrote with the style and patience to reach them.

Boring believed that the army’s mental tests had served their purpose. However, he was troubled by how the Gaussian law of normal distribution was assumed to apply to almost every scale imaginable, including measures of human intelligence. The functionalists were blind to the limitations of their craft, and Boring relished the opportunity to explain why. Whereas Titchener objected to what the functionalists wanted to do, Boring objected to how they were doing it. According to Boring, this was the essence of effective criticism; it must be internal and not external. No reader would be impressed by a structuralist who criticized a functionalist for behaving like a functionalist. He told Stevens (1968), “You can criticize only by showing that the author has failed to do what he himself has set out to do . . . I have got to show that my method will do [his] self-appointed job . . . better than his own” (p. 600).

Boring began his internal criticism by tracing probability theory to its origins in a practice that most readers could appreciate: games of chance. In the case of a coin flip, it is not unreasonable to argue from probability to frequency; that is, if a coin is a uniform, homogeneous circular disc, one should expect a roughly equal number of heads and tails from a large number of tosses. With no reason to expect a bias in favor of one outcome or the other, most investigators would consider each one equally probable. This reasoning may work perfectly well for a gambler who understands the few physical constraints of a coin flip, but it will fail almost anyone else. In most cases, to argue that two outcomes are equally probable because one has no reason to assume otherwise is to argue from ignorance, to derive probability from thin air.

In his article, Boring turned to Johannes von Kries (1886), a German physiological psychologist and student of Hermann von Helmholtz, to illustrate the dangers of what von Kries called the principle of insufficient reason. Boring explained how von Kries invited his readers to consider whether iron is in the distant star Sirius. With no evidence in favor or against, the principle asserts that the probability that Sirius has iron must be one in two (i.e., either it does or it does not). Von Kries expected his reader to understand that ignorance is complete uncertainty and that with complete uncertainty there can be no probability. Boring quoted von Kries (1886) in German: “Die Aufstellung der gleich möglichen Fälle muss eine in zwingener Weise und ohne jede Willkür sich ergebende sein” (p. 11). Translated, von Kries stated that the establishment of equal possible outcomes
must be compelling and not arbitrary. At the time, it was common for American psychologists to assume that their audience could read at least some German, but Boring took the opportunity to emphasize the cultural origins of this wisdom, offering no translation in an era of vicious anti-German sentiment. He added, “There is no alchemy of probabilities that will change ignorance into knowledge” (Boring, 1920, p. 3). This is not to say that probabilities should never be estimated, but they should be informed and perhaps deferred until science provides more information. Boring chose his example well: Concurrent with the publication of von Kries’s book, English astronomer William Huggins was using spectroscopy to demystify the composition of stars. By 1920, a curious scholar knew that Sirius probably had a tiny bit of iron (i.e., a little science revealed that the odds were much greater than 50/50) and almost certainly had hydrogen and helium. Boring’s contemporary audience could see the absurdity in estimating probability from ignorance and the benefit of waiting for data to inform estimates.

Boring then accused the mental testers of estimating probability from ignorance through the overapplication of Pierre-Simon Laplace’s (1812) normal law of error. In particular, Boring blamed Belgian polymath Adolphe Quetelet (1849) for imposing the law on the measurement of human traits generally and then on Francis Galton (1869) for applying it to human intelligence specifically. Quetelet saw the law as naturally preordained:

"Everything occurs then as though there existed a type of man from which all other men differed more or less. . . . This mean varies among different people, and sometimes even within the limits of a single country, where two people of different origins may be mixed together."

According to Quetelet, the measurements were destined to follow the law, and when they did not, he was able to make them conform. For example, he noted that if a distribution appears skewed, it is not because the law has been violated but because there must be multiple normal distributions, with unequal means, within the larger distribution. He went on to review data from 100,000 French conscripts and noted that 28,620 had been rejected for failing to meet the minimum height requirement of 5’2”. His calculations, based on the assumption of a perfectly normal distribution, indicated that only 26,345 should have been rejected. Rather than accept that the men had been appropriately dismissed from a skewed distribution, Quetelet concluded that 2,275 had been victims of fraud (p. 98). Instead of allowing observation to describe the distribution, he demanded that the data bow to theory. Quetelet applied the normal distribution with insufficient reason, and others followed.

Francis Galton (1869) embraced Quetelet’s theory, and it is a cornerstone of his landmark text Hereditary Genius. In the introduction Galton declared, “The range of mental power between . . . the greatest and least of English intellects, is enormous. . . . The method I shall employ for discovering all this, is an application of the very curious law of ‘deviation from an average’” (p. 26). In all his mental test measurements, he claimed, “deviations from the average follow theoretical computations with remarkable accuracy” (p. 30).

Boring noted that it took Galton only 10 years to disavow his faith in the normal distribution (Galton, 1879), and he was probably pleased to report that Galton’s revelation came by way of German psychophysics and Fechner’s law. In its simplest form, the law states that sensation = log(stimulus intensity). An observer tasked with differentiating a standard stimulus and comparison stimuli that are either more or less intense than the standard will make larger errors in excess and errors in deficiency will not be equal in magnitude. Most psychologists of the time understood the law, and if perception, the gateway to the mind, did not conform to the normal law, it seemed unlikely that other mental functions would. Boring stated the obvious: A law derived from a finite set of examples in the physical world cannot be universally applied to the mental. He concluded his description of Galton’s revelation with exasperation: “If the earlier writers . . . had endeavored to first classify experience in deviations from the average . . . that curve would never have obtained its present position in the theory of errors” (Boring, 1920, p. 14). They had ignored data in favor of theory.

Having told the story about how the French (Quetelet) and English (Galton) became inappropriately enchanted by the normal distribution, Boring (1920) took another opportunity to salute the cau-
tious Germans, who were “not so much interested in the applicability of the normal law as [they were] in the facts. The appeal to facts is . . . always a protest against theory which is given a priori” (p. 17).

The normal law of error had been overapplied and Galton himself admitted it. Boring’s final task was to explain what should be obvious: There is no justification for assuming the normal law will apply to the (human-made) units of mental testing, and he cited a budding line of research that, in fact, shows that it does not (Trabue, 1916; Williams, Titchener, & Boring, 1918).

Even if a mental measurement does conform to the normal law, Boring (1920) argued that we may not be fortunate enough to guess and use the appropriate unit. He provided a simple example by considering the size distribution of common salt crystals and asking whether size should be measured by height or by weight. Because weight should be proportional to the cube of height, the normal law cannot possibly apply to both scales. Nature will not guarantee us a normal distribution in whatever we choose to measure, and even if a normal distribution is to be found, it may not apply to our chosen scale.

In closing the article, Boring (1920) did not deny the utility of the normal law of error but outlined criteria by which it should be cautiously applied. Like Titchener, who believed that psychology needed a better understanding of consciousness before venturing into applied fields, Boring believed that the functionalists needed more data and better mental scales before using the normal law to estimate probability. He did not dismiss mental testing and concluded with words of encouragement:

There is nothing new in the contention that mental measurement is impossible, whereas now we do gain the assurance that rank-orders at least are validly demonstrable. And there is a great deal that can be done with rank-orders. . . . The serial constraints that do not presuppose a unit, yield less intricate resultants, but they present a rougher picture that represents truly the rough material which they describe. (pp. 32–33)

Although he never named his mentor in the normal law paper, Boring (1920) knew that _AJP_ was Titchener’s platform, and he had used it to, at least partly, endorse functionalism. The young professor had asserted himself but not without directly praising the wisdom of the German tradition and, perhaps, excusing introspection by indirectly explaining why its largely descriptive mental “measurements” may not yield to mathematical models. Boring offered Titchener all the respect he could afford.

The article cemented Boring’s reputation, and he believed that it impressed philosopher R. B. Perry enough to earn him an invitation to Harvard in 1922 when Clark’s political climate soured. Forty years after the article’s publication, however, Boring lamented that it “had never been effective, although the changing stream of scientific opinion (had) continued to flow in . . . the right direction” (Boring, 1961, p. 32). Reflecting on the 1920s, Boring stated that he “contributed no important research” (p. 51) but “found that (he) could get (his) crucial problems worked out at other universities if (he) would discuss the problems in publication” (p. 46). This statement belies Boring’s claim that his normal law paper had been ineffective. It secured his position—at Harvard and in the psychological community—and gave him the confidence to leave the lab work to his graduate students and focus on the “crucial problems.” He knew he had an audience. He had become “mixed in with the stream of American psychology, [he had become] its agent” (p. 51).

_Losing the Language_

By the 1920s, the mental testers, Gestalt psychologists, and behaviorists were running at their own pace and in their own directions, with no regard to pedigree. The trend was not new. Decades before, when James Mark Baldwin (1895) wrote his scathing critique of introspective methodology, Titchener (1896) responded by addressing the Princeton professor’s insolence when he replied, “I cannot think that his attitude to a long line of predecessors in the field is either scientifically or ethically defensible” (p. 241). Decorum and deference mattered to Titchener but not to the Americans who were disillusioned by introspection’s restrictive methodology. Titchener was destined to be disappointed, and from the moment he joined the staff in 1895, had no reason to expect _AJP_’s readers to blindly revere psychology’s intellectual ancestry. An editorial that bears his name declared that _AJP_ was a decidedly “American” journal.
and that it would not honor “discipleship to past or present leaders, or excessive deference to European thinkers” (Hall, Sandford, & Titchener, 1895, p. 4).

Nothing compromised American deference to European thinkers like the Great War. The impact on the German economy and universities was incalculable, and through the 1920s, an American psychologist would be hard pressed to keep up with the work coming from proliferating domestic laboratories, much less European ones. In a brief communication titled “Do American Psychologists Read European Psychology?” published in the AJP “Notes and Discussions” section, Boring (1928) presented results from a survey of 114 American libraries (31 public and 83 university) to show that, among foreign serials, German-language journals had lost favor to British ones, even though “German psychology (had) long led British psychology” (p. 674). The article’s title is misleading. Boring was not surprised to find that an American psychologist has access to the English journal Mind, “which he will seldom want” (p. 675), nor was he surprised to see that the libraries did not favor French- or Italian-language journals, but the unpopularity of the “new” German Gestalt journal, Psychologische Forschung, was unforgivable. Today, a reader might assume that Boring was simply bemoaning American provincialism. He was frustrated, but his historical perspective was conspicuously absent. Boring did not mention the anti-German fervor that swept the United States in the postwar years or the fact that some American professors were concerned that a German victory could come by way of scientific, if not political, domination. George A. Miller, a mathematician at the University of Illinois, called for American universities to commit themselves to their own research and break ties with German science and the German language:

Our students should not have to feel that the great majority of the best expository works relating to their subject are to be found only in the language of a people of low ideals imbued with a morbid desire to dominate the world at any cost. (Miller, 1918, p. 177)

Before the Supreme Court ruled such laws unconstitutional in 1923, 22 states had banned the public use and teaching of German (Gordin, 2015, pp. 181–182). American interest in German psychology was not just declining; it had been actively opposed and, for a time, legally forbidden. Boring’s article is therefore a brief, cautious statement of facts. A decade after the Treaty of Versailles, Boring was gently asking his English-speaking audience to consider the consequences of losing access to German science. He closed with a note of resignation: “Perhaps there is nothing to do about it, unless the fault and remedy lie in our graduate schools” (Boring, 1928, p. 675). American interest in the German language never recovered (Gordin, 2015, pp. 181–182), and the next generation of American psychologists did not have the ability or the inclination to read Wundt in his own words. Boring’s disappointment was probably compounded by the fact that his mentor, his link to the German tradition, had died the year before.

Losing “The Great Man”

“Seldom did (Titchener) distinguish between his wisdom and his convictions, and he never hid either” (Boring, 1961, p. 23). From the beginning, most could have guessed that Titchener and AJP would part ways. When his former student Karl Dallenbach purchased the journal from Hall in 1920, Titchener assumed that he had a philosophical and professional ally, and for a while he did. After all, like Boring, Dallenbach joined the military’s mental testers, but unlike Boring, he returned to Titchener and Cornell. However, Titchener resigned from AJP abruptly in 1925 amid circumstances that may have been as much about personalities as philosophies. Titchener assumed he had full editorial control, but, unknown to most, Dallenbach had taken a significant financial risk to purchase AJP, and “when fiscal wisdom came into conflict with [Titchener’s] editorial policy, fiscal realism won” (Boring, 1961, p. 42). Nonetheless, Titchener’s departure left a tremendous editorial void that Dallenbach scrambled to fill with Titchener’s former students Madison Bentley, Margaret Floy Washburn, and Boring. Titchener responded by founding a competing journal and turning his back on AJP, “predict[ing] shipwreck for a vessel with four rudders” (Boring, 1958, p. 15). Titchener died 2 years later.

Dallenbach asked Boring (1927) to prepare Titchener’s obituary, the first of many that he wrote for AJP. Both men must have understood that, to some degree, their reputations developed at Titchener’s
expense, and in 1927 the protest against structural psychology was still fresh and *AJP*’s future was still uncertain. Titchener’s obituary is an appropriate record of life events and accomplishments that concludes with a note of uncertainty, stating that “the evaluation of Titchener’s psychology can be left to posterity. . . . A century hence it will be possible to say just where his psychology belongs in the history of science” (Boring, 1927, pp. 505–506).

Boring withheld judgment, and at the time he was equally uncertain about psychology’s future. If one were to remove the references to Titchener from the previous quote, it would capture the sentiment that closes the first edition of *A History of Experimental Psychology* (Boring, 1929). It would read like this: “The evaluation of . . . psychology can be left to posterity. . . . A century hence it will be possible to say just where . . . psychology belongs in the history of science.” Two years after Titchener’s death, Boring was disappointed with psychology for not having produced a great man to unify the discipline in the way, for example, Charles Darwin had unified biology. He declared, “Psychology has never had a great man to itself. . . . There are signs that psychologists are ready for a great man or a great event . . . but the great event has not yet occurred” (Boring, 1929, p. 660). Perhaps Titchener was supposed to be psychology’s great man, but he had inspired division rather than unification. In 1927, Boring could not place Titchener in the history of psychology, and in 1929 he could not place psychology in the history of science. He was disappointed with the man and the discipline.

By 1950, he was more optimistic. When Boring (1950) published the second edition of *A History of Experimental Psychology*, he had the benefit of hindsight and had retreated from the great man theory of history. Psychology had matured and established itself as a science. Gestalt psychology, behaviorism, and the study of motivation had all borne fruit, and the proliferation of departments, journals, and textbooks stood as testimony to the discipline’s strength. Boring’s book increased by a third from the first edition to the second. Most importantly, psychology had not needed, and would not need, a single, galvanizing leader. It had progressed by integrating the accomplishments of many people, past and present; it had “matured, not like a person who never picks up new ancestors but like a family which, when a scion marries, acquires suddenly all the ancestors of the new spouse” (Boring, 1950, p. xiii). Two decades had given Boring the opportunity to reconsider psychology and Titchener; he did not need to be the great man, but the obituary had already been written. “Can history be revised? Yes” (Boring, 1950, p. xiii).

**Revising History Through Biography**

The AJP thrived under Dallenbach’s direction, and he and Boring cemented a partnership and friendship that lasted for decades. The bond was so strong that each agreed to write the other’s necrology. Instead, Boring got the opportunity to write Dallenbach’s *vita* (Boring, 1958) on the 50th anniversary of his career. In it, he painted a picture of another student who, like him, was lured into psychology by Titchener, but he had some explaining to do.

Nobody knew the Dallenbach–Titchener relationship better than Boring, and Dallenbach’s biography gave him an unusual opportunity to address two legacies and to reconcile the two men he admired most. He began with portraits of character. The 40-page biography of Dallenbach names Titchener more than one hundred times and includes a page-width photograph of the Englishman seated next to a table piled high with papers. To his former students, Titchener was an intellectual giant and the praise that Boring offers Dallenbach is measured first in how well he matched Titchener’s personal attributes:

No one . . . re-presented Titchener to the psychological world more nearly than has Dallenbach: . . . his basic social values, his sense of propriety, his initial formalism . . . his dominance and also its complement, his kindness, his intense interest in the personal welfare of his students; his capacity for indefatigable meticulous hard work . . . he still carries with him many of these Titchenerian values after all these years. (Boring, 1958, p. 11)

“Titchener,” the mortal noun, had become “Titchenerian,” the immortal adjective, manifested in Dallenbach. But if these men were so similar, and if Dallenbach so admired Titchener, what could explain the split? It is with an odd sense of relief that Boring describes Titchener’s deadly affliction; a
brain tumor gave the former students a resolution to the 1925 crisis that most other illnesses could not:

In 1927, two years after [he left] the JOURNAL, Titchener died of a brain tumor, which provided ample explanation for his capricious forgetfulness and his inconstant judgment in the year or two immediately preceding. . . . As soon as (Dallenbach) knew of the brain lesion, his loyalty to Titchener’s memory and ideals surged back and it has never diminished, though changing times have altered the form of its expression. (p. 15)

It is not clear that Boring’s loyalty surged back as quickly as Dallenbach’s supposedly did. In the original obituary, Boring (1927) did not address the AJP controversy or Titchener’s cause of death, although he surely knew about the former and probably knew about the latter. Titchener was Boring’s link to psychology’s experimental origins, but by the time he died, the sage had become isolated. Of course, the isolation was partly self-imposed. Titchener had left the American Psychological Association decades earlier and, after 1911, he seldom traveled except to meet with the Experimentalists (Boring, 1927). From 1912 to 1917, he toiled over but ultimately abandoned a systematic text; Boring supposed that he was unable to nurture it beyond a state of incubation (p. 501). The uncertainty was reflected in the Cornell lab, which had an ever-changing language; “a graduate three years absent had on returning to learn psychology again” (Boring, 1927, p. 500). The Titchener that Boring and Dallenbach knew as students had faded through the 1920s and certainly was not there when he stormed away from AJP in 1925. When Boring wrote the obituary in 1927, Titchener’s legacy was incomplete and uncertain. After all, to write a conclusive obituary within a year, or even within a decade, of the subject’s death would have violated Boring’s own rule for historical writing: “I speak with confidence up to twenty years ago; I speak, but with less assurance, of the next decade; whatever I say for the most recent decade is based on gratuitous courage” (Boring, 1950, xv). Thirty years allowed Boring to speak, confidently and finally, about Titchener’s science.

“On the record Titchener died in 1927, yet he was still enough alive for Dallenbach to take him along when he left Cornell in 1948” (Boring, 1958, p. 11). Titchener was not just a man; he was an idea that could manifest itself in his disciples. However, Dallenbach’s work, although always linked to sensation and perception, seldom adhered to the structuralist approach of introspecting the generalized human mind; he studied the perceptual skills of those with visual and auditory impairments, designed experiments that did not demand trained observers, and conducted groundbreaking research on sleep and memory (Jenkins & Dallenbach, 1924). Boring needed to be thorough if he expected his readers to see Titchener in Dallenbach and to understand that Titchener’s psychology could be found in the methods, if not the hypotheses. He started with the training.

Dallenbach’s psychological journey began when, as an undergraduate at the University of Illinois, he drew the attention of Titchener’s former student John Wallace Baird. When Titchener visited, Baird showed him Dallenbach’s meticulous notebooks, and although the undergraduate had intended to pursue law, “Baird and Titchener had now become Karl’s two idols in psychology, there was no place for him to go but to Titchener at Cornell” (Boring, 1958, p. 9). He completed a PhD in 1913.

Having defended his thesis, Dallenbach intended to obtain an MD at the University of Pittsburgh, but again, Titchener redirected him, and he took a faculty position at the University of Oregon. “It is hard to explain this ‘magnetic’ power that Titchener had over (his) disciples” (p. 12). Dallenbach enjoyed the University of Oregon, but it “had no money for ‘personal research,’ and that, it turned out, meant self-initiated research, research that the investigator had thought up on his own and wanted to do” (p. 13). It seemed the University of Oregon wanted only practical research, but Titchenerian Dallenbach was no functionalist. Titchener brought him back to Cornell in 1916 (p. 13).

Boring (1958, 1961) portrayed Dallenbach as Titchener’s loyal and like-minded disciple up until the AJP crisis of 1925. Titchener, “always self-confident, regarded himself as ruler through natural right even though one of his subjects had provided the cash” (Boring, 1961, p. 42). The arrangement was doomed to fail, but Dallenbach’s loyalty somehow sustained it for 5 years. Eventually, fiscal pragmatism
trumped editorial dogma, so Titchener walked away from Dallenbach, leaving only his science behind.

To Boring, Titchener was gone long before 1927, but his pre-illness ideals had already been transferred to Dallenbach. Nonetheless, the student needed to construct his own environment and shake the last remnants of the material Titchener. He needed to leave Cornell, where, despite being Sage Professor of Psychology, “Dallenbach could not see himself as having attained the goal that the Titchener-image demanded” (Boring, 1958, p. 17). An offer to chair the psychology department at the University of Texas allowed him to establish Titchenerian psychology anew.

When Dallenbach arrived in Texas in 1948, he found a university flush with oil money and ready to expand its psychology department. He found “clinical psychology predominant [at the University of Texas],” but “he was determined that experimental psychology should be the core of the department” (Boring, 1958, p. 18). When Boring described the layout of Dallenbach’s lab, he presented it as an expansion of the Cornell lab:

Dallenbach interpreted experimental psychology in Titchenerian terms, with more stress on sensation and perception than on motivation and personality. . . . The special rooms on the ground floor of the new laboratory clearly reflected the influence of Titchener, who had divided the Cornell Laboratory by sense-departments. (Boring, 1958, p. 19)

However, Dallenbach expanded outward from his Titchenerian foundation, but Boring justified each transgression:

There was . . . one un-Titchenerian feature (in the new lab), a suite of rooms for comparative psychology, but the father image would scarcely even have scowled, for mammals were housed in another building and the apes eight miles away. (p. 19)

Boring continued with more un-Titchenerian features:

Titchener held that experimental psychology deals with the human, adult, normal, generalized mind, and this view affected Dallenbach’s conception of the proper laboratory policy and also his policy for The American Journal of Psychology. There are exceptions. The rules are not rigid. Dallenbach used cockroaches as subjects in one of his best experiments. More than once he studied the capacities of children, and by no means did he confine his research to introspection, though his predilection for the investigation of sensation kept him well within the Titchenerian framework. (p. 21)

But what was the “Titchenerian framework”? Titchener resented labels, and although others called him “structural” and “introspectional,” Boring “never heard him refer to his school by any other word than ‘we’” (Boring, 1927, p. 497). In Dallenbach’s biography, Boring deemphasized the word introspection in favor of attention. After all, “Titchener sought to settle the problem of attention by declaring its status as an attribute of sensation” (Boring, 1927, p. 499), and Dallenbach continued this work while, for the most part, shedding Titchenerian terms such as quality, intensity, duration, extent, and clearness.

Boring saw no contradiction. “You could, of course, perform experiments that dealt with the conditions of attention and the effectiveness of intended distractors upon attention without yourself taking up a position on whether sensations really do possess an attribute of clearness” (Boring, 1958, p. 24). The study of attention may not have been a functionalist pursuit, but in 1958, it certainly was not anachronistic, and it certainly fit, at least as Boring saw it, the Titchenerian model. Dallenbach was to be commended for remaining true to pure, objective science in the age of applied research.

When Dallenbach, the Titchener image before him, stuck to sensory psychology, he was swimming against the current of the times. It was a handicap, but he never faltered in the face of difficulty. He stuck to his principles. (p. 40)

In Dallenbach’s biography, Boring used AJP’s format to reconsider and artfully shape Titchener’s legacy by showing how it was alive and well in Dallenbach. More importantly, the article showcases the historian at the peak of his career. When he wrote Dallenbach’s vita in 1958, Boring was on his way to becoming the “unofficial biographer of psychology’s
great”; having just completed memoirs for Robert Yerkes and Lewis Terman, he did one more for Karl Lashley before refusing additional requests (Boring, 1961, pp. 77–78). In his own way, Boring had emulated Titchener, not as an investigator but as a “historian \textit{par excellence}” (Boring, 1927, p. x).

In closing the preface to the first edition of his history book, Boring apologetically stated, “[Titchener] should have written this book, and it is with great diffidence that I offer a poor substitute” (Boring, 1929, p. x). The second edition (Boring, 1950) is also dedicated to Titchener but without apology. By the time he wrote Dallenbach’s biography, Boring was a seasoned craftsman. He closed the article by describing Dallenbach’s greatest achievement, the steady 37-year stewardship of \textit{AJP}, and in lauding this tremendous feat, he lamented that “the pity is that it does not show up clearly on history’s pages” (Boring, 1958, p. 36). The words are their own clever contradiction.

\textbf{But anyone can write History. Can he not do Research?}

That was Boring’s “pet paranoia.” He feared that those who admired his historical work quietly assumed he was incompetent in the lab (Boring, 1961, p. 15). Indeed, he was content to delegate and let his students do the work, and receive the credit, but Boring was a tireless supervisor and a ruthless editor. Any literature-based estimation of Boring’s scientific output would certainly fall short because, although he contributed to student work, he “would put his name on a paper only if he was the major contributor” (Stevens, 1968, p. 600). This stubborn position led to heated arguments between Boring and Stevens (p. 601), but it also placed a series of experimental reports, published in \textit{AJP} between January 1940 and January 1941, in sharp contrast against the rest of his writing. These four articles, coauthored with lab assistant Alfred H. Holway, present Boring as an active experimentalist, focused on a single phenomenon and surprisingly willing to claim authorship (Holway & Boring, 1940a, 1940b, 1940c, 1941). At the midpoint of his career, the established historian designed and executed a number of elegant experiments that remain staples of undergraduate textbooks and laboratory exercises (Gallagher & Hoefling, 2013; Goldstein, 2014). A simple psychophysical function, Emmert’s Law, was at the heart of this work, and in “Size Constancy and Emmert’s Law,” published in \textit{AJP}’s “Notes and Discussions” section, Boring (1940) provided a concise foundation for the four experimental articles.

To Boring, size constancy and Emmert’s Law were two sides of the same coin. Perceptual psychologists understood that an object’s apparent size remains constant when viewed from different distances, regardless of how much the image’s angular size changes. This is presumably because the observer somehow processes the inverse relationship between angular size and distance. Emmert’s Law explained, among other things, why a visual afterimage appears larger on a distant surface than it does on a close one. That is, an image with a stable angular size will vary in apparent size depending on the perceived distance. True to his psychophysicist roots, Boring presented Emmert’s Law as an equation that linked perception and the physical environment: $s = krd$, where $s =$ apparent size of the stimulus, $k =$ the constant of proportionality, $r =$ size of the retinal image (extrapolated from the image’s angular size), and $d =$ distance between observer and stimulus. The equation was so simple, and the link to size constancy so clear, that experiment seemed almost redundant. Why, then, would this particular problem motivate Boring to experimentation and authorship when, in most cases, he “could get [his] crucial problems worked out at other universities [by discussing them] in publication”? (Boring, 1961, p. 46). The answer was simple: The Gestaltists were skeptical.

Kurt Koffka (1935) addressed size constancy in his classic \textit{Principles of Gestalt Psychology}, but he complained, “We have to the present day no complete knowledge of the quantitative relations” between perceived size and distance (p. 91). Although all of the Holway and Boring articles could be interpreted as responses to Koffka, one article, “Determinants of Apparent Visual Size with Distance Variant” (Holway & Boring, 1941) explicitly targeted the Gestaltists’ challenge.

To conduct their experiments, the authors and three associates worked through the night, blissfully “free of interruptions” (Boring, 1961, p. 58), to create circumstances in which observers could make size judgments in the absence of distance cues. With an observer seated at the intersection of two dark corridors (Figure 1), the experimenters projected one
luminous circle in one hallway, on an 8’ × 8’ screen, located between 10′ and 120′ from the observer; this was the standard stimulus that, in all trials and regardless of distance, subtended a visual angle of 1°. In the second hallway, another screen displayed a comparison circle, always 10′ from the observer. The observers then needed to adjust the size of the comparison circle until it matched the standard. They viewed the circles with both eyes, and, although in the dark, stray light still illuminated surfaces of the corridor, providing “a sensory ground for the perception of the stimulus” (p. 30). With the features of the hallways visible, and the benefit of binocular disparity, observers accurately assessed distances and, as Emmert’s Law predicted, easily matched the comparison stimulus to the true physical size of the standard stimulus. On subsequent trials, the experimenters removed environmental depth cues, and as they did, the size estimates were no longer good indicators of the standard stimulus’ physical size. First, they removed binocular cues by occluding one eye. Then, they diminished the remaining monocular cues by having the participants view the standard circle through an artificial pupil and, finally, through an extendable reduction tunnel (a tube, actually) that blocked the corridor surfaces. In this final condition, observers were unable to judge distance to the standard stimulus and consequently made poor estimates of physical size but accurately matched the circles according to angular size. In the absence of depth cues, Emmert’s Law held but was reduced to \( s = kr \). The authors were triumphant:

For all that has been said by Gestalt psychologists against the validity of [Emmert’s Law], it would nevertheless appear that, when no relevant datum other than retinal size is available, then the perception of size will after all vary solely with the visual angle. (Holway & Boring, 1941, p. 36)

Without depth cues, perceived size matched the relative size of the proximal stimulus. Titchener would have been pleased. The historian had returned to the lab to answer a challenge from the Gestalt school, but there was more. To Boring, the Titchenerian disciple, positivism was a hidden pillar of Gestalt psychology. In his 10-page review of Koffka’s (1935) text, Boring expressed frustration in the writer’s willingness to discuss immediate experience without naming it. “Has he not perhaps repressed immediate experience because he hates positivism, therefore admitting it only in disguise?” (Boring, 1936, p. 66). The apparent size experiments were conceived shortly after the publication of this passionate review; for Boring, an exploration of immediate experience, especially as it pertained to size constancy, proved irresistible.

Boring and Holway expanded their work by developing elaborate contraptions to study the “moon illusion,” the perception that the moon is large on the horizon and small in culmination (Boring & Holway, 1940; Holway & Boring, 1940c). They found that the illusion is reversed when the observer is supine and must therefore depend on the angle of regard. This conclusion was subsequently challenged (Kaufman & Rock, 1962) and then reaffirmed (Suzuki, 2007), but with the exception of a few summative reports, Boring had little more to add to the story. Although one article, “The Moon Illusion and the Angle of Regard” (Holway & Boring, 1940c), concludes with a promising description of how the authors planned to test the illusion while observers viewed the sun, Boring never published the findings or returned to the lab with so much determination. World war, once again, pulled him toward more practical matters, and he assumed more teaching responsibilities as Harvard’s junior faculty answered the call to service. Warfare favored applied research, and December 6, 1941, was “the last day when pure scholarship could be undertaken with a clear conscience” (Boring, 1961, p. 60).

**Figure 1.** Illustration from Holway and Boring (1941) showing the corridor plan. Observer, O, is located at the intersection of two corridors. SC indicates the position of the screen displaying the comparison stimulus located at a constant distance \( D_c = 10' \) from O. SS, at a distance \( D_s \) from O, indicates one of the positions occupied by the standard stimulus, which always subtended a visual angle of 1°. Distance from O to the standard was varied from 10′ to 120′. \( P_c \) and \( P_s \) indicate the positions of the projectors.
Conclusion
Throughout his career, Edwin Garrigues Boring struggled to redefine his commitment to a figure and a school of psychology that, in the eyes of many, had been abandoned. True to AJP’s editorial mandate of 1895 (Hall et al., 1895), Boring’s manuscripts avoid deference without being indifferent. My intent has been to show how “the historical approach to the understanding of scientific fact is what differentiates the scholar in science from the mere experimenter” (Boring, 1961, p. 3). Boring was no mere experimenter, and AJP’s various publication formats allowed him to display the “many polished facets [that] make [the] gem” (Stevens, 1968, p. 606). In 1917, the student was fascinated and pleased with his mastery of unusual apparatus and operating against a background of German physiology; in 1920, the recently commissioned methodologist understood the excitement of the testing movement but exposed its theoretical flaws; in 1927 and, again, in 1958, the biographer, struggling to understand a beloved mentor, wrote flaws; in 1927 and, again, in 1958, the biographer, struggling to understand Gestalt theory and then rewrote an obituary; in 1940, the positivist experimenter seemed to understand Gestalt theory but its practitioners and used data to challenge the school from within. Boring presented many facets, but each one reflected the historian, providing perspective and working to unify the discipline.

NOTES
Dedicated to David P. M. Northmore, who once asked, “Why don’t we try to replicate the Boring experiment?”

Thanks to Roger K. Thomas, Christine Savicky, and the anonymous reviewers for commenting on earlier drafts of this manuscript.

Address correspondence about this article to Shawn P. Gallagher, Department of Psychology, 111 Byerly Hall, Millersville University, P.O. Box 1002 (1 S. George St.), Millersville, PA 17551 (e-mail: Shawn.Gallagher@Millersville.edu).

REFERENCES


Many educational demonstrations of memory and recall employ word lists and number strings; items that lend themselves to semantic organization and “chunking.” By applying taste recall to the adaptive memory paradigm, which evaluates memory from a survival-based evolutionary perspective, we have developed a simple, inexpensive exercise that defies mnemonic strategies. Most adaptive memory studies have evaluated recall of words encountered while imagining survival and non-survival scenarios. Here, we’ve left the lexical domain and hypothesized that taste memory, as measured by recognition, would be best when acquisition occurs under imagined threat of personal harm, namely poisoning. We tested participants individually while they evaluated eight teas in one of three conditions: in one, they evaluated the toxicity of the tea (survival condition), in a second, they considered the marketability of the tea and, in the third, they evaluated the bitterness of the tea. After a filler task, a surprise recognition task required the participants to taste and identify the eight original teas from a group of 16 that included eight novel teas. The survival condition led to better recognition than the bitterness condition but, surprisingly, it did not yield better recognition than the marketing condition. A second experiment employed a streamlined design more appropriate for classroom settings and failed to support the hypothesis that planning enhanced recognition in survival scenarios. This simple technique has, at least, revealed a robust levels-of-processing effect for taste recognition and invites students to consider the adaptive advantages of all forms of memory.

Key words: taste; memory; recall; adaptive memory; levels-of-processing; survival processing; classroom demonstration

Memory and recall are common topics in general psychology courses as well as others that explore perception, learning, and cognition. Cited work and suggested demonstrations, however, often revolve around a few familiar paradigms like recalling word lists and number strings, recognizing faces, and reporting past events. A current and popular cognitive psychology textbook cites 781 references and while at least 300 of these sources describe memory research, none of the studies addresses taste memory (Goldstein, 2015). This is not surprising, considering the importance of remembering sights and sounds and the ease with which visual and auditory recall experiments can be designed, but students should be aware of the multimodal nature of memory and educators should explore ways to demonstrate recall across the senses. This report describes a novel method that applies taste memory to an established experimental framework.

Specific forms of human memory may have evolved to serve specific survival-based functions and one experimental paradigm attempts to take adaptive mechanisms into account (Nairne et al., 2007). These “adaptive memory” studies have shown a mnemonic advantage for survival processing that is superior to some well-established encoding conditions including mental imagery, self-reference, relational processing, and intentional learning (Nairne and Pandeirada, 2008; 2010; Nairne et al., 2008). Nairne et al. (2007), asked participants to imagine that they had been stranded on an unfamiliar grassland. They then viewed a list of words and provided a score for each one to indicate its importance for surviving in such an environment. Control conditions included one in which participants were asked to score the words while imagining a move to a new home in a foreign country, and another in which participants were simply asked to rate the pleasantness of each word. A surprise free-recall task given at the conclusion of the experiment showed superior word recall for the adaptive memory condition compared to the other two. Other investigators have replicated these findings both within- and between-subjects and with alternate scenarios meant to control for schematic processing (Kang et al., 2008; Weinstein et al., 2008). Some see these results as evidence that selective pressures have shaped memory and that the effects of this shaping are manifest in superior recall for items easily associated with survival (Nairne et al., 2009).

Several researchers have challenged or refined the survival processing explanation, igniting a rich debate (Butler et al., 2009; Otgaar and Smeets, 2010; Burns et al., 2011; Otgaar et al., 2010; Nairne and Pandeirada, 2011). Among the most intriguing are two studies suggesting that the survival advantage does not appear in implicit memory (Tse and Altarriba, 2010) or face recognition (Savine et al., 2011) and another that proposes a special form of iconic memory for threatening visual stimuli (Kuhbandner et al., 2011). These results underscore the importance of teaching memory as multi-modal and testing the possible survival advantages for different forms of memory across a wide variety of experimental situations.

Nairne et al. (2007) employed a grasslands scenario on the assumption that adaptive memory emerged during the Pleistocene era of human hunter-gatherers. Although
Describe memory not as a single function, but as aifty with a label,ely identify (e.g. experiment, f e asked imagine scenarios and rate tea flavors -ent, students should be -esse.

Compare and contrast the Describe how recall suitably for testing students en ma

in shallow processing.

After completing a distracter task, our participants took a

adaptive memory the

the

other objects or events (deep processing) than it will be for items that are encoded while the participant only considers the item’s features (shallow processing). We had one adaptive memory deep processing task, one non-adaptive deep processing task, and one shallow processing task. After completing a distracter task, our participants took a surprise taste recognition test in which we presented both previously encountered and novel stimuli. We hypothesized that, compared to those in the other two conditions, the participants in the imaginary poison content condition would be better at discriminating between recently encountered and novel teas. Additionally, recognition would be better for those who engaged in non-adaptive deep processing compared to those who engaged in shallow processing.

A second experiment employed a modified procedure suitable for testing students en masse. Using a similar but streamlined technique, we tested the hypothesis that a future-directed temporal orientation (planning) could enhance the survival processing effect (Klein et al., 2011).

PRIMARY LEARNING OBJECTIVES
Upon completion of this experiment, students should be able to:
1. Describe memory not as a single function, but as a collection of systems shaped by natural selection.
2. Describe the adaptive uses of the sensory systems and, with these in mind, propose experiments aimed at exploring the adaptive memory of each.
3. Describe how recall (and, presumably, encoding) can be difficult when the stimuli are not easily to label.
4. Compare and contrast the adaptive memory effect and levels-of-processing effect and consider whether or not they are, in fact, different effects or a common effect that can be exhibited by different techniques.

MATERIALS AND METHODS: EXPERIMENT 1

Participants
These experiments complied with the standards of Millersville University’s Institutional Review Board. In the first experiment, fifty-three undergraduates participated in 30-minute sessions in exchange for course credit. Participants were tested individually in a classroom environment by fellow undergraduates.

Materials
We obtained 32 different loose-leaf teas from a local market and brewed each one by steeping 3 tablespoons in 32 ounces of hot water for four minutes. We allowed them to cool, numbered them, and then stored them at room temperature (21°C), in sealable plastic containers for at least 12 hours before use.

Procedure
Prior to data collection, a group of eleven undergraduates volunteered to screen the 32 teas. We gave them 5 ml of each tea to taste and consume and asked each to describe or label them according to as many characteristics as they could subjectively identify (e.g., tastes minty, smells like lemon, tastes very bitter). We were concerned that tea recognition could be confounded by labeling; for example, one could report recognizing a particular tea flavor because they originally encoded the stimulus with a label, such as “lemony” or “unusually bitter,” and not because they recognized the actual stimulus. To minimize labeling effects, we chose the 16 teas that the volunteers described or labeled most inconsistently.

<table>
<thead>
<tr>
<th>Set A</th>
<th>Set B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chai*</td>
<td>Apricot</td>
</tr>
<tr>
<td>China Black*</td>
<td>Assam</td>
</tr>
<tr>
<td>Darjeeling</td>
<td>Ceylon Ultima</td>
</tr>
<tr>
<td>Earl Grey</td>
<td>English Breakfast</td>
</tr>
<tr>
<td>Lemon Spice*</td>
<td>Giulia*</td>
</tr>
<tr>
<td>Mango*</td>
<td>Imperial Green</td>
</tr>
<tr>
<td>Oolong</td>
<td>Turk Caravan*</td>
</tr>
<tr>
<td>Russian Ancai</td>
<td>Yorkshire Gold**</td>
</tr>
</tbody>
</table>

Table 1. Teas Used. Teas were made by White Coffee (Long Island City, NY) except *Metropolitan Tea Company (Cheektowaga, NY) and **Brands of Britain (San Ramon, CA).
We divided the selected teas into two sets of eight (Table 1) and presented each participant with one of the two sets (Set A, n = 29; Set B, n = 24) to taste and swallow for the encoding phase. After selecting a tea set, we randomly assigned participants to one of three encoding conditions and an experimenter read the corresponding scenario aloud. The three scenarios were as follows:

**Condition 1**
(Survival Condition, n = 18): Imagine that all tea leaves naturally contain a certain amount of poison and that, even though the poisonous properties of these teas have been neutralized during processing for human consumption, some taste of the original poison remains. You are at a prime age for tasting the poison present in the tea leaves, even though the tea is no longer toxic. Please rate each tea on a scale of 0 to 10 based on how much poison you think was in the tea leaves prior to processing, with 0 signifying no poison, 5 signifying a moderate amount of poison, and 10 signifying a large amount of poison.

**Condition 2**
(Marketing Condition, n = 16): Imagine that the university is instituting a new program which will have students grow teas in the university greenhouse and then sell them to the community in a new campus store. You have been chosen to rate a selection of teas based on how well you think each one will sell. Please rate each tea on a scale of 0 to 10 based on how well you think each tea will sell with 0 signifying that the tea will not sell at all, 5 signifying that the tea will sell moderately well, and 10 signifying that the tea will sell very well.

**Condition 3**
(Bitterness Condition, n = 19): Please rate each tea on a scale of 0 to 10 based on how bitter you think it is, with 0 signifying that the tea is not bitter at all, 5 signifying that the tea is moderately bitter, and 10signifying that the tea is extremely bitter.

We used an alternating strategy to assign participants to the sets and we counterbalanced presentation sequences within sets. We then blindfolded the participant and presented the eight teas in individual plastic cups containing 5 ml of liquid at room temperature (21°C). After tasting and rating a tea, the participant rinsed their mouth with 100 ml of water, spit the water out, and waited 30 seconds for the next tea. At the conclusion of the rating phase, we instructed the participant to remove the blindfold and complete a tea familiarity questionnaire which simply asked how often they drank tea (0, 1-3, 4-6, or more than 6 times/week) and how many different flavors they typically sampled each week.

**RESULTS: EXPERIMENT 1**
We used an alpha level of 0.05 for all statistical tests. Figure 1 shows the mean recognition score, measured as d' per MacMillan and Creelman (1991), for the three conditions. One participant's score was more than two standard deviations above the mean for that condition and was therefore excluded from the analyses. Participants in the bitterness condition performed at chance level. A single factor analysis of variance (ANOVA) revealed a significant effect for condition, \( F(2,49) = 7.15, \ MSE = 0.07, \ p = 0.002, \eta^2_p = 0.23 \). Scheffé tests for multiple comparisons showed that while recognition scores differed significantly between the survival condition and the bitterness condition, \( p = .011 \), as well as between the marketing condition and the bitterness condition, \( p = .006 \), there was no significant difference between the survival and marketing condition scores, \( p = 0.954 \).

We explored the possibility of a link between tea familiarity and recognition with a single factor ANOVA that revealed no significant differences among groups of participants who reported drinking tea 0 times/week, 1-3 times/week, and 4-6 times/week, \( F(2,48) = 0.73, \ MSE = 0.10, \ p = 0.487 \). (No participants reported consuming tea more than six times per week.) ANOVAs were also used to verify that the effect of condition seen across all participants existed within each of the two tea sets (A and B). Recognition (d') was compared for participants in the three conditions who tasted teas in Set A, \( F(2,26) = 3.73, \ MSE = 0.06, \ p = 0.038 \), as well as Set B, \( F(2,20) = 7.18, \ MSE = 0.06, \ p = 0.005 \). Post hoc analyses confirm the same effect of condition within both tea sets.

**MATERIALS AND METHODS: EXPERIMENT 2**
We modified the previous procedure to test the effect of planning on survival processing scenarios and streamline
data collection for classroom situations.

Participants
One investigator tested forty-one undergraduates in two large groups as part of a demonstration for a cognitive psychology class. The students had studied the sensory systems and attention but had yet to study memory.

Materials
We reserved a single classroom and prepared individual desks for testing up to 25 participants. We prepared teas as described in Experiment 1 and distributed 5 ml quantities in opaque plastic 30ml cups. Single sheets of paper described one of two experimental scenarios as well as a common distractor task that instructed the students to calculate the mean of twenty single-digit numbers. Each desk had eight cups of tea and one instruction sheet placed face-down. The four teas to be consumed for the rating phase were labelled A-D and visible on the desktops while the four cups to be used in the recognition phase, also labelled A-D, were behind the other four and draped with paper towels. We used a Latin square method to create 16 unique rating sets. Two teas in each corresponding recognition set were randomly selected from the four in the rating set while the other two were selected from the remaining teas.

Procedure
Having prepared each desk with one of the 16 possible tea sets, the investigator led the students into the classroom, gave them pencils, and told them to read their individual instruction sheets which included one of the following scenarios:

Condition 1
(Survival without Planning Condition, \( n = 19 \)): Imagine that you are lost in the wilderness, far from home and without any water. You have come across an abandoned campsite and find four heavy jars that appear to contain four different teas. You are not sure if the teas are safe to drink, but you are very thirsty now and need to drink something in order to survive. You decide to sample each tea and then immediately drink from the jar of the one that you believe is safest. Please taste and rate each tea on a scale of 0 to 10 based on how safe you think each tea is with 0 signifying unsafe, 5 signifying moderately safe, and 10 signifying that the tea is very safe.

Participants drank the teas at their own pace, provided scores on the instruction sheet on spaces labelled A-D, and completed the distractor task. The investigator then distributed a second sheet of paper that instructed them to “Gently remove the cover from the four remaining cups, try each one and indicate if this tea was in the first group of four or if it is a new tea. Circle the appropriate answer.” A list of teas, A-D, followed along with the text: “This tea was in the first group of four. / This is a new tea.” A third sheet asked a final question; “What strategies (if any) did you use to remember the teas?” The responses remained at the corresponding desks when the students were dismissed.

RESULTS: EXPERIMENT 2
Whereas Experiment 1 required participants to make sixteen binary decisions, Experiment 2 required them to make only four. Fourteen participants (34%) performed perfectly and the resultant ceiling effect, combined with the fact that four binary choices yield only five possible recognition score (\( d' \)) values, precluded data analysis with parametric tests. Because of the low sample sizes, we collapsed the data into below chance (\( d' < 0 \)), chance (\( d' = 0 \)) and above chance (\( d' > 0 \)) scores (Figure 2). A 2x3 Freeman-Halton Fisher exact test did not show a significant difference between the two conditions, \( p = 0.282 \).

When asked to describe the strategies used to recall the teas, 32 participants stated that they relied on taste and/or smell, but only five mentioned the use of descriptive labels like bitter, and watered-down. None stated that they recognized a specific tea. We saw no evidence of consistent labelling and nine participants reported having no strategy at all. Six of these nine performed at above chance and two of them performed perfectly despite having no insight to their ability.

Figure 2. Frequency distributions of recognition scores for the participants in Experiment 2.

DISCUSSION
Student investigators and participants found these experiments simple and tolerable, if not enjoyable. Although most participants had performed some form of word recall memory task as part of a classroom demonstration in previous high school or college courses,
none had previously performed a memory task involving taste. The process, of course, presents opportunities to discuss methodological and analytical issues but the results should lend themselves to rich classroom discussions about the multi-modal nature of memory, levels-of-processing theory, and the challenge of recalling (and, presumably, encoding) stimuli that are not easily labelled.

Unfortunately, we still do not know if and how survival processing scenarios apply to taste recall. For example, in Experiment 1, contrary to previous findings, our adaptive memory (Survival) condition failed to produce superior recognition compared to a non-adaptive marketing condition. However, the poison and marketing conditions elicited much better recognition than the bitterness condition.

One interpretation of the Experiment 1 results, and perhaps the most parsimonious one, is that we have demonstrated a standard levels-of-processing effect (Craik and Lockhart, 1972; Challis et al., 1996). The survival and marketing conditions required imaginative elaboration (deep processing) whereas the bitterness condition required only sensory analysis (shallow processing). This interpretation suggests that while our data show no obvious adaptive memory effect, the levels-of-processing effect is robust.

Another possible explanation for the lack of an adaptive memory effect in Experiment 1 could be that the encoding processes for tastes and words are qualitatively different. For example, one recent study found that the adaptive memory advantage for words could be accounted for by fluctuations in both relational and item-specific processing (Burns et al. 2011). Relational processing between the tea flavors in our study and the encoding scenarios would have been difficult, given that the tea flavors were chosen because they defied labels. Congruency between encoded stimuli and processing tasks may also play a role in adaptive memory; recall is best when words are congruent with the encoding scenario (Butler et al., 2009; but see Nairne and Pandeirada, 2011). In Experiment 1, however, the scenario most congruent with the tea flavors was arguably the bitterness condition, in which recognition was no better than chance.

Another possibility is that both the poison and marketability conditions induced adaptive memory since financial stability can be thought of as fitness-relevant by modern day standards. While previous experiments (Kang et al., 2008; Nairne and Pandeirada, 2010; Weinstein et al., 2008) have shown that the survival advantage for words is linked to encoding scenarios that are ancestral (e.g., grasslands survival) as opposed to contemporary (e.g., city survival, planning a bank heist), Soderstrom and McCabe (2011) found no differences in recall between ancestral and modern scenarios. Although the personal acquisition of wealth could benefit survival in modern societies, our marketing scenario made no mention of direct personal gain. Our participants were working for the benefit of an institution and any expectation of reward would have been assumed and based on subjective interpretation alone. An adaptive memory benefit cannot be ruled out in the marketing condition but, if it is present, it should not be as large as it is in the survival condition, with its clear and direct link to personal fitness. Of course, any investigation of adaptive memory must consider the limited external validity of the lab where students, hopefully, don’t feel endangered.

In Experiment 2, we aimed to test another hypothesis using a method suitable for classroom demonstrations. Klein et al. (2011) suggested that the future-directed temporal orientation of the survival scenarios used by Nairne et al., (2007) explained much of the survival advantage. That is, recall is enhanced when participants are encouraged to plan for the future (as they are in most survival scenarios). We found that, although our participants, who tasted only eight samples, had higher recognition scores than those in Experiment 1, who tasted 24, they failed to demonstrate a clear future-directed temporal orientation (planning) effect when the two groups were compared. Both of our Experiment 2 scenarios had explicit survival components and, perhaps, a “survival effect” lifted recognition to a common performance ceiling that masked the “planning effect.” We can only conclude that we failed to detect a significant additive benefit for planning.

Like the original eleven volunteers who were invited to help us choose the teas, the experimental participants struggled to label the teas in a way that could possibly serve recall. We surveyed Experiment 2 participants to see if they could express any insight on how they were able to recognize the teas. Although most reported that they used olfaction and taste, and a few stated that they noted qualities like bitterness, none mentioned a label that could be attributed to a specific tea. These results can be used to illustrate how recall can be good - and sometimes perfect - in the absence of insight.

Exercises like these may lead cognitive psychology or perception students to a more comprehensive understanding of the various forms of memory and how they have been shaped by natural selection. Memory is multi-modal but, as a matter of convenience, most textbook examples and lab demonstrations employ visual or auditory/lexical stimuli. Students should understand that different forms of memory involve different mechanisms and serve many different purposes. We have developed a simple, inexpensive technique that illustrates memory in a seldom-explored sensory domain. To our knowledge, this investigation represents the first application of the Nairne et al., (2007) strategy to taste memory and we believe it holds great promise for both instruction and exploratory investigation. Instructors can easily simplify or expand the procedure or try other taste stimuli (flavored jelly beans are convenient, but we found them too easy to label). Instructors can also develop new imaginary scenarios to target specific cognitive tasks, like planning.

We still do not know if adaptive memory can be clearly demonstrated with taste stimuli, but this simple and cost-effective experimental technique shines a light on a seldom studied form of memory, it yields a levels-of-processing effect, and opens the door to discussions about the evolutionary origins of cognition.
REFERENCES


Klein SB, Robertson TE, Delton AW (2011) The future-orientation of memory: planning as a key component mediating the high levels of recall found with survival processing. Memory 19:121-139.


Received July 08, 2016; revised February 02, 2017; accepted February 03, 2017.

This work was supported in part by a Millersville University Undergraduate Research Grant.

We thank Richard Carlson, Stuart McCaughey, and the anonymous reviewers for their helpful comments on earlier drafts of this manuscript. We also thank Henry L. Roediger III for suggesting experimental avenues for this novel technique.

Henry L. Hallcock and Heather Garman conducted Experiment 1 while undergraduates and Chelsea L. Fleeger collected the data for Experiment 2 while a graduate assistant at Millersville University of Pennsylvania.

Address correspondence to: Shawn P. Gallagher, 210B Byerly Hall, Millersville University, P.O. Box 1002, Millersville, PA 17551. E-mail: Shawn.Gallagher@Millersville.edu.

Copyright © 2017 Faculty for Undergraduate Neuroscience www.funjournal.org
Although textbooks are still assigned in many undergraduate science courses, it is now not uncommon, even in some of the earliest courses in the curriculum, to supplement texts with primary source readings from the scientific literature. Not only does reading these articles help students develop an understanding of specific course content, it also helps foster an ability to engage with the discipline the way its practitioners do. One challenge with this approach, however, is that it can be difficult for instructors to select appropriate readings on topics outside of their areas of expertise as would be required in a survey course, for example. Here we present a subset of the papers that were offered in response to a request for the “most amazing papers in neuroscience” that appeared on the listserv of the Faculty for Undergraduate Neuroscience (FUN). Each contributor was subsequently asked to describe briefly the content of their recommended papers, their pedagogical value, and the audiences for which these papers are best suited. Our goal is to provide readers with sufficient information to decide whether such articles might be useful in their own classes. It is not our intention that any article within this collection will provide the final word on an area of investigation, nor that this collection will provide the final word for the discipline as a whole. Rather, this article is a collection of papers that have proven themselves valuable in the hands of these particular educators. Indeed, it is our hope that this collection represents the inaugural offering of what will become a regular feature in this journal, so that we can continue to benefit from the diverse expertise of the FUN community.

Key words: teaching; scientific literature; neuroscience literature; primary sources; classic papers.

Although the textbook has enjoyed a long period of dominance in the science classroom, trends in the direction of more active, methods-based teaching have bolstered the use of the raw material of science: the journal article. The benefits of teaching neuroscience (or any discipline) using its own literature are diverse and numerous. Perhaps the most obvious benefit has to do with the flexibility and accessibility papers afford. Decisions about whether to include one or more papers can be made quickly and these can be added to a course, as circumstances demand, even when the course is already underway. With the ever-rising costs of textbooks and college in general, a related benefit is that such papers do not lead to additional financial burdens for students. Individual papers or collections of them can serve to either supplement textbooks or, under some conditions, substitute for them altogether. Textbooks are very efficient at conveying large quantities of information; however, they tend to achieve that efficiency by sacrificing depth. Papers allow students to gain a clearer understanding of the methods and practices used in research, and provide an opportunity for a more critical assessment of a study’s conclusions (e.g., Hoskins, 2008; Willard and Brasier, 2014). This can be valuable in that it helps support the development of information literacy, critical thinking, and a general scientific disposition in students (e.g., Dirks and Cunningham, 2006; Hoskins et al., 2007; Hoskins et al., 2011; Kozeracki et al., 2006).

Late last summer one of us (Grisham) received an interesting request from a group of students as they neared the end of their summer research experience. Having read numerous articles relevant to their specific research projects, these students now wanted to read the most amazing neuroscience article ever. We are not certain they appreciated how tall an order this was. Is there a single most amazing neuroscience article ever? How would such a title be decided? When this request was put to the members of the Faculty for Undergraduate Neuroscience (FUN) by way of the organization’s listserv, more than a dozen nominations were made in short order. Although some of these papers could be considered landmarks (e.g., according to Google Scholar, Hodgkin and Huxley’s 1952d has been cited more than 16,000 times), others were far more contemporary. Not surprisingly, however, there was little evidence of consensus (although Hubel and Wiesel’s 1962 paper on the physiology of the visual cortex, cited some 9,800 times, received multiple nominations). The nominations that were received are undoubtedly just the tip of the iceberg.

Given the nature of our training, it can be difficult to identify key readings on topics outside of our immediate areas of expertise. One of us (Harrington) was reminded of a quote that was prominently displayed in the lab of one of his undergraduate professors, Vincent LoLordo: “I am so
small, and the literature...so vast.” In teaching practice, however, and this is especially true for those who teach survey courses, there is an expectation that we can find our way through this vast literature. The purpose of this article is to provide diverse recommendations for pedagogically valuable papers—recommendations made by undergraduate neuroscience educators, for undergraduate neuroscience educators—in order to promote a “collective expansion” of our appreciation of the neuroscientific literature. Each of the contributors has offered what they identified to be among the most amazing papers in a particular corner of the literature. Those who responded to the listserv request were asked to submit short descriptions of their recommended papers. In addition to describing the general content of these papers, they were also asked to describe how these papers have been useful in their teaching (what we have here termed their ‘value’). Finally, they were asked to describe the appropriate audience for their recommended papers.

We have organized the fourteen submissions included here by topic to the extent that was possible. The first four submissions address more fundamental issues including neural transmission (Gizerian), electrical excitability and K+ channels (Vilinsky), long-term potentiation (Brasier), and adult neurogenesis (Lom). These papers are followed by two concerning biological rhythms, the first addressing the effects of time-cue deprivation (Hagenauer) and the second related to sensory control of circadian rhythms (Gallagher). From here we turn to a paper about the study of the visual cortex (Olivo), and two papers that address plasticity: sensory plasticity following brain rewiring (Harrington) and cognitive plasticity following brain damage (Gordon). Following this are three papers related to the endocrine system including the effects of steroids (Sandstrom), neural correlates of sexual orientation (Grisham), and sex differences in spatial abilities (Stough). Like the paper reviewed by Stough, the final two submissions also relate to the hippocampus but instead of its role in spatial behavior they focus on memory, specifically on the induction of false memories (Linden) and the transplanting of memories (Wiest) by hippocampal stimulation. We hope that these brief descriptions will be sufficient to guide decisions about whether to include some of these papers in your courses. Moreover, we hope that collections like this one will become a regular feature of JUNE in the future.

How neurons conduct messages
Contributor: Samantha Gizerian
Topics: Membrane currents; membrane potential; ionic current; action potential; conduction; Loligo
Description: In a series of five elegant papers, Hodgkin and Huxley (and Katz) describe the movement of current through the squid (Loligo forbesi) giant axon as well as the relationship between membrane potentials and currents. These papers together were the first to describe the electrical properties of neurons and neuronal membranes. In addition to their invaluable contribution to science, these papers represent the development of technologies still in use today. At 400-800 µm in diameter, the Loligo giant axon was the first nerve structure discovered that was large enough to be penetrated by a microelectrode and, thus, could be investigated using the tools available at the time (1939-1952). In the first paper of this series, Hodgkin and Huxley describe the electrode, amplifier, and signal recorder they created and then demonstrate proof of concept for both the current clamp and voltage clamp techniques. Modern current/voltage clamp experiments are done with equipment based on these early instruments. The next three papers in the series use the techniques set forth in the initial paper to describe the currents that travel through neuronal membranes, the ions whose movement are the basis of those currents, membrane potentials, and the relationship between membrane potentials and the movement of ions through the neuronal membrane. The fifth and final paper in the series summarizes all of the previous results into the mathematical model that serves as the foundation for our understanding of membrane properties and action potential production and propagation today.

Value: This series of papers is invaluable to the history of neuroscience as well as to our understanding of the electrical properties of neurons. Hodgkin and Huxley’s findings serve as the foundation of modern neurophysiology and have broadly influenced scientists in many disciplines. Moreover, studying these papers gives students a unique insight into how data analysis is used to build models. The electrical properties of neurons, as presented in most textbooks, are represented by a series of increasingly complicated mathematical equations. Students, even those with a strong calculus background, often find these equations cumbersome to manipulate and apply because they have little understanding of the physiological processes represented. In reading these papers, students can approach the problem from the other direction. That is, how does the physiology of the neuron serve as the basis for the mathematical model? Students typically know about as much about membrane currents at the beginning of a course as Hodgkin and Huxley did at the beginning of their studies, so it is easy for students to walk alongside these pioneers in their own journey of discovery, learning how neurons work as they build and apply the mathematical model of electrical signaling in neurons and, thus, gain a deeper understanding both of the function of neurons and the basis of many of our investigations of them. Moreover, if the appropriate facilities are available, students can reconstruct Hodgkin and Huxley’s experiment as a laboratory exercise, collecting and analyzing data de novo.

Audience: These papers would be suitable for an upper-level course (or an introductory graduate course) on electrophysiology or biophysics, or in a neural physiology unit in an upper-level anatomy/physiology course. It is important that students have a strong background in math and physics, including differential calculus and electricity/magnetism in order to work through the mathematics of the model, so these papers are less suitable for lower division or introductory classes in
Discovering the mechanisms of electrical excitability

Contributor: Ilya Vilinsky

Topics: Voltage-gated ion channels; genetics; biochemistry

References: Kamb et al., 1987; Tempel et al., 1987; Wei et al., 1990; Zhou et al., 2001

Description: The first two papers in this series (Kamb et al., 1987; Tempel et al., 1987) are independent accounts of how “forward genetics” was first used to fish out the first voltage-gated K+ channel sequence in Drosophila. In forward genetics, investigators start with an interesting phenotype and use it as a “hook” to find the corresponding gene. Electrophysiology on the mutant line known as “Shaker,” where the flies display characteristic leg twitches when anesthetized, reveals defects in potassium currents. The predicted structure of the molecule responsible bore the hallmarks of a transmembrane protein and fit within a framework of how K+ channels were thought to work. For students, it is especially interesting to read these classic papers in light of our current knowledge of K+ channel structure and function. A fun exercise is to see how many of the early predictions from the initial sequence have been proven correct.

Wei et al. (1990) extends the fast-growing field of K+ channel physiology by using sequence homology to find related channels. The K+ channel family is vast. In fact, it is the most abundant member of the voltage-gated ion superfamily. Wei et al. describe four major types of K+ channels—Shaker, Shab, Shal and Shaw—in Drosophila, mouse, and by extension all animals (the naming of these genes, and the naming schemes for Drosophila genes in general, make for interesting stories on their own).

Zhou et al. (2001) exemplifies more recent structural and biochemical work on K+ channels, a field that has expanded greatly since the late 1980’s and now includes computational theorists, biochemists, and ecologists. The MacKinnon group described a high-resolution crystal structure of a K+ channel for the first time. This paper, surprisingly readable by undergraduates despite the highly specialized techniques used, is one of the reasons Rod MacKinnon was awarded the 2003 Nobel Prize in Chemistry. Some of the questions addressed by this and similar reports include: 1. How can an ion channel be so selective, especially for ions like K+ and Na+, which have similar physical characteristics? 2. Given this extreme selectivity, how is it that K+ channels have such a high conductance, with the speed of K+ ion transport being similar to that of K+ ion diffusion through water? 3. How does the K+ channel actually change shape in response to electrostatic charge across the membrane? 4 How does the detailed knowledge gained from the high-resolution structure affect the view of channel function inspired by previous experiments?

Value: This series of papers ranges from “classical” to “cutting-edge.” Voltage-gated ion channels drive electrical excitability in neurons, and thus determine how information is processed in the brain. This is especially true for voltage-gated K+ channels; these molecules are astoundingly diverse, yet share fundamental functional principles and largely determine the excitation characteristics of neurons. The story of how K+ channel structure was characterized is a great example of interdisciplinary research, incorporating genetics, electrophysiology, evolutionary biology, genomics, and biochemistry. In my courses, I use these papers to drive home the importance of utilizing multiple levels of analysis and diverse model systems.

Audience: I have taught selections from these papers in a neurophysiology lab course, where they served as a great complement to the applied work, and allowed students to put their experimental results into context. These papers would work well in a mid-level to advanced undergraduate neuroscience or neurobiology course, and are perfect for graduate level courses.

Is LTP expressed pre- or post-synaptically?

Contributor: DJ Brasier

Topics: Plasticity; LTP; synapse function; electrophysiology

References: Kauer et al., 1988; Malinow and Tsien, 1990; Stevens and Wang, 1994; Liao et al., 1995

Description: This material is incredibly fun to teach because it is as much a human history as it is a scientific one. Prior to beginning, students should have a background understanding of synaptic release and transmission, including AMPA and NMDA receptors. In the drama that will unfold, the pre-synaptic cells are the CA3 neurons in the hippocampus and the post-synaptic cells are the CA1 neurons. Students get to learn more detail about synaptic physiology (one vs. multiple points of contact and release probability vs. post-synaptic sensitivity) in the context of the debate. Typically, I begin with the Stevens and Wang (1994) and Malinow and Tsien (1990) studies to explore the pre-synaptic side of the debate. These studies, in particular, require students to review basic probability and statistics. Then, I present apparently contradictory results from Kauer et al. (1988). We discuss the assumptions and caveats of each study. Students are frequently asked to evaluate and re-evaluate their positions as the discussion unfolds. Typically, after the Stevens/Wang and Malinow/Tsien papers are discussed, the majority of the class believes LTP is pre-synaptic. Subsequent discussion of the Kauer et al. study usually leaves the class split with many students unsure or believing both changes happen. I often conclude with the discovery of silent synapses (Liao et al., 1995). This is finding especially dramatic because it provides a novel theoretical framework that upends some of the assumptions made by the Stevens/Wang and Malinow/Tsien studies; a postsynaptic change can explain most – but not all – of the results the pre-synaptic camp relied on to support their view. After this, a majority of students tend to believe the post-synaptic theory, but a minority still cites particular unanswered questions. I do not insist that students accept the consensus postsynaptic view, but encourage them to evaluate the data and come to their own conclusions.
Value: Students have the opportunity to follow an historic story in neuroscience. They experience difficulty reconciling two contradictory models each with its own supporting evidence. The personalities of the scientists involved can also be discussed. One central figure, Roberto Malinow, is especially interesting as he provides a rare case of someone experimentally overturning his own opinion. The resolution of the controversy provides a great example of how new data can force a re-evaluation of past assumptions and suddenly allow a single consistent model to explain seemingly contradictory pieces of data. Although the material is difficult for students at all levels, the insight into the scientific process is profound (Willard and Brasier, 2014). Also, a good deal of experimental cellular neuroscience is explored and students are given a chance to apply math to neuroscience. Finally, although students are not required to learn the current consensus view that LTP at this synapse is postsynaptic, at the end I do tell them that that is the consensus; the minority who feel that this consensus is not completely satisfying (there are some results it cannot fully explain) are encouraged to share their views. Students also begin to consider not only the value of data, but the reproducibility of data.

Audience: The sequence as described works for introductory students with no specific background other than synaptic transmission which is discussed leading up to this. The sequence takes three 50-minute class periods on top of the pre-requisite knowledge of how synapses work, AMPA vs. NMDA receptors, and a basic introduction to LTP. For introductory students, a good deal of class time is spent on the mathematical and theoretical foundations of the work as well as some superficial explanations of the methods of data collection and analysis. The key to success with this difficult material is targeted homework assignments before each class period to prepare the students for thinking about the data. More advanced students with a stronger statistics background can go further and explore other synapses (Weisskopf et al., 1995) or continuing challenges to the post-synaptic model (Enoki et al., 2009).

Neurogenesis in the adult human brain
Contributor: Barbara Lom
Topics: Postnatal neurogenesis; neuronal differentiation; staining, tracing, and imaging techniques
Reference: Eriksson et al., 1998
Description: This paper is a simple and powerful clinical study that asked the question, “Is the adult human brain capable of making new neurons?” in response to the longstanding view that neuron loss was irreversible in the primate brain. (Adult neurogenesis in many other vertebrates had long been known and its absence in primates had been hypothesized as a potential evolutionary trade-off.) In the 1990s new evidence of adult neurogenesis in non-human primates began to emerge, acknowledging earlier evidence of neurogenesis in the adult primate brain that had gone largely ignored. To answer this important question Eriksson et al. were fortunate to have access to rare postmortem brain tissue of cancer patients who had consented to receive injections of BrdU, a marker of dividing cells, to assess tumor proliferation near the end of their lives. BrdU is a widely used synthetic nucleoside analog of thymidine (T) that can incorporate into the DNA of S-phase cells undergoing DNA synthesis. Due to its short half-life as a monomer, but long stability when incorporated into a new DNA strand, BrdU offers a unique opportunity for scientists to obtain a snapshot of cells preparing to divide at the time of BrdU administration. Given BrdU’s ability to integrate into the genome, it is a potential mutagen, thus, not appropriate for most human studies, yet widely used in animal studies of neurogenesis. Examining the brains of five consenting cancer patients after their natural deaths (roughly two weeks to two years after BrdU administration), this team of scientists in Sweden and California report a singular and striking result: hippocampal cells in the subventricular zone (SVZ), hilus, and granule cell layer (GCL) had incorporated BrdU. Thus, these images provided the first direct evidence that the adult human hippocampus is capable of generating new neurons, many decades into its life. The research team also combined BrdU staining with immunostaining for specific and widely used neuronal markers (NeuN, NSE, Calbindin) to confirm that BrdU-stained cells also stained for these markers of differentiated neurons, suggesting neuronal differentiation had occurred.

Value: This paper firmly put out of business the popular conception that humans are born with all the neurons they will ever have. Even though the age of this paper is now approaching the age of college students, in my recent experiences many undergraduates have still heard from at least one source that adult neurogenesis is impossible in humans. Consequently, introducing this paper as a paradigm-shattering example motivates considerable student engagement. This paper can also be paired nicely with other papers examining neurogenesis in rodents and non-human primates, opening up lively conversations on the utility of animal models and species differences. In addition, this paper is particularly valuable as an example of a clinical research study which thereby stimulates natural and engaging discussions of critical research issues such informed consent, human subjects institutional review boards (HSIRBs), institutional animal care and use committees (IACUCs), appropriate sample sizes, and other important considerations for responsible conduct and scientific rigor in contemporary research.

Audience: Students in my 200-level seminar (Neuroscience of Exercise) and 300-level lab courses (Cellular and Molecular Neuroscience) have read this paper with ease and enthusiasm. The paper is accessible in part because it is short and simple; using just two related staining techniques (BrdU labeling of mitotic cells and immuno-staining of neuronal markers) I expect this paper to be similarly interesting, accessible, and relevant in just about any undergraduate neuroscience course as well as in cell biology and developmental biology courses.

Living without time: Internal timekeeping in students isolated in a WWII bunker
Contributor: Megan Hagenaer
Topics: Circadian rhythms; sleep; chronobiology
Reference: Aschoff, 1965
Description: Like many classic papers, this one is not only a forceful scientific argument, but also a personal account of an adventure exploring the unknown. It describes the rationale, methods, and results for a series of studies in which German students volunteered to live in complete timeless isolation in an underground WWII bunker for 3-4 weeks to discover whether the human body was capable of independently tracking time by means of a biological clock. The paper introduces all of the major concepts of modern chronobiology, including free-running period, entrainment, zeitgebers, and desynchrony. It is also full of fascinating details regarding what the experiment felt like to the subjects—from their initial optimism regarding how much studying they would accomplish while living in total isolation, to the system of double doors for the delivery of goods and messages from the outside world, and the inclusion of beer as part of their daily provisions. The author, Jurgen Aschoff, is considered a father of chronobiology, and one of my favorite parts of this paper is his description of his own experiences in the bunker, trying out the experimental set-up. He describes his disorientation in response to waking up in isolation and having no idea how long he had slept, as well as his complete surprise when he emerged from the bunker on the “last morning” of the experiment and discovered that it was actually 3 p.m. Through the figures, including a beautiful chart of the daily fluctuations in metabolites in Aschoff’s own urine, we can clearly see evidence of the human body generating its own daily physiological schedule in isolation, and how it slowly drifts later relative to the outside world due to the complete absence of environmental time cues. We are also introduced to the evolutionary adaptiveness of a self-sustained timekeeping system, as well as the importance of biological clocks for human health. To make this last point, Aschoff presents evidence from an individual who had his sleep/wake cycle spontaneously desynchronize from his other physiological rhythms while in the bunker. On the days when his rhythms were properly re-synchronized, his diary notes that he felt “especially well and fit.” Using these data, Aschoff correctly predicts that forced internal desynchronization may explain the malaise felt by shift-workers, astronauts, and jet-lagged international travelers. In the end, it is impossible to read this paper without wondering whether you would be willing to take the challenge, and (in the name of science!) insert a rectal thermometer and enter an underground bunker to experience true timelessness.
Value: I have used Aschoff (1965) as the first paper in a series of class periods aimed at introducing both the fundamental concepts of biological rhythms and skill of active reading. Since this is the first paper in the series, I typically recommend that the students start by reading a two-page popular science article that provides a colorful, illustrated description of the history of circadian biology and the bunker experiments (Globig, 2007). I also provide a brief introduction to the research question and basic rhythm terminology (e.g., oscillator, frequency/period, phase, amplitude), and a few pieces of advice on how to extract the most important information from scientific papers.

The text of Aschoff (1965) is unusual for a scientific paper because it is short (five pages) and relatively unintimidating. In contrast, the figures can be quite challenging, so on the day that we discuss the paper, I have the students initially work through the paper in groups with a particular focus on deciphering and explaining the most important figures (Figs. 1-4 and 7). I structure the lesson this way because over the years I have found that approximately 1/5 of my upper-level science students still have serious difficulties interpreting graphs (even scatterplots or bar charts). Approximately 15 minutes into the class period, we come back together as a class and work our way through the key concepts, methods, results, and conclusions in the paper. My goal for this exercise is to encourage students to treat scientific writing and figures as a puzzle to decipher strategically, and to create an atmosphere where students feel comfortable building their own understanding of the concepts instead of simply hiding their ignorance by parroting the paper’s own formal scientific language and figure legends. I also use the theme of self-quantification and exploration in the Aschoff paper to introduce the first project for the semester: tracking personal sleep/wake rhythms using free, downloadable smart-phone applications (e.g., Sleep Cycle, Sleep as Android, SleepBot) or commonly-sold wrist actigraphy (e.g., Jawbone Up, Fitbit, iWatch).
Audience: I have used this paper in a 400-level seminar that I teach on sleep and circadian rhythms using classic primary literature, but I believe that it could be easily adapted for a unit in an introductory neuroscience course.

Circadian rhythms are driven by photosensitive retinal ganglion cells
Contributor: Shawn P. Gallagher
Topics: Sensory and motor systems; vision; retina; photoreceptors; biological rhythms; sleep; SCN anatomy, physiology, neurochemistry
References: Freedman et al., 1999; Berson et al., 2002
Description: These two short papers describe the prediction and subsequent discovery of light-sensitive retinal ganglion cells that project to the suprachiasmatic nucleus (SCN) and influence circadian behavior. In the first report, Freedman et al. (1999) conducted behavioral experiments with blind transgenic mice. Despite having no rods or cones, the mice exhibited normal circadian wheel-running behavior that vanished when the eyes were removed. This study, using a combination of transgensics, behavior, and the crude but effective practice of enucleation, made a compelling case that something in the eye, other than rods or cones, could detect changes in ambient illumination. The second report, by Berson et al. (2002), takes the next step and provides evidence that the circadian clock is set by a subset of retinal ganglion cells that contain the photopigment melanopsin and project directly to the SCN. Using retrograde tracings (hypothalamus to retina) in rats, the authors marked the cells and recorded from them in isolated retina.
preparations. The recordings showed that, unlike the unmarked ganglion cells, these cells had unusually slow response times and were photosensitive, even when they were functionally disconnected from rods and cones. These ganglion cells, although inappropriate for image-forming visual pathways, are suitable for providing the SCN with information about slow-changing, ambient levels of illumination.

**Value:** These studies elucidate the link between the mammalian retina and circadian rhythms. Taken together, the papers also present an excellent example of progressive science. One group conducts behavioral studies and makes a prediction while the next group completes the story with anatomical and electrophysiological evidence. In the classroom, these papers could bridge a description of the retina to discussions of the hypothalamus, circadian rhythms, or parallel processing in the optic nerve. Even novice neuroscience students should be familiar with basic retinal anatomy and be impressed by the discovery of photoreceptive ganglion cells. For psychology students, the results could be used to address the clinical significance of these cells since they may present a key to understanding seasonal affective disorder. Students interested in comparative neuroanatomy could explore the evolutionary history of melanopsin, a pigment that is present in the pineal gland of non-mammalian vertebrates. Darwin, himself, was troubled by his inability to imagine intermediate stages of the eye's evolution; I think he would have liked the story of the ganglion cells that monitor sunrise and sunset.

**Audience:** These papers are clear and describe experiments that tell a simple story. The significance of the findings, however, can be discussed at many different levels. I recommend these papers for any course that introduces the anatomy of the retina. Students in basic psychology and neuroanatomy courses should understand how some retinal cells serve functions that lie, perhaps exclusively, beneath conscious visual perception. More advanced students, like those in a mid-level neuroscience class, can compare the electrophysiology of the photosensitive and non-photosensitive ganglion cells to understand how the different response types serve different functions. Finally, students in experimental design classes should appreciate how the many techniques employed in these studies converge on a single, profound discovery.

### Structure and function of the mammalian visual cortex

**Contributor:** Richard Olivo  
**Topics:** Cortical physiology; receptive fields; visual cortex  
**Description:** This paper, David Hubel's 1981 Nobel Prize address, is a clearly written account of his collaboration with Torsten Wiesel to unravel the structure and function of the primary visual cortex (V1) in cats and monkeys. It is written from a personal viewpoint, explaining the decisions that were made and why they made them. It covers the physiology of single unit recordings, including a number of figures from Hubel and Wiesel's early papers (e.g., Hubel and Wiesel, 1962; 1968) showing responses to oriented bars and edges, as well as anatomical figures showing layers in V1, ocular dominance columns, and even cytochrome oxidase blobs. It also includes their original summary figures showing models of synaptic circuits and the famous "ice cube" model of a cortical module. A few aspects have been refined by subsequent research ("hypercomplex" cells are now regarded as an extreme form of complex cells, and the "ice cube" model that shows wide ocular dominance columns perpendicular to narrow orientation columns demonstrates the concept but not the actual microanatomy of V1), but most of the information remains valid. The paper presents a detailed overview of what many consider the most important research program into the mammalian cortex, written by a pioneer in the field.  
**Value:** Although many people would regard Hubel and Wiesel's two massive research papers on primary visual cortex in cat (1962) and monkey (1968) as the true classics of this era, this review of their work is in its own way a classic that serves students very well. While it is not as detailed as the original research papers, it does provide many original figures embedded in the context of the overall research program. The review covers both their physiological experiments to record from and classify single units in primary visual cortex (V1), and also their experimental attempts to determine the functional architecture of V1. The physiological models of how simple and complex cells might be driven by excitatory input from their presynaptic elements are also included, which have remained viable, if simplified, models of V1's neural circuitry. The anatomical experiments have been superseded by newer optical techniques that more clearly reveal the overlap of ocular dominance bands and orientation pinwheels, but the "ice cube" module they proposed is of historical importance and still provides basic insight into the organization of V1. Finally, the paper includes anecdotes of Hubel and Wiesel's personal experience, starting as postdocs with Stephen Kuffler at Johns Hopkins before they moved to Harvard Medical School. The personal accounts are a further reflection of Hubel's clear and unpretentious writing style that makes this a very accessible paper for students.  
**Audience:** I have used this paper as a reading assignment in an upper-level Neuropsychology course, where we spend several weeks on visual processing from retina through extrastriate cortex. The paper provides an appropriately detailed supplement to the relatively brief account of visual cortex in most textbooks; it hits the sweet spot between overly simplified textbook accounts and the original research papers.

### Seeing with a rewired auditory cortex

**Contributor:** Ian Harrington  
**Topics:** Plasticity; cross-modal rewiring; cortical receptive fields; animal behavior
References: Sharma et al., 2000; von Melchner et al., 2000

Description: Although any number of papers by this group could have been considered for inclusion here, I have found that these two papers, published in the same issue of Nature, work particularly well together. The first paper, by Sharma et al. (2000), addressed whether cortical receptive field properties are determined by different inputs or reflect characteristics of the fields themselves. Considered another way, does auditory cortex look like auditory cortex regardless of the modality of its inputs? To address this question, neonatal ferrets had the projections from their eyes redirected to the auditory thalamus. This changed the modality of the input to the auditory cortex while maintaining the integrity of the projections from the thalamus to the cortex. The study demonstrated that cells in the rewired auditory cortex were not only visually responsive, but that their tuning for orientation and their local connections were similar to those found in normal visual cortex. The second paper, by von Melchner et al. (2000), addressed a natural follow-up question raised by the previous one: When an animal with a rewired auditory cortex is exposed to visual stimuli, does it have visual or auditory experiences? In this study, ferrets were only rewired unilaterally to allow the animals to serve as their own controls. The animals were trained to make one response to centrally presented sounds, and another response to lights presented contralateral to their intact visual pathway. Once the animals were performing the task well, visual stimuli were presented from the central location and one contralateral to the rewired pathway. The animals were tested again following the destruction of all visual pathways other than the novel one from the auditory thalamus to the auditory cortex. The results showed that when visual stimuli were presented to the rewired auditory cortex alone, they were experienced visually. As was suggested by Sharma et al. (2000), although the rewired auditory cortex is not an exact reproduction of the normal visual cortex (suggesting some intrinsic influences), it shares certain characteristics and is able to support visual experiences.

Value: With only a small risk of hyperbole, these papers have it all: interesting surgical interventions, plasticity, cortical physiology (optical imaging and some single-cell recording), retrograde tracing of cortical connections, complex behavioral testing (with good experimental controls), animal psychophysics, and lesions. Perhaps the greatest value of these papers is to demonstrate the multidisciplinary approaches that are necessary to address complex questions in neuroscience. The methods can be challenging for undergraduate students to follow but these challenges are not insurmountable. The papers, given the format of Nature, are fairly brief, but are well written and include clear graphs and other figures. Because of the format, however, some methodological details are referred to other sources. These papers can also be read at different levels. I have mentioned some of the key findings of these studies in 5-10 minutes of class time (or had students work to understand a single data figure in small groups), but could also imagine spending one or more class periods working through the details of the papers with students.

Audience: I have used these papers in several courses including a 200-level Brain & Behavior and a 300-level Sensation & Perception, but could also imagine them being used in other upper level courses, especially in a senior seminar. As mentioned above, the papers and their findings can be pitched at several levels, as their use demands.

The prefrontal cortex and moral judgments

Contributor: Rupa Gupta Gordon

Topics: Human cognition and behavior; decision making and reasoning; moral judgments; prefrontal cortex; plasticity

References: Koenigs et al., 2007; Taber-Thomas et al., 2014

Description: These two studies address the role of the ventromedial prefrontal cortex (vmPFC) in moral decision-making using the lesion method. The first, by Koenigs et al. (2007), compares the moral judgments of patients with adult-onset vmPFC damage to healthy comparison participants on personal versus impersonal moral dilemmas. These two forms of dilemma differ in that the personal form requires direct action (e.g., pushing a fat man off of a footbridge) rather than indirect action (e.g., pushing a button that diverts a train to a different track) in the interests of saving lives. Patients with vmPFC damage are more likely to endorse utilitarian actions in personal moral dilemmas, due to the lack of emotional response to the personal aspect of the moral dilemma. This finding suggests that the emotionally aversive reaction typically experienced when considering personal moral dilemmas depends upon the vmPFC. The second article, by Taber-Thomas et al. (2014), builds upon this line of research by studying the effect of developmental vmPFC damage on moral judgments. Unlike adult-onset vmPFC patients, those with developmental vmPFC exhibit more self-serving behavior (e.g., pushing an annoying boss off of a building). This demonstrates the importance of the vmPFC for learning social and moral norms during development. However, once learned, the ability to use knowledge of these norms can occur in adults independent of the vmPFC, as adult-onset vmPFC patients do not endorse self-serving situations, but the vmPFC must be intact during development for the acquisition of intact moral knowledge.

Value: Not only are these excellent examples of lesion studies using groups rather than single cases, but they also address a topic that evokes a flurry of debate in class. Furthermore, it leads nicely to a discussion of the role of free will in moral responsibility and the influence of neuroscience on other disciplines like law. It is beneficial for teaching about the research process, as it demonstrates the progression of a systematic line of research across time. Beyond the topic, there are valuable teaching opportunities in these articles to demonstrate basic research concepts. For example, in Koenigs et al. (2007), students can discuss how “personal” vs. “impersonal” moral judgments were operationally defined.
based on the content of the story, while “high” and “low”
conflict moral judgments were operationally defined based
on the consistency of participants’ responses.
**Audience:** I have used these articles in an upper level
seminar course on Cognitive Neuropsychology, where
students read and analyze primary literature. However, the
content is also appropriate for an introductory course in
human neuroscience.

**Steroids as a rejuvenating or anti-aging agent**

**Contributor:** Noah Sandstrom

**Topics:** Testosterone; steroids; aging; human behavior;
history of neuroscience

**Reference:** Brown-Séquard, 1889

**Description:** The world of professional sports is fraught
with cases of athletes seeking to gain a competitive edge
through the use of performance enhancing drugs. In many
instances, the drugs of choice are anabolic steroids (e.g.,
testosterone). In recent years, steroid allegations have
been made about Barry Bonds, Jose Canseco, Lance
Armstrong, and countless others. At the same time, a
legitimate scientific literature has explored the potential
clinical utility of steroid replacement/supplementation for a
variety of conditions (e.g., Alzheimer’s disease, hypogonadism) as well an intervention against naturally occurring declines in androgen production associated with
aging. This report by Brown-Séquard is an absolute
classic in which the author engages in self-experimentation
to explore whether administration of extracts from the
testicles of animals (guinea pigs and dogs) might positively
impact some of the abilities and faculties that he notes
have been waning with age. Brown-Séquard, noting these
deficiencies (e.g., a developing inability to concentrate,
constipation, fatigue, forgetfulness) uses a certain logic,
misguided as we may now understand it to be, to design a
study in which he grinds up testicles from animals, filters
them (no sense in injecting anything gross!), and injects
the extract into his bloodstream. He soon reports
remarkable changes in his intellect, his stamina, and his
powers of defecation and urination.

**Value:** This is a wonderfully engaging paper that has value
in several important regards. First, it speaks to the long
history of interest and research in the effects of gonadal
steroids on human behaviour. These ideas are at the
foundation of the steroid scandals that plague so much of
professional sports – an area of interest to many students.
Second, it provides a rich case with which to begin
discussing issues of experimental design and clinical trials.
The “study” had no controls. The researcher wasn’t blind;
the subject wasn’t blind. In fact, they were the same
person and we can quite confidently conclude that much of
the effect that was reported was placebo in nature. But the
paper can be a wonderful tool to start students on the
assignment of trying to design an appropriate clinical trial
to explore the fundamental questions of interest to Brown-
Séquard. The paper is a not a state-of-the-art paper –
rather, the complete opposite. It is a classic. Don’t use it
to educate as to the current state of knowledge regarding
hormone replacement therapy. Instead, use it to introduce
the topic and get people to appreciate that some of the
same questions we find fascinating today are the same
ones that researchers were intrigued by over 100 years
ago.

**Audience:** I have used this paper in an upper-level seminar on Hormones & Behavior. It’s on the reading list
for day 1 alongside a couple of news reports or sports
magazine articles on steroid abuse. I use it primarily to
introduce the concept of hormones influencing behaviour
but we revisit the topic later in the term when we talk
specifically about hormones and cognition (and talk about
current research in that area). I could also imagine it being
used in a research methods class when talking about
clinical trials.

**Neural correlates of human sexual orientation**

**Contributor:** William Grisham

**Topics:** Homeostatic and neuroendocrine systems;
anatomy; sexual orientation

**References:** LeVay, 1991; Byne et al., 2001

**Description:** These papers were selected because they
both investigate differences in a hypothalamic nucleus that is
related/correlated with differences in sexual orientation.
The second is an attempt at a replication of the first, which
we almost never see in neuroscience. In the original
who found marked sex difference in two cell groups in the
anterior hypothalamicus of humans. The interstitial nuclei
do not have the same questions that researchers were
intrigued by over 100 years ago.

**Value:** Despite a beautifully written and well-reasoned
discussion in which LeVay clearly defines the limitations of
the conclusions, this article is still severely criticized in both
academic and non-academic circles. Indeed, although it
was published nearly 25 years ago, critiques can still be
found on the web with ill-founded allegations about what
the data actually mean and what the article actually says.
These critiques and allegations, however, provide good
starting points for discussions. I don’t lead the students to
the refutations, but rather asked them to figure out if the
critiques or allegations are valid or not. These critiques
and allegations are listed below. 1) The study shows that
homosexual men are “born that way”—sexual orientation is
either genetic and/or congenital. Refutation: LeVay makes
it clear that the finding is a correlate and that the difference
could either be a cause or a consequence of engaging in
homosexual sex. 2) All of the homosexual men in LeVay’s
study had died of AIDS, so sexual orientation and HIV+ status are confounded. Refutation: A sub-group of the heterosexual men in LeVay’s study had also died of AIDS, and the difference between this sub-group of heterosexual men with AIDS vs. homosexual men was still present. Also, LeVay found there was no correlation between the volume of INAH 3 and the length of survival from the time of HIV+ diagnosis. 3) Promiscuity could actually be responsible for the decrease of INAH 3 size in homosexual men rather than sexual orientation. Refutation: LeVay admits that it could be a possible explanation. 4) LeVay is openly homosexual, so his results cannot be trusted. Refutation: The study was done blind. 5) INAH 3 is much too small to be measured reliably. Refutation: My students and I measure much smaller objects (neuron soma sizes) with great reliability (Grisham et al., 2003)—it just takes the right lens on a microscope. 6) The heterosexual HIV+ men in LeVay’s study were actually “in the closet” and should have been assigned to the homosexual group. Reassigning all of the allegedly heterosexual HIV+ men would make the difference in INAH size between homosexual and heterosexual men disappear. Refutation: Combining the data across these groups would indeed markedly reduce the difference between homosexual and heterosexual men. Nonetheless, the question would then be why there is a difference between HIV+ men who were identified as homosexual on their medical records versus those who were not. 7) The final two arguments revolve around statistical considerations. The first is that there is some overlap between the groups in INAH 3 size, therefore the differences aren’t real because every last individual wasn’t different. Refutation: Notably, we perform statistics on differences between group means, so this is possible. LeVay discusses these outliers and suggests that sexual orientation may not be the only variable that determines the size of this nucleus. 8) Byne et al. did not replicate LeVay’s finding. Refutation: I have my students take the values from Table 3 of Byne et al. and run a simple t-test on INAH 3 volume between the homosexual vs. heterosexual. (Students will have to figure out the standard deviation, but they have the standard error of the mean and the sample size, so they can.) When doing this, students will find that the t-test actually does show a significant difference. Byne et al. used a post-hoc Tukey–Kramer HSD test, which did not reveal the difference. This can generate discussions about statistical power, whether or not stringent criteria are appropriate in statistical testing, and how sacred the 0.05 criterion should be. As a footnote, Garcia-Falgueras and Swaab (2008) found similar results with male-to-female transsexuals: they had a smaller INAH 3 volume than did controls.

**Investigating sex differences in spatial ability using multiple approaches**

*Contributor: Shara Stough*

**Topics:** Sex differences; neuroethology; animal behavior and cognition; spatial learning; hippocampus

**References:** Gaulin and FitzGerald, 1986; Jacobs, Gaulin, Sherry, and Hoffman, 1990

**Description:** On average, males demonstrate superior spatial navigation skills compared to females. These sex differences are demonstrated across species. The papers I chose for this collection attempt to answer the question of why these differences exist and point to a possible neurobiological basis for this sexually dimorphic behavior. The first paper, by Gaulin and FitzGerald (1986), makes use of two closely related species, meadow voles and pine voles, with distinct mating systems to test the evolutionary hypothesis that differences in spatial ability arise due to the larger home ranges of males in polygynous species. In a field study, the authors first measure the home ranges of male and female voles of each species using implanted radio transmitters. They find that in the polygynous meadow voles, males range much farther than females, but in the monogamous pine voles, males and females have similar home ranges. The researchers then re-capture the monitored voles to test their spatial ability in a maze in the lab. As expected, the male meadow voles demonstrate better maze performance than female meadow voles, while male and female pine voles perform similarly in the maze. The second paper, by Jacobs et al. (1990), provides evidence suggesting that differences in hippocampal volume may play a role in the sexual-dimorphism observed in spatial ability in meadow voles. In this study, the researchers measured the relative hippocampal volume (hippocampal volume/brain volume) of wild-caught male and female meadow voles and pine voles. As predicted from previous behavioral results, male meadow voles had larger hippocampal volumes than female meadow voles and there was no difference between hippocampal volumes in male and female pine voles, providing a possible neurobiological basis for the observed differences in behavior.

**Value:** I think the greatest value of this set of papers is that they demonstrate the utility of investigating a question from multiple perspectives. The papers use both field and laboratory studies, and move from an investigation of behavior to the neurobiology underlying behavior. The papers are fairly straightforward, so students have a relatively easy time understanding the experiments and following the logic connecting each approach. This gives us the opportunity to move beyond simple understanding to discuss the strengths and weaknesses of each approach and to appreciate how evidence can be strengthened with the combination of multiple approaches. The papers also present fairly low-hanging fruit for students to identify follow-up experiments that would provide stronger support for the authors’ hypotheses. The papers are selected from a time point early enough in this line of research and leave

"..."
Creating false memories

**Contributor:** Monica Linden

**Topics:** Learning and memory; hippocampus; amygdala; fear; optogenetics; associative learning

**References:** Reijmers et al., 2007; Liu et al., 2012; Ramirez et al., 2013

**Description:** There are multiple papers by this group that could be considered including review articles, however the Ramirez et al. article is a tractable primary source article and is a slightly more exciting finding than the Liu et al. article. The Ramirez paper focuses on how the researchers can create a false memory of fear in a “safe” box by activating the neurons representing the “safe” box while the animal is learning to fear a “scary” box. This is accomplished using c-fos-tTA mice in combination with an AAV-TRE-ChR2-mCherry virus to express channelrhodopsin with temporal specificity in the hippocampus. There are several versions of the experiment, but in general, channelrhodopsin is expressed in neurons active in a “safe” context. These neurons are then reactivated using light while the animal is fear conditioned in a separate context. The animal’s response is then tested in the safe context, where we now see a fear response. The researchers compare the results of the experiment when the virus is injected into the dentate gyrus versus the CA1 region of the hippocampus, showing that these results are observed for dentate gyrus but not for CA1. They also use fluorescent imaging to compare expression patterns of the channelrhodopsin-expressing neurons and cFos. In an additional experiment, the researchers show that conditioned place avoidance can be induced using a similar protocol wherein the animal is fear-conditioned in a separate context, but will express a fear memory for the context reactivated during the fear conditioning. These results produce the remarkable conclusion that neurons reactivated during the delivery of an unconditioned stimulus can create a false associative fear memory to a conditioned stimulus that was not present during the delivery of the unconditioned stimulus.

**Value:** This paper excites the students as it exposes them to cutting-edge technology coupled with clear experimental design and easy-to-understand behavioral experiments. Students are drawn in by the idea of the “Marilyn Monroe Experiment” (i.e., can you artificially give the memory of one night with Marilyn Monroe?), as they see a real implementation of mice fearing a location where they have never received a shock. They also enjoy learning about optogenetics. While the technique itself is complicated, using Reijmers et al. (2007) and Liu et al. (2012) as background material allows students to appreciate the need for both temporal and anatomical specificity in gene expression. Furthermore, reviewing the results of Liu et al. helps the students understand the basic setup of the experiments in Ramirez et al. With guidance, the students can understand how each type of specificity is accomplished and why it is necessary. The paper also uses clear figures to illustrate the behavioral paradigms, so it is quite easy to follow what is happening during each iteration of the experiment. Furthermore, this paper can serve as a nice capstone in a learning and memory class because it brings together the functioning of the hippocampus with fear conditioning in an exciting, new way. Because the paper is in *Science*, it is a condensed format that is not overwhelming to the students. At the same time, it is useful to direct the students to some of the supplementary figures. (Alternatively, interpretation of the supplementary figures can make for good exam questions!)

**Audience:** I have used this paper in a junior/senior level Neurobiology of Learning and Memory course. I could imagine this paper being used in a variety of upper level courses including senior seminars. The methods could also be used in a techniques-focused course. Additionally, the findings could be discussed in a lower-level course or a non-majors course as a way to get students excited about the future of neuroscience research.

Knowledge transplant

**Contributor:** Michael C. Wiest

**Topics:** Learning and memory; rodent; hippocampus; micro stimulation; electrical stimulation; memory-transfer; ensemble; neuro-prosthetic

**Reference:** Deadwyler et al., 2013

**Description:** Deadwyler et al. (2013) demonstrates a transfer of task-relevant knowledge from a well-trained donor rat to a relatively naïve recipient rat in the form of neural activity patterns induced by multi-site electrical micro-stimulation in the recipient hippocampus. The task is a well-studied delayed non-match to sample (DNMS) task that involves remembering the position of a sample lever over the course of a 1 to 60 second delay. Distributed spiking activity patterns were measured using multi-site recordings in the donor rat, and these formed the basis for a computational model of the hippocampal “ensemble codes” or spatial activity patterns corresponding to successful and unsuccessful memory encoding during the sample phase of the task. “Successful” activity patterns were then induced in the donor hippocampus by multi-site electrical stimulation, resulting in dramatic performance improvements compared to un-stimulated trials or stimulation with unfavorable activity patterns. This suggests the exciting possibility of transferring memories from one brain to another in order to enhance memory performance or recover lost memory functions — as distinct from transferring immediate sensory information or motor commands as in a related “brain-to-brain interface” paper (Pais-Vieira et al., 2013). Another paper from the Deadwyler and Hampson group (Hampson et al., 2013)
shows that neuro-prosthetic memory enhancement is feasible in primates; however that paper does not show transfer between animals as in the rat paper above (Deadwyler et al., 2013).

**Value:** The idea of directly inducing specific experiences and knowledge by brain stimulation is naturally exciting. It evokes images from popular science fiction movies like The Matrix and Inception and can lead one to fascinating philosophical issues. But for more practical-minded students, it is not that difficult to imagine huge medical and societal benefits if we become able to implant skills and knowledge as needed in people — people with memory loss or people with important and difficult jobs. Given these motivations for studying the paper, there is also a substantial pedagogical payoff in terms of understanding the experimental design, the multi-channel recording and stimulation methods, distinct functional roles for hippocampal areas CA1 and CA3, and the concept of an ensemble code that undergoes transformations over time that predict success or failure in the task. The “non-linear multiple-input multiple-output” model in the paper may be taken as an example of the importance of mathematical and computational modeling approaches in systems neuroscience.

**Audience:** I have had a small group of senior undergraduate neuroscience majors present Deadwyler et al. (2013) for discussion in our senior capstone seminar course. I consider it a challenging paper even for that relatively advanced group of students. However, I think the gist of the paper is accessible and valuable even if some technical details are skimmed over. For example, students can get some appreciation of how computational modeling can be useful from Deadwyler et al. (2013), even if they are not immediately motivated to master those methods themselves. Thus, I think this paper could be discussed in lower-level classes, but in that case I would probably not expect students to read the whole paper and I would present some of the core ideas myself rather than expecting students to lead the discussion. For non-majors, I would probably present a digested version of the results rather than assigning sections of the paper.

**REFERENCES**


Brown-Séquard CE (1889) Note on the effects produced on man by subcutaneous injections of a liquid obtained from the testicles of animals. Lancet 2:105-107.


characterization of Shaker, a Drosophila gene that encodes a potassium channel. Cell 50:405-413.

Received August 13, 2015; accepted August 22, 2015.

The authors thank the membership of the Faculty for Undergraduate Neuroscience whose open discussions make this kind of collaboration possible.

Address general correspondence to: Dr. Ian A. Harrington, Department of Psychology and Program in Neuroscience, 639 38th St., Augustana College, Rock Island, IL 61201. Email: ianharrington@augustana.edu

Specific correspondence can be sent to individual contributors: D. J. Brasier, dbrasier@cmu.edu; Shawn P. Gallagher, Shawn.Gallagher@millersville.edu; Samantha S. Gizerian, sgizerian@vetmed.wsu.edu; Rupa Gordon, rugordon@augustana.edu; William Grisham, grisham@lifesci.ucla.edu; Megan Hagenauer, hagenaue@umich.edu; Monica Linden, monica_linden@brown.edu; Barbara Lom, balom@davidson.edu; Richard Olivo, rolivo@smith.edu; Noah J. Sandstrom, nsandstr@williams.edu; Shara Stough, sharastough@augustana.edu; Ilya Vilinsky, vilinsia@ucmail.uc.edu; Michael Wiest, mwiest@wellesley.edu.
The validity of the Major Field Test in Psychology as a programme assessment tool

Shawn P. Gallagher & Shaun P. Cook

The Major Field Test in Psychology (MFT) is a standardised test designed to assess subject mastery at the conclusion of an undergraduate career. Eighty-one graduating majors completed the MFT and 56 of them also took a multiple-choice exam of questions drawn randomly from an introductory psychology test bank. Like the MFT, the constructed exam was divided into four subscales. A second sample of 29 novice majors also completed the MFT. For the advanced majors, total and subscale scores from the two tests were highly correlated (all p<0.01). Advanced majors scored above national norms and were significantly better than the novices. However, few courses could be linked to performance in the associated MFT subscales but general academic indices like grade point average were highly predictive. The novice MFT scores were similar to the norms for advanced majors. We conclude that the MFT measures broad, basic understanding, not advanced mastery.

Keywords: Major Field Test; programme assessment; undergraduate psychology programmes.
assessed as part of a Major Field Test (MFT), for psychology as well as several other undergraduate subject areas. These tests are ‘designed to measure the critical knowledge and understanding obtained by students in a major field of study (ETS, 2013)’. A report published by ETS not only suggests that their products can be critical components of institutional assessment but also calls for their regular administration at all accredited US institutions:

Post-secondary education today is not driven by hard evidence of its effectiveness… A nationwide system of accountability needs to be developed within the context of efforts to monitor and improve higher education… We recommend that… accrediting agencies be charged with integrating a nationwide system of assessing student learning into… ongoing reviews of institutions of higher education (ETS, 2006).

For US institutions, especially those who rely on government support and are annually scrutinised by a tax-paying public, the message from ETS is clear: The MFT is not only a useful tool for measuring ‘critical knowledge and understanding’ but it is also an objective, quantifiable way to for a reviewing body, governmental or otherwise, to address accountability and, possibly, influence university rankings. The aim of this study was to evaluate the validity of the psychology MFT, a standardised multiple-choice test, as a useful tool for internal or external departmental assessment. Additionally, the strategies that we have employed can be applied to any such test that purports to assess advanced understanding in a discipline.

Our psychology department serves over 400 undergraduate majors and aspires to provide each one with a broad disciplinary foundation. One element of our curricular review process includes the periodic administration of the MFT. More than 7000 students at 200 institutions have taken the 140-question multiple-choice test since 2010 (see the Appendix). ETS identifies participating institutions as those with five or more students taking the test in the 2010–2012 period but there is no specific recruitment protocol. Some institutions administer it regularly to all graduating students (Stoloff & Feeney, 2002) but the cost of $25 US per exam might lead others, like ours, to administer the test only periodically and to only a sample of students. The MFT test questions cover a wide range of topics in psychology, and test results are reported with a single total score as well as subscores for Learning and Cognition (LRN), Perception, Sensory, Physiology, Comparative, and Ethology (PHYS), Clinical, Abnormal, and Personality (CLIN), and Developmental and Social (DEV). Although the objective results generated by a widely-used test give us a convenient set of numbers to analyse and present, a review of the literature prompted us to take a closer look at the validity of the test (Frazier & Edmonds, 2002; Stoloff & Feeney, 2002).

Thousands of psychology majors from hundreds of institutions have taken the MFT, and many departments have used MFT test results to evaluate their programmes and revise their curricula (Dolinsky & Kelley, 2010; ETS, 2013; Frazier & Edmonds, 2002; Stoloff & Feeney, 2002). According to ETS (2005), the MFT is a tool designed to ‘assess mastery of concepts, principles, and knowledge expected of students at the conclusion of an academic major in specific subject areas’ (p.1). Presumably, performance on the test will improve as a student progresses through the undergraduate major and acquires a deeper and more comprehensive understanding of psychology. Advanced majors should outperform novices. However, the reports from those using the MFT in programme assessment are, at best, inconclusive about the factors that lead to high test scores. It is not surprising that MFT scores consistently correlate highly and significantly with general indicators of academic proficiency like grade point aver-
ages (GPAs) and SAT1 (College Board) scores (Dolinsky & Kelley, 2010; Stoloff & Feeney, 2002), but some found that these measures have more predictive value than discipline-specific indicators of acquired knowledge like the number of psychology courses taken (Dolinsky & Kelley, 2010; Stoloff & Feeney, 2002). In fact, success on the MFT has been linked to performance in only a few courses (Stoloff & Feeney, 2002) or, as Frazier and Edmonds (2002) concluded, just one. Although some courses cover subjects named in the MFT subscale titles, Stoloff and Feeney (2002) concluded that students who take courses like Abnormal Psychology, Social Psychology, and Developmental Psychology scored no better on the respective subscales than those who do not.

The published literature challenges the assertion that MFT scores reflect knowledge acquired at ‘the conclusion of an academic major’ in psychology. Although our informal review of past MFT questions suggested that the psychology test had high face validity, it also revealed that many of the questions dealt with foundational material covered in our introductory psychology course (see the Appendix for sample questions). Results found in the published literature (e.g. Frazier & Edmonds, 2002; Stoloff & Feeney, 2002) also suggest that the MFT may be assessing a student’s mastery of basic concepts that are then reinforced in advanced courses. We, therefore, hypothesised that MFT total scores and subscores would correlate with results from an assessment test that we constructed with questions randomly drawn from an introductory psychology textbook’s test bank. We also hypothesised that, despite the fact that our advanced majors historically perform at or above the national MFT means, a sample of novice psychology majors would also score close to the normative means.

Method

Programme

Our undergraduate programme requires psychology majors to complete a minimum of 33 credit hours in psychology. Most students are between the ages of 18 and 22 years and complete the undergraduate programme of 120 credits in four years. Typical courses are worth three credit hours and hold three weekly classroom hours over a 14-week period. Lab courses carry four credits and require additional classroom hours that involve running experiments or learning computer applications. All psychology students are required to take a three-credit course in General Psychology and then two sequential four-credit courses in Statistics and Experimental Design. They must then choose at least one four-credit advanced lab course in Child Development, Cognitive Psychology, Learning and Motivation, or Sensation and Perception. Students select the remainder of courses according to their interests and career objectives.

Percentage grades for each course are converted to grade points that fall on a scale from 0 (65 per cent or lower) to 4.0 (95 per cent or higher). These points are then used to calculate a grade point average for all courses (overall GPA) as well as psychology courses in particular (major GPA). Psychology majors who fail to achieve a major GPA of 2.25 by the end of their second year are typically dismissed from the programme. Psychology majors who fail to achieve an overall GPA of 2.0 are not awarded degrees.

Participants

Eighty-one advanced psychology majors, defined as students who were within one month of completing the required minimum of 33 credits hours in psychology (M=41.1 credits, SD=8.27), took the MFT as well as other assessments and surveys. A second sample of novice majors (N=29),

---

1 The SAT is a standardised test used for admissions at most undergraduate institutions in the US. It contains three major sections: Critical Reading, Mathematics, and Writing.
having completed no more than six credit hours of psychology (two courses), also took the MFT. We recruited the students two weeks in advance and told them that they were going to take a test of general subject knowledge. They were not instructed to study for the tests and were given course credit for participation, regardless of performance.

**Materials**

The assessment tests included the web-based version of the MFT and a 100-question introductory psychology final exam (IPFE) that we created with ExamView® software (Version 5.2.0 FS Creations/now eInstruction, Denton TX) that accompanied an introductory psychology textbook by James Kalat (2005). The questions are published in the book’s test bank (Meine & Kalat, 2005). This textbook was currently not in use at the university and was chosen because of the author’s expertise in the development of standardised tests including the Graduate Record Examination (GRE) in psychology, also produced by ETS (Kalat & Matlin, 2000). The ExamView® software automatically selected 100 multiple-choice questions from a final exam question pool and compiled them in the form of a single exam. We then used the software to convert the exam to an HTML file that could be loaded onto a university server and administered via campus computers. The authors independently categorised each IPFE question into no more than two of the MFT subgroups, and only questions nominated by both authors were included in a given subscale score. Distributions were as follows; 23 questions were classified under LRN, 18 were classified under PHYS, 21 were classified under CLIN, and 17 were classified under DEV. Four questions were scored under two of the four subgroups, and the remaining 25 questions pertained to other areas, such as ethics and methodology, and were used only in the total score calculation. Surprisingly, the proportions of questions allotted to each subgroup in the IPFE were not significantly different from those of the MFT ($\chi^2(4)=2.06, \ p>.05$) which also includes questions (about 20 per cent) that are not specific to any subgroup and used only in the total score calculation (ETS, 2005).

**Procedure**

We administered tests with campus computers during single, uninterrupted blocks of no more than two hours. Although proctors were available, no student requested assistance after receiving instructions and starting a test. We counterbalanced the order of the MFT and IPFE for the advanced majors at intervals of between two and 14 days. Students took the online MFT according to ETS protocol, and results were automatically submitted to and scored by ETS. We then accessed overall and subscale scores through the ETS website. We could not access answers to specific test questions or determine how many questions had been answered. When students completed the IPFE, the software scored the tests and emailed results to the first author. We entered the responses into a spreadsheet application that calculated subscale scores.

**Results**

Table 1 summarises mean GPA, SAT and MFT scores for the advanced and novice majors. MFT scores are reported on a scale from 120 to 200; raw scores are not provided. For the advanced students, major GPAs ($M=3.14, \ SD=0.57$) were similar to their overall GPAs ($M=3.15, \ SD=0.52$), and the mean overall GPA did not differ significantly from that of the previous year’s graduates ($z=1.35, \ p=0.18$, two-tailed). MFT total and subscale scores were significantly correlated with all GPA and SAT measures (Table 2).

The advanced majors ($N=81$) performed well on the MFT ($M=160.7, \ SD=15.0$, Range: 126 to 196) and the total and subscale score means were all above the corresponding normative means with the DEV subscale mean being the only one not significantly so (all other $p<0.05$, two-tailed $z, N=7077$). Of the advanced majors who completed the
Table 1: Comparison of GPA, SAT, and MFT total and subscale scores for Advanced Majors (>=33 psychology credits completed), Novice Majors (<=6 psychology credits completed), and the Normative Sample (N=7077).

<table>
<thead>
<tr>
<th></th>
<th>Advanced Majors M (SD)</th>
<th>Novice Majors M (SD)</th>
<th>Normative Sample M (SD)</th>
<th>Advanced–Novice</th>
<th>Normative–Novice</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPA Overall</td>
<td>3.15 (0.52)</td>
<td>3.29 (0.59)</td>
<td>-</td>
<td>t (107)=−1.21</td>
<td>−</td>
</tr>
<tr>
<td>SAT Reading</td>
<td>526.0 (79.5)</td>
<td>555.6 (73.1)</td>
<td>-</td>
<td>t (78)=−1.41</td>
<td>−</td>
</tr>
<tr>
<td>SAT Math</td>
<td>509.2 (89.0)</td>
<td>535.0 (70.9)</td>
<td>-</td>
<td>t (78)=−1.13</td>
<td>−</td>
</tr>
<tr>
<td>SAT Writing</td>
<td>522.0 (74.2)</td>
<td>543.9 (81.3)</td>
<td>-</td>
<td>t (71)=−1.06</td>
<td>−</td>
</tr>
<tr>
<td>MFT Total</td>
<td>160.7 (15.0)</td>
<td>152.3 (12.9)</td>
<td>156.1 (15.5)</td>
<td>t (108)=2.69*, d=0.26</td>
<td>z=1.32</td>
</tr>
<tr>
<td>MFT: LRN</td>
<td>59.1 (15.5)</td>
<td>52.0 (12.8)</td>
<td>55.6 (15.2)</td>
<td>t (108)=2.22*, d=0.21</td>
<td>z=1.28</td>
</tr>
<tr>
<td>MFT: PHYS</td>
<td>62.8 (15.5)</td>
<td>56.2 (11.9)</td>
<td>56.1 (14.7)</td>
<td>t (108)=2.09*, d=0.20</td>
<td>z=−0.04</td>
</tr>
<tr>
<td>MFT: CLIN</td>
<td>60.9 (14.9)</td>
<td>53.3 (14.8)</td>
<td>56.4 (15.4)</td>
<td>t (108)=2.37*, d=0.23</td>
<td>z=1.08</td>
</tr>
<tr>
<td>MFT: DEV</td>
<td>57.9 (14.3)</td>
<td>52.5 (15.3)</td>
<td>55.5 (14.8)</td>
<td>t (108)=1.71*, d=0.16</td>
<td>z=1.09</td>
</tr>
</tbody>
</table>

MFT=Major Field Test; LRN=learning and cognition; PHYS=perception, sensation, physiology, comparative, and ethology; CLIN=clinical, abnormal, and personality; DEV=developmental and social; Major GPA=grade point average for psychology courses; Overall GPA=grade point average for all courses. *p<0.05, one-tailed.
Table 2: Advanced Majors: Pearson correlations of Introductory Psychology Final Exam Scores and Academic Scales with MFT scores.

<table>
<thead>
<tr>
<th></th>
<th>MFT: Total</th>
<th>MFT: LRN</th>
<th>MFT: PHYS</th>
<th>MFT: CLIN</th>
<th>MFT: DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPFE: Total</td>
<td>r (54)</td>
<td>.83**</td>
<td>.75**</td>
<td>.71**</td>
<td>.77**</td>
</tr>
<tr>
<td>IPFE: LRN</td>
<td>r (54)</td>
<td>.79**</td>
<td>.74**</td>
<td>.66**</td>
<td>.71**</td>
</tr>
<tr>
<td>IPFE: PHYS</td>
<td>r (54)</td>
<td>.71**</td>
<td>.66**</td>
<td>.68**</td>
<td>.63**</td>
</tr>
<tr>
<td>IPFE: CLIN</td>
<td>r (54)</td>
<td>.72**</td>
<td>.66**</td>
<td>.61**</td>
<td>.69**</td>
</tr>
<tr>
<td>IPFE: DEV</td>
<td>r (54)</td>
<td>.68**</td>
<td>.56**</td>
<td>.63**</td>
<td>.62**</td>
</tr>
<tr>
<td>Major GPA</td>
<td>r (79)</td>
<td>.73**</td>
<td>.70**</td>
<td>.56**</td>
<td>.67**</td>
</tr>
<tr>
<td>Overall GPA</td>
<td>r (78)</td>
<td>.69**</td>
<td>.65**</td>
<td>.53**</td>
<td>.64**</td>
</tr>
<tr>
<td>Major Credits</td>
<td>r (79)</td>
<td>.41**</td>
<td>.43**</td>
<td>.42**</td>
<td>.30**</td>
</tr>
<tr>
<td>SAT Reading</td>
<td>r (60)</td>
<td>.59**</td>
<td>.52**</td>
<td>.44**</td>
<td>.50**</td>
</tr>
<tr>
<td>SAT Math</td>
<td>r (60)</td>
<td>.47**</td>
<td>.44**</td>
<td>.44**</td>
<td>.30*</td>
</tr>
<tr>
<td>SAT Writing</td>
<td>r (53)</td>
<td>.58**</td>
<td>.50**</td>
<td>.39**</td>
<td>.51**</td>
</tr>
</tbody>
</table>

IPFE=Introductory Psychology Final Exam based on questions selected randomly from an introductory text (Kalat, 2005) test bank; MFT=Major Field Test; LRN=learning and cognition; PHYS=perception, sensation, physiology, comparative, and ethology; CLIN=clinical, abnormal, and personality; DEV=developmental and social; Major GPA=grade point average for psychology courses; Overall GPA=grade point average for all courses. Major Credits=Psychology Credits Completed.

*p<0.05, **p<0.01, correlations for corresponding subscale scores are in bold.
Table 3: Novice Majors: Pearson correlations of Academic Scales with MFT scores.

<table>
<thead>
<tr>
<th></th>
<th>MFT: Total</th>
<th>MFT: LRN</th>
<th>MFT: PHYS</th>
<th>MFT: CLIN</th>
<th>MFT: DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall GPA</td>
<td>r (29)</td>
<td>.58**</td>
<td>.59**</td>
<td>.21</td>
<td>.45*</td>
</tr>
<tr>
<td>SAT Reading</td>
<td>r (18)</td>
<td>.75**</td>
<td>.70**</td>
<td>.65**</td>
<td>.62**</td>
</tr>
<tr>
<td>SAT Math</td>
<td>r (18)</td>
<td>.77**</td>
<td>.70**</td>
<td>.79**</td>
<td>.62**</td>
</tr>
<tr>
<td>SAT Writing</td>
<td>r (18)</td>
<td>.63**</td>
<td>.70**</td>
<td>.37</td>
<td>.48*</td>
</tr>
</tbody>
</table>

MFT=Major Field Test; LRN=learning and cognition; PHYS=perception, sensation, physiology, comparative, and ethology; CLIN=clinical, abnormal, and personality; DEV=developmental and social; Overall GPA=grade point average for all courses.

*p<0.05, **p<0.01
MFT, 56 also completed the IPFE. Table 2 shows the Pearson correlations between the elements of the MFT and IPFE. The MFT total scores and subtest scores correlated significantly not only with their IPFE counterparts but also with every other element of the IPFE. The IPFE LRN and PHYS subscale scores correlated best with their MFT counterparts. Major GPA, overall GPA, and major credits completed also correlated significantly with all MFT scores (all \(p<0.05\)).

All the novice majors (\(N=29\)), had completed a course in introductory psychology, and 11 of the 29 had also completed a course in child development. For the novice majors, MFT total and subscale means were all significantly lower than those of the advanced majors (\(M=152.3, SD=12.9, \text{ Range: 131 to 174}\)), although effect sizes were modest (Cohen’s \(d\)). Novice means were not significantly lower than the normative means provided by ETS (all \(p>0.05\), one-tailed \(z\)). Five novice majors had MFT total scores that placed them above the 80th percentile for the individual national normative scores (ETS, 2013). MFT total scores and most subscale scores were highly correlated with GPA and SAT scores (Table 2).

We performed a stepwise linear regression analysis to predict MFT total scores from Major GPA, Major Credits, SAT Reading, SAT Math, and SAT Writing scores. Overall GPA was highly collinear with Major GPA (\(r=.92\)) and, therefore, excluded from the analysis. Each of the predictor variables had a significant (\(p<.05\)) zero-order correlation with MFT total score (Table 2), but only Major GPA (\(\beta=.44\)), Major Credits (\(\beta=.28\)), and SAT Writing (\(\beta=.29\)) had significant partial effects in the full model (all \(p<.01\)). The model employing only these three predictors accounted for 62 per cent of the variance in MFT total score, (\(F(3,51)=27.96, p<.001\)).

Like Stoloff and Feeney (2002), we examined the effect of completing specific courses on MFT performance for all courses completed by 10 per cent to 90 per cent of our advanced majors. This analysis did not allow us to look at the effects of the three required courses taken by all advanced students (General Psychology, Statistics and Experimental Design I & II) and we excluded broad content courses that could be applied to all four subscales (History and Systems, Tests and Measurements). Using \(t\)-tests to compare the mean scores of those who did and did not take specific courses, we found no content-specific effects for 16 courses (all \(p>0.10\)) which, based on course descriptions, should contribute to specific MFT subscale scores (Table 4). The literature provides little evidence linking specific courses to subscale scores but Dolinsky and Kelley (2010) noted dramatic improvements in PHYS subscale scores after the implementation of a required second-year course in physiological psychology. In our case, those who took physiological psychology (\(N=29\)), outperformed their peers in all MFT measures (all \(p<.01\), one-tailed \(t\)). Those taking physiological psychology also had higher major GPAs (\(t(79)=2.34, p=0.01\), one-tailed) and completed more credits (\(t(79)=5.76, p<0.01\), one-tailed) than their counterparts. A similar broad, nonspecific effect was observed for History and Systems. Both of these courses tend to attract students with graduate school aspirations.

In a final, exaggerated effort to find a link between course content and an MFT subscale score, we decided to compare two specific groups of advanced majors. Students who take physiological psychology often have plans to pursue graduate training in experimental psychology and we were not surprised by the performance of this highly motivated group. We compared these students to a second group of motivated students with postgraduate plans in clinical, rather than experimental, psychology. In our case, such students can be found by identifying those who have taken both abnormal psychology and personality theory, two topics that fall firmly within CLIN subscale of the MFT. We compared students who took both of these courses but not physiological...
Table 4: Courses taken by 10 to 90 per cent of advanced majors for which no content-specific effect on MFT was observed. Each course is listed along with the abbreviation of the subscales for which a specific effect might be anticipated. Required courses taken by all students (General Psychology, Statistics and Experimental Design I & II) and those with broad content (History and Systems of Psychology, Tests and Measurements) were not included.

<table>
<thead>
<tr>
<th>Course</th>
<th>Subscales</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abnormal Psychology – CLIN</td>
<td></td>
</tr>
<tr>
<td>Behavior Modification – LRN, CLIN</td>
<td></td>
</tr>
<tr>
<td>Childhood Disorders – CLIN, DEV</td>
<td></td>
</tr>
<tr>
<td>Cognitive Psychology – LRN, PHYS</td>
<td></td>
</tr>
<tr>
<td>Counseling Strategies – CLIN</td>
<td></td>
</tr>
<tr>
<td>Child Development Lab – DEV</td>
<td></td>
</tr>
<tr>
<td>Drug Addiction – CLIN, PHYS</td>
<td></td>
</tr>
<tr>
<td>Family Systems – CLIN</td>
<td></td>
</tr>
<tr>
<td>Health Psychology – PHYS</td>
<td></td>
</tr>
<tr>
<td>Human Adjustment – CLIN</td>
<td></td>
</tr>
<tr>
<td>Human Relations – CLIN, DEV</td>
<td></td>
</tr>
<tr>
<td>Industrial/Organisational Psychology – DEV</td>
<td></td>
</tr>
<tr>
<td>Learning and Motivation – LRN, PHYS</td>
<td></td>
</tr>
<tr>
<td>Lifespan and Human Development – DEV</td>
<td></td>
</tr>
<tr>
<td>Personality Theory – CLIN</td>
<td></td>
</tr>
<tr>
<td>Sensation and Perception – PHYS</td>
<td></td>
</tr>
<tr>
<td>Social Psychology – DEV</td>
<td></td>
</tr>
</tbody>
</table>

LRN=learning and cognition; PHYS=perception, sensation, physiology, comparative, and ethology; CLIN=clinical, abnormal, and personality; DEV=developmental and social.

psychology ($N=37$, $M=63.27$, $SD=13.4$) to those who took physiological psychology but not abnormal psychology or personality theory ($N=16$, $M=60.81$, $SD=19.9$). Contrary to our hypothesis, those who took physiological psychology but not the two clinical courses had a higher mean CLIN subscale score ($M=66.7$, $SD=11.01$) than those who took the clinical courses but not physiological psychology ($M=58.5$, $SD=15.3$). These two groups did not differ significantly in mean major GPA ($t(51)=1.16$, $p=.25$, two-tailed), overall GPA ($t(51)=.836$, $p=.41$, two-tailed), or SAT scores (all $p>0.1$, two-tailed $t$). The physiological psychology group did, however, complete more major credits ($t(51)=5.31$, $p<0.01$). For these matched samples, we suspect that the seemingly paradoxical result is due to exposure to broad, basic content.

**Discussion**
Our department has been using the MFT for several years as part of our periodic review process. Although the results are objective and simple to compile and present, published reports compelled us to evaluate the test’s validity (Frazier & Edmonds, 2002; Stoloff & Feeney, 2002). Previous studies have challenged the notion that multiple choice exams can measure critical thinking skills (Stanger-Hall, 2012) and found that they typically over-estimate general content mastery (Funk & Dickson, 2011). The results
of our investigation are consistent with these findings and have led us to four conclusions. First, MFT total and subscale scores can be predicted by performance on a comprehensive introductory psychology exam. Second, although our advanced students performed well on the MFT, as predicted by Stoloff and Feeney (2002), we found no link between specific courses and MFT subscale scores. Third, novice majors can perform well on the MFT despite having completed no more than two psychology courses. Finally, performance on the MFT appears to be a function of overall academic potential as measured by standardised tests and GPA.

In our opinion, the sample questions provided by ETS (Appendix A) and the questions that we have seen on the actual tests assess basic, introductory-level knowledge. Although we observed significantly different MFT scores between novice and advanced students, the question of what led to these differences remains. We suspect that these differences are due to the fact that the advanced students have had more opportunities to review and discuss basic material common in introductory courses. The students who do well on the MFT might well have a deep understanding of psychology, but they also have had many opportunities to review core content; basic Freudian principles get reinforced through courses in Personality Theory or Abnormal Psychology and basic Neuroanatomy is reviewed in courses like Sensation and Perception. Our top-scoring novices are exceptional students who have performed well in all of their college courses and have probably mastered the material they have encountered so far. However, regardless of how exceptional a novice student happens to be, a test ‘designed to assess mastery of concepts, principles, and knowledge expected of students at the conclusion of an academic major’ (ETS, 2005) should contain material well beyond his or her reach. Although one of our five top-scoring novice students took an advanced placement psychology course in high school, none of the others had any supplemental psychology education or training; two had completed only General Psychology, and three had completed General Psychology and Child Development. We suspect that exceptional novice students from any university would do well on the MFT, and the gaps we observed between advanced and novice scores might be even smaller for departments with two-semester introductory psychology requirements. Although they will not have data that they can directly compare to a normative sample, we recommend that, if a department aspires to measure the broad, basic content knowledge of its advanced majors, they construct an IPFE of their own with questions from a comprehensive introductory psychology textbook.

Although our specific claim that the MFT fails to validly assess mastery might be new, our results are consistent with many previous reports. Like Stoloff and Feeney (2002), we found that although psychology credits completed was significantly correlated with MFT total, GPA, either Overall or Major, accounted for approximately 50 per cent of the variance in MFT total while credits completed accounted for less than 20 per cent of the variance. In another study exploring MFT correlates, Frazier and Edmonds (2002) found that, of all the academic indicators evaluated, MFT scores were predicted only by performance in a single in-depth, and closely mentored, research methods course. The authors noted that those who excel in this class are typically those planning on graduate school and, we suggest, that they were the ones who made a point of mastering the basics from the beginning.

In a four-year longitudinal study, Dolinsky and Kelly (2010) noted impressive gains in MFT scores after converting Abnormal Psychology, Physiological Psychology, and History and Systems from optional to required courses. Although it is likely that the course in Abnormal Psychology contributed to the gains in the CLIN subscale scores and that Physiological
Psychology helped the scores on the PHYS subscale, the LRN and DEV subscales exhibited similar gains without any curricular modifications in the associated sub-disciplines. It is possible that the History and Systems course, due to the topic’s broad nature, facilitated MFT gains across the four subscales. Although History and Systems should introduce even the most advanced students to new material, it is impossible to teach the course without reviewing key elements from across psychology. Dolinsky and Kelly seem aware of the benefits of reviewing basic material and note that their department briefly contemplated adding a capstone review-and-synthesis course or a ‘mini-review’ aimed at improving MFT scores. Although they concluded that dedicating a single course to a standardised test would be an inappropriate allocation of resources, it probably would have worked.

Our homemade assessment test, the IPFE, was generated with software that randomly drew questions from an introductory psychology textbook’s final exam test bank. Having little knowledge about how MFT questions were selected, we concluded that the ExamView® software would be the least biased arbiter of content. As it turned out, the proportional distribution of content was remarkably close to that of the MFT, possibly due to the fact that the textbook author has been affiliated with ETS. It is, however, noteworthy that the IPFE was much shorter than the MFT and that the strength of the correlation between the subscale pairs was proportional to the number of IPFE questions in the category. We did not want to bias the IPFE by adding or subtracting questions, but we believe that the MFT subscale correlations could have been even higher had we made the IPFE longer and actively managed content with input from across our department.

If the MFT is primarily measuring basic knowledge in psychology, this fact will remain hidden each time a department’s top students live up to faculty expectations by outperforming their peers. To the casual observer, the test appears to be doing what it has been designed to do. However, we have found that MFT scores can be accurately predicted by performance on an introductory psychology exam, the subscale scores cannot be clearly linked to courses, and the test cannot clearly differentiate between novices and a large normative sample of graduating majors. Our department is now faced with the challenge of deciding how these and future MFT results will be used. We will probably continue to employ the test in a limited fashion for the sake of comparative analysis but, in its current state, the MFT will have little impact on our department’s curricular development. As a result of this investigation, we have redoubled our efforts to evaluate our own course offerings and requirements. We are also constructing our own assessment tests which will incorporate some multiple choice questions but also include data analysis and critical reviews of published studies. Our test will, of course, be administered at a fraction of the MFT's cost.

Comprehensive and effective programme review is a multi-dimensional endeavor and each component of the process must be closely scrutinised and validated (Dunn et al., 2007). Like many universities, our institution is facing increased pressure, from both internal and external sources, to objectively demonstrate programme effectiveness, and the results from assessment initiatives can have profound implications. Courses can be offered or withdrawn and faculty bodies can expand or contract based on data generated by quantitative instruments like the MFT. Furthermore, the administration of these tests can command a significant amount of money, faculty and student hours, and facility space. Our hope is that the resources dedicated to any assessment initiative are in proportion to its demonstrated validity.
Author Notes
We thank Rich Carlson and Bea Grosh for reviewing and commenting on earlier drafts of this manuscript.

This work was supported in part by the Millersville University Department of Planning, Assessment and Analysis, Dr Lisa R. Shibley, Assistant Vice President.

This work was presented in part at the annual meeting of the Eastern Psychological Association, Cambridge, Massachusetts, 12 March 2011.

References

Dolinsky, B. & Kelley, J.M. (2010). For better or worse: Using an objective programme assessment measure to enhance an undergraduate psychology programme. Teaching of Psychology, 37, 252–256.


About the Authors
Shawn P. Gallagher & Shaun P. Cook
Millersville University of Pennsylvania.

Correspondence
Shawn P. Gallagher
Department of Psychology,
145 Byerly Hall,
Millersville University,
P.O. Box 1002,
Millersville, PA 17551.
Email: sgallagher@millersville.edu
Appendix: Major Field test in Psychology sample questions (ETS, 2003).

1. Breland and Breland trained pigs to carry wooden coins in their mouths to a piggy bank. This sequence was reinforced with food. After some weeks, however, the trained pigs began to root the coins with their noses, treating them like pieces of food. This can be best characterised as an example of:
   (A) avoidance responding;
   (B) conditional responding;
   (C) superstitious behaviour;
   (D) instinctive drift;
   (E) delayed conditioning.

2. Which of the following therapeutic interventions places the most emphasis on gaining insight into early childhood relationships?
   (A) Systematic desensitisation.
   (B) Behaviour modification.
   (C) Family therapy.
   (D) Gestalt therapy.
   (E) Psychoanalysis.

Questions 3 and 4 are based on the following passage.
A psychologist investigated the developmental relationship between the average daily amount of television viewing and the reading skills of children. Parents of children in four age groups (6-year-olds, 7-year-olds, 8-year-olds, and 9-year-olds) were asked to record the number of hours their children watched television for a six-month period. The psychologist also gave the children reading-speed and reading-comprehension tests on a monthly basis for the six-month period. Analyses of the data reveal the following correlations:

<table>
<thead>
<tr>
<th>Age</th>
<th>Correlation between Hours of Television Viewing and Reading Speed</th>
<th>Correlation between Hours of Television Viewing and Reading Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>0.13</td>
<td>−0.32</td>
</tr>
<tr>
<td>7</td>
<td>−0.03</td>
<td>−0.38</td>
</tr>
<tr>
<td>8</td>
<td>0.07</td>
<td>−0.41</td>
</tr>
<tr>
<td>9</td>
<td>−0.05</td>
<td>−0.49</td>
</tr>
</tbody>
</table>

3. The pattern of results above suggests which of the following about television watching?
   (A) It increases reading comprehension but does not increase reading speed.
   (B) It has a stronger relationship to reading speed than to reading comprehension.
   (C) It has a stronger relationship to reading comprehension than to reading speed.
   (D) It diminishes the relationship between reading speed and reading comprehension.
   (E) It has a diminishing relationship to reading skills as the child grows older.
4. Based on the correlational data, the psychologist claims that television viewing significantly reduces reading skills. This claim can be justly criticised because:
(A) children younger than 10 years prefer television viewing to reading and the sample in the study is, therefore, biased;
(B) a cross-sectional study cannot provide information about longitudinal development effects;
(C) reading comprehension is more difficult to assess than is the amount of television viewing;
(D) television viewing need not impede the acquisition and utilisation of reading skills;
(E) correlational data do not justify inferences about causes.

5. A juror in a criminal case believes that the defendant’s illegal act cannot be explained or excused by extenuating circumstances. According to attribution theory, the juror is most probably:
(A) making stereotypical judgments;
(B) emphasising dispositional factors;
(C) discounting altruistic motivation;
(D) overestimating situational causes;
(E) relying on compliance to the law.

6. Lesions in Broca’s area of the association cortex are most likely to result in which of the following disorders?
(A) Expressive aphasia.
(B) Visual agnosia.
(C) Apraxia.
(D) Agraphia.
(E) Alexia.

7. A clinical psychologist is conducting a diagnostic interview with a client. Her impression that the client is suffering from paranoid schizophrenia would be supported by the presence of each of the following symptoms EXCEPT:
(A) poor contact with reality;
(B) systematised delusions;
(C) social withdrawal;
(D) panic attacks;
(E) loose associations.

8. In the course of learning their first language, young children may say ‘goed’ (for ‘went’) and ‘man’s (for ‘men’). These kinds of errors suggest that young children tend to:
(A) pay little attention to what they hear;
(B) overgeneralise the regularities that they hear in language;
(C) produce words that they do not comprehend;
(D) use correct grammatical forms only after they have developed a large vocabulary;
(E) imitate the grammatical errors of adults.
A Size–Distance Scaling Demonstration Based on the Holway–Boring Experiment

Shawn P. Gallagher¹ and Crystal L. Hoefling¹

Abstract
We explored size–distance scaling with a demonstration based on the classic Holway–Boring experiment. Undergraduate psychology majors estimated the sizes of two glowing paper circles under two conditions. In the first condition, the environment was dark and, with no depth cues available, participants ranked the circles according to their angular sizes. In the second condition, the environment was illuminated and, with depth cues available, the students ranked the circles according to actual physical size. The demonstration replicated the key elements of the original experiment, and objective and subjective measures indicated that it improved understanding of size–distance scaling. We also describe variants of the experiment suitable for different instructional environments.

Keywords
depth perception, size–distance scaling, size constancy, Holway and Boring

Among the most impressive feats of vision is the perception of depth and size. Textbooks and instructors commonly group these two processes because size perception can facilitate depth perception and vice versa. The size–distance scaling equation summarizes the relationship between the two functions:

\[ S = kRD, \]

where \( S \) is the perceived size of an object, \( R \) is the retinal image size, \( D \) is the perceived distance between the observer and object, and \( k \) is a scaling constant (Boring, 1940; Goldstein, 2010). The values \( R \) and \( D \) are inversely related and, as targets approach or retreat, the perceived size \( (S) \) remains constant.

Discussing the classic Holway–Boring (1941) experiment gives instructors an opportunity to deconstruct the size–distance scaling equation and demonstrate the importance of depth perception in estimating size. By removing depth cues, Holway and Boring (1941) concluded that the only remaining cue, the size of the retinal image, determined the perceived size. A modified version of the size–distance scaling equation shows the relationship between size perception and retinal image size in the absence of depth cues:

\[ S = kR. \]

With an observer positioned at the intersection of two hallways, Holway and Boring (1941) projected a luminous standard circle in one hallway, on an 8-ft × 8-ft screen positioned anywhere from 10 to 120 ft from the observer. In the other hallway, a second screen displayed a comparison circle 10 ft from the observer. The experimenters instructed observers to adjust the diameter of the comparison circle until the two circles were equal in size. Although the experimenters conducted their work at night, in a dark hallway, stray light illuminated surfaces of the corridor, providing “a sensory ground for the perception of the stimulus” (Holway & Boring, 1941, p. 30). Our efforts to recreate the original environment, in a hallway devoid of stray light, suggest that participants could also see the edges of the projection screens. With the features of the hallways visible, the original participants used depth cues like linear perspective to estimate the distance between themselves and the two stimuli. In this condition, observers easily matched the comparison stimulus to the true physical size of the standard stimulus, and estimates could be explained as a product of retinal image size and perceived distance, or \( S = kRD. \)

On subsequent trials, the experimenters proceeded to remove environmental depth cues and, as they did, the size estimates were no longer good indicators of physical size. First, the experimenters removed binocular depth cues by occluding one of the observer’s eyes. Then, they removed monocular depth cues by having the participants view the standard circle through an extendable occlusion tunnel that blocked all peripheral environmental depth cues, including the edges of the projection screen. In this final condition, observers were unable to judge distance to the standard stimulus and made poor estimates of

¹ Department of Psychology, Millersville University of Pennsylvania, Millersville, PA, USA

Corresponding Author:
Shawn P. Gallagher, Department of Psychology, 145 Byerly Hall, Millersville University, PO Box 1002, Millersville, PA 17551, USA.
Email: sgallagher@millersville.edu
physical size, but accurately matched the circles according to retinal image size, summarized as $S = kR$.

Although others have described effective ways to demonstrate size–distance scaling (Kunkel, 1993; Lumsden, 1976), we wanted to develop a method that captured elements of the original Holway–Boring (1941) experiment for students in sensation and perception. Like Holway and Boring (1941), we asked our participants to compare the sizes of two glow-in-the-dark circles. Instead of having participants manipulate and match the stimuli, and instead of progressively removing depth cues, we asked them to simply rank two circles according to size under two different conditions, with the room lights on and with the room lights off.

We assessed the pedagogical effectiveness of the demonstration with a short quiz administered before and after the exercise. We also hypothesized that subjective reports would support the objective findings. The value of the demonstration, of course, depends on how well it replicates the results of the original experiment (Holway & Boring, 1941, p. 30). Therefore, we also hypothesized that participants would rank the circles according to physical size when the lights were on and depth cues were available, as predicted by $S = kRD$ and that participants would rank the circles according to angular size, which is proportional to retinal image size, when the lights were off and depth cues were not available, as predicted by $S = kR$. We tested these hypotheses with a naive group of psychology majors, who had not studied the Holway–Boring (1941) experiment, as well as a group of sensation and perception students who knew our objectives.

**Method**

**Participants**

This experiment complied with the standards of Millersville University’s Institutional Review Board. Fifty-six naive psychology majors who had not studied the Holway–Boring (1941) experiment participated in exchange for course credit. A second group of 27 students from a sensation and perception class performed the experiment as part of a unit on depth perception and size–distance scaling.

**Materials**

We conducted the experiment in a university building that had several classrooms and offices on each floor. We placed the circles in a 12-ft × 8-ft office room and, although the room had no exterior windows, the door to the room had a window that made the interior visible from the hallway. We occluded this window with cardboard except for a 2-cm$^2$ aperture, 120 cm above the floor. It was small enough to prevent binocular viewing. The office room was not cleared or prepared in any particular way; desks, chairs, computers, and bookshelves were easily visible through the aperture and served as environmental depth cues when the room’s interior was illuminated.

We cut two circles from glow-in-the-dark paper and positioned them in the room, as shown in Table 1 and Figure 1. We taped each circle to a $1 bill and then taped the bill to the end of a 120-cm long dowel rod. A wooden base secured the bottom of each rod. The dollar bills served as a familiar size depth cue visible only in the illuminated condition and may have aided size estimations in our study in the same way that the two identical projector screens probably facilitated size judgments in the original study (Holway & Boring, 1941). The center of each circle was 120 cm above the floor and positioned to make the line of sight from the viewing aperture normal to each luminous face. When viewed through the aperture, the circles appeared side by side along the horizontal and when the room interior was not illuminated, the circles were the only visible objects in the room (see Figure 2). We chose sizes and locations that made the physical size of Circle A larger than that of Circle B and the angular size of Circle B larger than that of Circle A (see Table 1).

**Table 1. Features of the Stimulus Circles.**

<table>
<thead>
<tr>
<th>Circle</th>
<th>Distance to Aperture (cm)</th>
<th>Angular Size (Diameter in Centimeter)</th>
<th>Actual Size (Height in Degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>376</td>
<td>8.3</td>
<td>1.27</td>
</tr>
<tr>
<td>B</td>
<td>60</td>
<td>3.8</td>
<td>3.63</td>
</tr>
</tbody>
</table>

Note. Angular size in degrees = 57.29 (diameter of circle)/(distance to aperture).
Naïve students \((n = 56)\) participated in counterbalanced groups of six to eight. We told them that they were going to view and identify the larger of the two paper circles. We asked the participants to form a line in the hallway outside the room containing the circles and approximately 10 ft from the viewing aperture. We then gave each participant a black marker and a piece of paper with the letters A and B printed left to right in 100-point Arial font. We instructed them not to communicate with each other and monitored them during the experiment. One by one, we led them to the aperture where they could see two yellow circles. We told them that Circle A was on the left and Circle B was on the right and asked them to identify the larger circle by circling the corresponding letter on the paper. The experimenter then took the paper, handed the participant an identical sheet, and summoned the next participant. After all the group members had made their first estimates, we told the participants that we needed a moment to set up the next part of the experiment. One experimenter led the participants around a nearby corner where they could no longer see the door to the room, and the other experimenter changed the experimental condition by turning the room lights on or off. This delay was no more than 30 s, but it allowed the participants to entertain the possibility that the experimenter rearranged or changed the circles. Finally, the participants made estimates under the second condition.

The sensation and perception students participated one day after reading about the Holway–Boring (1941) experiment and taking a short pretest quiz (Appendix) made up of questions from the textbook’s question bank (Wurst, 2007). Because knowledge of the size–distance equations could bias these students, we thought that they would be likely to anticipate our intentions and, after seeing the true size and locations of the circles in the illuminated condition, would not be deceived by the absence of depth cues in the dark condition. We therefore tested all of these students in the dark condition first. An informal posttest survey confirmed that most assumed that the experimenters had not manipulated the circles between conditions. These students took the quiz again on the following day, at the beginning of the next class meeting.

**Results**

For the 25 sensation and perception students who completed the pretest and posttest quiz, posttest scores \((M = 4.04, SD = 0.98)\) were significantly higher than pretest scores \((M = 2.88, SD = 1.17)\), \(t(24) = 3.82, p = .001\) (two-tailed), \(d = 0.76\). Of the 26 responding sensation and perception students, 22 agreed that the demonstration improved their understanding of the Holway–Boring (1941) experiment by answering \(d\) or \(e\) on Item 6 of the posttest quiz. The five students who reported that the demonstration did not improve their understanding stated that they did so because the textbook description of the original experiment was sufficient.

As expected, most participants chose Circle A as the larger circle when the room was illuminated, but most chose Circle B when the room was dark and depth cues were not available. There was a significant relationship between viewing condition and circle choice for the counterbalanced naïve group, \(\chi^2(1, n = 56) = 45.56, p < .001, \varphi = 0.90\) (see Figure 3), as well as the sensation and perception students, \(\chi^2(1, n = 27) = 19.32, p < .001, \varphi = 0.85\).
Discussion

The pretest–posttest quiz results suggest that this demonstration enhanced understanding of the classic Holway–Boring (1941) experiment. History effects, of course, threaten the validity of any pretest–posttest investigation, but this simple, low-cost exercise successfully captured the key elements of the Holway–Boring experiment, yielded similar results, and subjective reports were positive.

Without depth cues, the size–distance scaling equation predicts that perceived size will be proportional to angular size which is, in turn, proportional to the size of the image on the retina. In the dark condition, the participants perceived only the angular sizes of the glowing circles and, as predicted, most chose Circle B as the larger of the two. In the illuminated condition, participants were able to use depth cues to estimate actual physical size, which is directly proportional to the product of retinal image size (R) and perceived distance (D), assuming accurate estimates of distance. As predicted, when participants could see the interior of the room, and especially the dollar bills taped to the backs of the circles, most were not misled by the fact that angular size was a poor indicator of physical size; the circle projecting the smallest retinal image was perceived as distant and correctly identified as the physically largest whereas the circle casting the largest retinal image was perceived as very close and correctly identified as the physically smallest.

Results from the sensation and perception class indicate that the size–distance scaling effect is still apparent in a biased sample. Interestingly, Holway and Boring (1941) were participants in their own study and also exhibited the effect, despite knowing the experimental objectives. Sensation and perception instructors may prefer to conduct this exercise before introducing size–distance scaling in order to collect an unbiased set of data and then present the results while discussing the Holway–Boring experiment, but we believe that the demonstration is a more effective educational exercise when presented in context. Our results indicate that the size–distance scaling effect is apparent in both naïve and informed samples.

This demonstration is simple and we estimate that, with planning, two experimenters can test groups of up to 100 students in less than an hour. We have found that comparative judgments are easier and data collection is faster with glowing squares that allow participants to compare the heights of their vertical edges. We used circles in an attempt to approximate as much of the original experiment as possible. If running the experiment as described here is still impossible, instructors can use a large-group simulation by showing students photographs taken from the perspective of the aperture in the illuminated condition and then simulating the dark condition with an image similar to the one at the bottom of Figure 2. The experiment can also be set up in a variety of windowless rooms. We have replicated the experiment in large computer labs and small storage closets.

The design presents many opportunities for elaborations, which could include asking the participants to estimate distance to the circles in order to determine if they perceive the circles as equidistant from the aperture in the dark condition and to see if and how depth estimates change with condition. Students could incorporate these estimates into the size–distance scaling equations and explain their perceptions mathematically.

Appendix

Assessment questions edited and drawn from an instructors manual (Wurst, 2007) for the course textbook (Goldstein, 2010). Correct answers are underlined. Item 6 appeared only on the posttest and was not used to assess understanding of size–distance scaling.

1. Holway and Boring found that
   a. size constancy holds under all viewing conditions.
   b. the law of visual angle does not work in humans.
   c. size constancy occurred if participants could see the features of the hallway.
   d. size constancy does not occur under binocular viewing conditions.
   e. more than one of the above.

2. An object’s angular size is determined by
   a. its physical size alone.
   b. its placement in the visual space.
   c. the distance between the object and the observer.
   d. the speed at which it is moving.
   e. more than one of the above.

3. In the absence of depth information, size estimates are based largely on
   a. size constancy.
   b. the actual physical size of an object.
   c. the physical distance of an object.
   d. the angular size of an object.
   e. more than one of the above.

4. The equation for size–distance scaling is $S = kRD$. The term D stands for
   a. perceived distance between observer and object.
   b. actual distance between observer and object.
   c. true physical diameter of the object.
   d. size of the distal stimulus.
   e. more than one of the above.

5. You step outside the building after class and spot your friend’s car as it leaves the parking lot. Which terms of the size–distance scaling equation, $S = kRD$, change as you watch the car pull away?
   a. S and R.
   b. S and D.
   c. R and D.
   d. only one term changes.
   e. all of the terms change.
6. Select the answer that describes how much you agree or disagree with the following statement: This demonstration improved my understanding of the Holway–Boring experiment.
   a. Strongly disagree
   b. Disagree somewhat
   c. Neutral
   d. Agree somewhat
   e. Strongly agree

Acknowledgments
We thank Bea Grosh and the anonymous reviewers for their helpful comments on earlier drafts of this article. We also thank Kate Irwin and Briana Feole for their assistance in data collection.

Declaration of Conflicting Interests
The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

References
Donor Risk Factors for Graft Failure in the Cornea Donor Study

Joel Sugar, MD,* Monty Montoya, MBA,† Mariya Dontchev, MPH,‡ Jean Paul Tanner, MPH,‡ Roy Beck, MD, PhD,‡ Robin Gal, MSPH,‡ Shawn Gallagher, PhD,§ Ronald Gaster, MD,† Ellen Heck, MS, MA,¶ Edward J. Holland, MD,** Craig Kollman, PhD,‡ Jackie Malling,†† Mark J. Mannis, MD,‡‡ and Jason Woody, §§ for the Group Cornea Donor Study Investigator Group

Purpose: The purpose of this study was to assess the relationship between donor factors and 5-year corneal graft survival in the Cornea Donor Study.

Methods: Donor corneas met criteria established by the Eye Bank Association of America, had an endothelial cell density of 2300 to 3300/mm², and were determined to be of good to excellent quality by the eye banks. Donor corneas were assigned using a random approach and surgeons were masked to information about the donor cornea including donor age. Surgery and postoperative care were performed according to the surgeons’ usual routines and subjects were followed for 5 years. Donor and donor cornea factors were evaluated for their association with graft failure, which was defined as a regraft or a cloudy cornea that was sufficiently opaque to compromise vision for a minimum of 3 consecutive months.

Results: Graft failure was not significantly associated with the type of tissue retrieval (enucleation versus in situ), processing factors, timing of use of the cornea, or characteristics of the donor or the donor cornea. Adjusting for donor age did not affect the results.

Conclusion: Donor and donor cornea characteristics do not impact graft survival rates for corneas comparable in quality to those used in this study.

Key Words: cornea donor, cornea transplant, graft failure (Cornea 2009;28:981–985)

INTRODUCTION

The Cornea Donor Study (CDS) recently published results demonstrating no difference in corneal graft survival at 5 years related to donor age.1 Although eye banks routinely gather information on both the donor and the quality of the corneal tissue before distributing tissue for transplantation, relatively little information is present in the literature concerning the impact of these factors on graft survival.2 In the CDS, information on cornea donors, tissue handling, and tissue findings was obtained and assessed to examine any potential impact on graft survival at 5 years. The results of these analyses are presented here.

MATERIALS AND METHODS

Study Protocol

Details of the CDS protocol have been reported previously1,3,4 and the key aspects are briefly summarized. The study protocol was approved by Institutional Review Boards for each eye bank and at each investigational site.

Eligible subjects were between 40 and 80 years old and had corneal disease associated with endothelial dysfunction and moderate risk of failure (principally Fuchs dystrophy and pseudophakic corneal edema). Written informed consent was obtained from each subject.

Eligible donor corneas met Eye Bank Association of America standards for human corneal transplantation.5,6 Eligibility criteria for the donor corneas assigned in the study are listed in Table 1. Eye banks obtained information about the cornea donor, including age, gender, race, history of diabetes, and cause of death from medical records, healthcare provider interviews, and family members. Type of tissue recovery, either as whole eye (enucleation) or corneoscleral rim removal (in situ), time from death to placement in preservative medium, body refrigeration time, and death to surgery time were recorded. Specific characteristics of the donor tissue were recorded, including epithelial slit lamp findings, stromal edema, arcus, folds in Descemet’s membrane, the presence of “snail tracks” (linear ruptures of endothelial cells), and endothelial morphology by specular microscopy.
Specular microscopy:
- Endothelial cell density 2300 to 3300 cells/mm²
- Polymorphism/polymegethism—none to no more than mild (slight)
- Guttae—no true guttae
- No evidence of central endothelial cell damage/trauma or dystrophy

Slit lamp examination criteria:
- Epithelium
  - Defects of 50% or less of epithelium
  - Haze—none to no more than moderate
  - Exposure—none to no more than moderate
- Stroma
  - Edema—none to no more than mild
  - Arcus—≥8.0-mm clear zone
- Descemet’s membrane
  - Folds—none to no more than few (mild)
- Endothelium
  - Snail tracks (endothelial stress lines)—none to no more than mild centrally
  - Guttae—no true guttae
- No evidence of central endothelial cell damage/trauma or dystrophy.

Clinical investigators and subjects were masked to all characteristics of the donor cornea, including age and endothelial cell density. Preoperative management, surgical technique, and postoperative care (including prescription of medications) were provided according to each investigator’s customary routine. The visit schedule during the first 6 postoperative months was at the discretion of the investigator. Thereafter, the minimum follow-up visit schedule included a visit between 6 and 12 months and then annual visits through 5 years. Because of the trial’s simple design, data collection at each visit was limited and included an assessment of graft clarity, signs of graft rejection, and intraocular pressure. The definition of graft failure, based on the definition used in the Collaborative Corneal Transplantation Studies,7,8 was a regraft or, in the absence of regraft, a cloudy cornea in which there was loss of central graft clarity sufficient to compromise vision for a minimum of 3 consecutive months. Further details of the classification scheme for graft failures have been published.1

Statistical Methods

The analysis included the 1090 eligible subjects in the CDS. Baseline endothelial cell density was evaluated in the Collaborative Corneal Transplantation Studies.7,8 was a regraft and 33 (24%) met the cloudy cornea failure criteria defined for the study without a regraft (30 had a cloudy cornea for at least 3 months and 3 had a cloudy cornea for less than 3 months without additional available follow up). Three graft failures were the result of primary donor failure, 8 uncorrectable refractive error, 48 graft rejection, 46 endothelial decompensation, and 30 other causes.

As shown in Table 2, graft failure rates were not significantly impacted by any donor characteristics (gender, history of diabetes, or cause of death); by any factors related to the type of tissue retrieval, processing, timing of use of the cornea (time from death to preservation or time from death to surgery); or by any characteristics of the donor cornea (presence of endothelial polymorphism, endothelial cell damage, Descemet folds, snail tracks, baseline endothelial cell density, epithelial defects, epithelial haze, epithelial exposure, stromal edema, or arcus). Adjusting for donor age did not affect the results (data not shown). When analyses were conducted separately for rejection and nonrejection graft failures, no baseline donor factors were found to be associated with the rate of graft failure based on our prespecified level of significance (data not shown).

DISCUSSION

Although a number of factors evaluated here have been assessed in other studies, this prospective study is one of the few to address their impact on graft survival. The limitation of this study is that the tissue selection criteria excluded extremes such as prolonged death to preservation time and death to surgery time. Only mild to moderate epithelial, stromal, Descemet, and endothelial variations were accepted. Nonetheless, the information obtained is useful in demonstrating the lack of any adverse impact of the abnormalities found in the ranges studied.

Data on donor cause of death and presence or absence of diabetes mellitus show no impact on 5-year graft outcomes. Earlier studies have shown a similar lack of impact of donor cause of death on 5-year graft survival.2 Recently, deaths resulting from cancer have been implicated in postoperative endophthalmitis,9 but there were no cases of endophthalmitis attributable to the donor cornea in the CDS. Additional studies have shown no reason to exclude donors with cancer,10 but these did not look at graft survival.
Graft outcome did not differ by method of tissue procurement (enucleation versus in situ retrieval). Rootman and coinvestigators reported no difference in initial donor tissue quality rating by procurement methodology, but they did not look at graft outcomes beyond primary graft failure. A recent study showed no difference in graft clarity at 3 months with either procurement method.

Timing of tissue procurement, refrigeration, and use has been studied in the past and has also been shown, within limited ranges, to have no effect on graft outcome, although prolonged storage times, not studied here, may well have a deleterious effect. Endothelial characteristics likewise had no impact on graft outcomes using the donor criteria of the CDS.

### TABLE 2. Baseline Donors Factors Predictive of Graft Failure (N = 1090)

<table>
<thead>
<tr>
<th>Donor Factors</th>
<th>5-yr Graft Failure ± 99% CI^</th>
<th>Hazard Ratio (99% confidence interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>0</td>
</tr>
<tr>
<td>Overall</td>
<td>1090</td>
<td></td>
</tr>
<tr>
<td>Donor characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>716</td>
<td>13% ± 3%</td>
</tr>
<tr>
<td>Female</td>
<td>374</td>
<td>15% ± 5%</td>
</tr>
<tr>
<td>History of diabetes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>891</td>
<td>14% ± 3%</td>
</tr>
<tr>
<td>Yes</td>
<td>199</td>
<td>12% ± 6%</td>
</tr>
<tr>
<td>Cause of death</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardio/Stroke</td>
<td>659</td>
<td>14% ± 4%</td>
</tr>
<tr>
<td>Cancer</td>
<td>207</td>
<td>18% ± 7%</td>
</tr>
<tr>
<td>Trauma</td>
<td>96</td>
<td>11% ± 9%</td>
</tr>
<tr>
<td>Respiratory</td>
<td>78</td>
<td>11% ± 9%</td>
</tr>
<tr>
<td>Other†</td>
<td>50</td>
<td>9% ± 12%</td>
</tr>
<tr>
<td>Retrieval and timing factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Tissue Retrieval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enuceleation</td>
<td>218</td>
<td>13% ± 6%</td>
</tr>
<tr>
<td>In situ</td>
<td>872</td>
<td>14% ± 3%</td>
</tr>
<tr>
<td>Time from Death to Preservation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–4 hours</td>
<td>206</td>
<td>17% ± 7%</td>
</tr>
<tr>
<td>&gt;4–8 hours</td>
<td>577</td>
<td>13% ± 4%</td>
</tr>
<tr>
<td>&gt;8–10 hours</td>
<td>165</td>
<td>12% ± 7%</td>
</tr>
<tr>
<td>&gt;10 hours</td>
<td>142</td>
<td>18% ± 9%</td>
</tr>
<tr>
<td>Body refrigerated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>255</td>
<td>15% ± 6%</td>
</tr>
<tr>
<td>Yes</td>
<td>835</td>
<td>13% ± 3%</td>
</tr>
<tr>
<td>Time from death to preservation by refrigeration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refrigerated 0–4 hours</td>
<td>107</td>
<td>16% ± 10%</td>
</tr>
<tr>
<td>Refrigerated &gt;4–8 hours</td>
<td>433</td>
<td>12% ± 4%</td>
</tr>
<tr>
<td>Refrigerated &gt;8 hours</td>
<td>295</td>
<td>15% ± 6%</td>
</tr>
<tr>
<td>Not Refrigerated 0–4 hours</td>
<td>99</td>
<td>17% ± 10%</td>
</tr>
<tr>
<td>Not Refrigerated &gt;4 hours</td>
<td>156</td>
<td>14% ± 8%</td>
</tr>
</tbody>
</table>
| Timing of tissue procurement, refrigeration, and use has been studied in the past and has also been shown, within limited ranges, to have no effect on graft outcome, although prolonged storage times, not studied here, may well have a deleterious effect. Endothelial characteristics likewise had no impact on graft outcomes using the donor criteria of the CDS.
### Table 2 continued

<table>
<thead>
<tr>
<th>Donor Factors</th>
<th>N</th>
<th>5-yr Graft Failure a</th>
<th>99% CI</th>
<th>Hazard Ratio (99% confidence interval)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donor cornea characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline endothelial cell density</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000 ≤ 2500</td>
<td>318</td>
<td>14% ± 5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2500 ≤ 2700</td>
<td>270</td>
<td>17% ± 6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2700 ≤ 3000</td>
<td>361</td>
<td>13% ± 5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥3000</td>
<td>141</td>
<td>9% ± 7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polymorphism/polymegathism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>951</td>
<td>13% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild/moderate</td>
<td>139</td>
<td>16% ± 9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endothelial cell damage – diffuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1054</td>
<td>14% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild/moderate</td>
<td>36</td>
<td>9% ± 13%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endothelial cell damage – peripheral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1038</td>
<td>14% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>52</td>
<td>8% ± 10%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epithelium – defects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clear and intact</td>
<td>525</td>
<td>12% ± 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defects on &lt;50% of epithelium</td>
<td>565</td>
<td>15% ± 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epithelium – haze</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>733</td>
<td>12% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>315</td>
<td>16% ± 6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate/severe</td>
<td>42</td>
<td>19% ± 17%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epithelium – exposure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>209</td>
<td>12% ± 6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild</td>
<td>661</td>
<td>16% ± 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate/severe</td>
<td>220</td>
<td>10% ± 5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stromal edema</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>817</td>
<td>14% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild/moderate</td>
<td>273</td>
<td>14% ± 6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcus present</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>380</td>
<td>17% ± 5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Present</td>
<td>710</td>
<td>12% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Descemet Folds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>376</td>
<td>13% ± 5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Few (mild/moderate</td>
<td>714</td>
<td>14% ± 4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skinal tracks – central</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1011</td>
<td>13% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild/moderate</td>
<td>79</td>
<td>19% ± 12%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skinal Tracks – diffuse</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>903</td>
<td>13% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild/moderate</td>
<td>187</td>
<td>17% ± 8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skinal tracks – peripheral</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>880</td>
<td>15% ± 3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mild/moderate/severe</td>
<td>210</td>
<td>8% ± 5%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a - CI = confidence interval
b - ○ signifies an upper 99% confidence limit > 3.00 (for moderate/severe epithelium – haze group the upper 99% confidence limit = 4.03)
c - 12 anoxia/asphyxiation, 10 renal, 6 hepatic, 4 neurological, 4 drug overdose, 3 gastrointestinal, 3 shock, 2 pancreatitis, 2 seizures, 2 undetermined, 1 adrenal insufficiency, 1 electrocution

d - Includes 17 subjects who received corneas where time from death to transplant longer than five days (15 were six days, 1 was seven days and 1 was eight days)
e - 658 values from the reading center and 432 from the eye bank. Results were similar when excluding the cases without a Reading Center value (data not shown).

f - Only one subject classified as moderate
g - Only one subject classified as severe
Epithelial and stromal changes were generally mild with only one case of severe epithelial exposure and one with moderate stromal edema. Although over half of the donor corneas had epithelial defects, these all involved less than 50% of the epithelium and had no statistically significant effect. Because of the relatively low prevalence of endothelial trauma related to tissue preparation, as manifested by snail tracks in the central cornea and present in only 7% of donor corneas, the impact of this type of trauma cannot be fully assessed.

Although study of more extreme alterations of tissue would be of benefit, this study demonstrates effectively, as have the other reports from the CDS, that all tissue meeting the donor criteria used performs equally well. Continued follow up of this cohort through 10 years is ongoing to assess any potential differences in longer-term survival.

REFERENCES

APPENDIX
A listing of the Cornea Donor Study Investigator Group, including clinical site investigators, eye bank staff, coordinating center staff, specular microscopy reading center staff, and committees, has been previously published online.1

The following CDS Publications Committee members independently reviewed and approved this article for submission: John Affeldt, MD, Michael W. Belin, MD, Terry E. Burris, MD, Richard Eifermann, MD, and Jonathan Macy, MD.
Human Science Fictionalized: A Novel, a Visual Narrative, and an Indie Film

Symposium at the Annual Meeting of the History of Science Society
In Atlanta, Georgia on November 6, 2016

In studies of science popularization, the focus is usually on nonfiction. But what about fictionalized portraits of science? This symposium looks at three attempts to bring the human and neuro sciences to the public through fiction. Among the questions explored are: How is the fact/fiction boundary negotiated? How do a “fact writer” and a “fiction writer” think about popularization differently? What are the different relationships that they have to their sources, or that they envision with their audiences? Our presenters: Andromeda Romano-Lax is a successful novelist whose most recent work, Behave (2016), dramatizes the life and career of Rosalie Rayner, wife and former student of behaviorist John Watson. Matteo Farinella is an illustrator and artist with a doctorate in neuroscience. His visual narrative, Neurocomic (2013, coauthored with Hana Roz), portrays the history of neuroscience through a young man’s voyage of discovery in a land of giant neurons and encounters with famous scientists. Gina Perry is an Australian journalist who used her investigative and narrative skills to write a Behind the Shock Machine (2013), a history of Stanley Milgram’s obedience studies. Now a doctoral student in psychology, she will review Experimenter, Michael Almereyda’s 2016 film about Milgram and his work. Our commentator is Nadine Weidman, a historian of science at Harvard University known for her work on public controversy and popularization in the 20th-century human sciences. Our Chair is John Carson, a historian at the University of Michigan and Director of Undergraduate Studies for its Program in Science, Technology, and Society. Organizer: Ben Harris, University of New Hampshire.

Developing Resources

Cajal Institute Seeks UNESCO Recognition to Establish a Museum for the Anatomist’s Legacy

Shawn P. Gallagher, Millersville University of Pennsylvania

Spain was on the periphery of European science and medicine in 1887 when Santiago Ramón y Cajal saw a demonstration of Camillo Golgi’s method for staining individual nerve cells. He was not necessarily a neuroanatomist at the time, but he was a skilled histologist with a passion for illustration. The Golgi technique revealed the beauty and diversity of neurons and Cajal found the forms—and their implications—irresistible. Within 2 years, he had a collection of slides and meticulous illustrations that allowed him to cross language barriers in a way that few scientific discoveries could. The images told stories of neural development, homology, and, of course, solidified the neuron doctrine in the minds of most of his peers. Cajal and Golgi shared the 1906 Nobel Prize in Physiology/Medicine and, in 1920, King Alfonso XIII of Spain signed a Royal Decree establishing the Cajal Institute, a research facility focused on Biology, Experimental Physiology, and Histology.

When he died in 1934, Cajal left the institution with his collection of laboratory instruments and over 4,000 photographs and illustrations: a treasure trove that, like no other, shows how an individual with the perspective of artist and scientist used form to illuminate function. For more than eight decades, the Cajal institute has carefully stored and maintained these materials, but, as a research institution, it does not have the resources or the mandate to create an exhibit or make the materials easily accessible to researchers. Therefore, the institute aims to have the Cajal Legacy recognized as a UNESCO (United Nations Educational, Scientific and Cultural Organization) World Heritage treasure. Such recognition would facilitate the establishment of a museum to display the archives and make them, along with items like notebooks and correspondence, readily available for research and educational programs. The institute asks its international colleagues to support its petition. More information can be found at http://www.cajal.csic.es/ingles/legado.html.

Correspondence should be addressed to Shawn P. Gallagher, Department of Psychology, 111 Byerly Hall, Millersville University, P.O. Box 1002 (1 S. George Street), Millersville, PA 17551. E-mail: Shawn.Gallagher@Millersville.edu

1 The author would like to offer congratulations to María Ángeles Langa as she completes more than 45 years of dedicated service to the Cajal Legacy.

Review by: Shawn P. Gallagher, Psychology, Millersville University of Pennsylvania

In the preface to *Forty Studies That Changed Psychology*, Roger R. Hock states that his goal is to “fill the gap between all those psychology textbooks and the research that made them possible.” Hock’s narrative arc is familiar and common to many introductory psychology textbooks, beginning with the biological basis of behavior and ending with social and abnormal psychology. He has organized psychology into ten broad categories and surveyed each one with four readings that summarize specific published reports. Hock’s intended audience is “anyone, in any course, who wishes a greater understanding of the true roots of psychology.”

Before delving into the studies, Hock takes care to explain his strategy for presenting each one, the importance of understanding scientific methodology and, quite presciently, the ethical guidelines of the American Psychological Association. From the outset, he admirably encourages his readers to reflect on the costs and benefits of the studies they are about to explore and consider how each one shaped our current understanding of psychology.

This book is as much an introduction to the scientific method as it is to the discipline of psychology. Each reading has the basic structure of a journal article (Theoretical Propositions, Method, Results, and Discussion) and, together, they familiarize the reader with the nature of printed scientific communication. Hock concludes each reading by putting the study in historical context and addressing subsequent criticisms and applications.

Chapter one presents four readings on the biological basis of behavior and the first reading describes the “split brain” work of Michael Gazzaniga and Roger W. Sperry. The reading presents foundational topics like basic neuroanatomy and perception and describes scientists using a clear, hypothesis-driven approach to explore the consequences of bisecting the corpus callosum. While a great place to start, it is important to note that this reading, like six others, is a synopsis of a report published in *Scientific American*, a popular science
magazine written for a general, educated audience. As such, Hock’s report is a review of a review and only slightly more accessible than the original.

Other readings in chapter one address the works of Mark Rosenzweig and Eleanor Gibson, in an effort to give readers an appreciation for how psychology has addressed nature-versus-nurture questions; again, these are common and appropriate first topics for an introductory course.

Chapters two and three address consciousness and learning and, like the readings in chapter one, describe controlled scientific studies. Watson’s work with “Little Albert,” an orphaned nine-month-old, while not the most scientific study reviewed, gives Hock a suitable platform for discussing the ethical treatment of children in psychological research. The next three chapters present studies in cognition, development, and motivation. A review of Edward C. Tolman’s work clearly explains how behavioral studies contributed to the foundation of modern cognitive psychology and invites students to consider how psychologists infer cognition from behavior. A description of Elizabeth Loftus’ work explains how psychology has informed the criminal justice system, and the reading on Masters and Johnson is appropriately framed within the sexual revolution. In most cases, the historical significance is obvious and fosters an appreciation for the interdisciplinary impact of psychological research. The fact that Masters and Johnson are the featured characters in a television series and that Stanley Milgram and Philip Zimbardo are about to become featured characters in major motion pictures validates not only the disciplinary but also the cultural significance of the inclusion of their studies.

Forty Studies That Changed Psychology does not have the substance of a core text but it can support primary course material and offer students the opportunity to explore the research mentioned in a primary textbook. No survey is perfectly comprehensive and instructors could argue incessantly over what should or should not be included in such a review. Rather than challenging the global merit of Hock’s selections, however, I would question how well a few of the readings fit his specific goal of providing a “glimpse into the birth and growth of the science of psychology.” For example, Anna Freud’s paper on defense mechanisms is little more than a theoretical proposition; there is no methodology to explain her theoretical proposition, and the summary of her position feels like an obligatory, albeit indirect, nod to her father, Sigmund Freud, who is otherwise overlooked. Similarly, the readings on the Rorschach and Thematic Apperception Tests are also mostly theoretical statements that offer students little science to discuss. Hock’s plan to present “a detailed account of the experimental design and methods used to carry out the research” is compromised by an attempt to cover the breadth of the discipline. These three readings describe ideas, not studies, and break the rhythm of a useful and otherwise consistent pedagogical process. They serve, somewhat paradoxically, as illustrations of how some of psychology’s biggest ideas were not scientifically derived.

Hock’s claim that this text is for “anyone, in any course” undermines his objectives and risks trivializing his accomplishment. In my professional experience, this book (like its previous U.S. editions) best serves students at the extremes of their undergraduate careers. As stated, the book is organized as though it was intended to complement an introductory psychology textbook and it reads like a friendly invitation to the field. By describing the methodology behind the big ideas, Hock takes novice students into the realm of psychological research, topic by topic. The readings can be used to bridge the gap between brief descriptions in an introductory textbook and the original, primary source reports which all psychology students should eventually explore. That said, for the most advanced students,
the text is a concise review that can remind them of material they learned months or years before. I recommend *Forty Studies That Changed Psychology* for students who are preparing for the U.S.-based Graduate Record Exam (GRE) subject test for Psychology.

Finally, I would note that only one reviewed study, an fMRI investigation by Philip Ross, was conducted in the past 25 years and this might give a novice the impression that the discipline is stagnant. Of course, the true impact of a study cannot be measured immediately but the field of neuroscience, alone, is replete with recent discipline-changing discoveries like the place cell work that won May-Britt Moser and Edvard I. Moser the 2014 Nobel Prize.

This is a useful and enjoyable text and, for more than twelve years, I have always had a copy of *Forty Studies That Changed Psychology* within easy reach, but the current text could be improved with a consistent scientific perspective and the inclusion of just a few more recent, revolutionary discoveries. Hopefully, that will change in subsequent editions.
hegemonic cultural imperialism inherent in the spread of English and its impact on global culture.

*World Englishes: The Study of New Linguistic Varieties* offers a brief yet thorough introduction to the diverse and fascinating field of English language acquisition and development. Themes such as morphology, syntactic theory, structural elements, and discourse particles are analyzed from a variety of different sociocultural perspectives giving the reader a comprehensive guide to sociolinguistics that is at once accessible for novices, yet thorough enough for experts looking for a brief handbook of English language acquisition.

**Bio**

Mary Alice Adams (PhD, University of Alabama, 2006) is an assistant professor in the Department of Speech at Louisiana Tech University, Ruston. Her research interests include the social effects of the media, visual rhetoric, and culture studies.


**Reviewed by:** Shawn P. Gallagher  
Millersville University, Millersville, PA, USA  
DOI: 10.1177/0261927X10397155

Today, it is widely accepted that since human language is a characteristic of a living organism, it must be studied and understood within the framework of evolution. Decades ago, however, linguistics and biology were on separate and, only occasionally, parallel paths, and interdisciplinary collaborations were rare. Recent advances within these and related disciplines have demonstrated that the fields must converge but, as a matter of course, the reconciliation of a traditional linguistic theory and evolutionary theory will be difficult and, at times, impossible. Although the title of Anna Kinsella’s book, *Language Evolution and Syntactic Theory*, suggests a broad review, it is in fact a very specific argument for why Noam Chomsky’s (1995) Minimalist Program (MP), which is an extension of Generative Grammar, is incompatible with evolutionary theory.

In a review article published in *Science*, Hauser, Chomsky, and Fitch (2002) outlined an approach for studying language evolution within a minimalist framework. The article offered a strategy for identifying the broad language faculties (FLB), the ones that are shared by humans and other species, and the narrow language faculties (FLN), those that are unique to humans. Most important, this approach assumes recursion, the ability to take a finite set of ideas and link or embed them to form an infinite array of thoughts, phrases, or expressions, is the only unique component of FLN. This
hypothesis, Kinsella argues, is not based on evidence but, instead, on an unscientific attempt to bridge the gap between established assumptions of MP and the evolutionary theory that has served biology and cognitive psychology so well. Although she begins her critique by reviewing other elements of minimalist theory, Kinsella’s points converge on a case against the recursion-only hypothesis. With this, she argues that MP and evolutionary theory cannot be reconciled.

Kinsella’s book opens with a brief review of both syntactic theory and evolutionary theory. She reminds us that evolution is not a goal-oriented process and that Chomsky’s (1995) question, “How perfect is language?” might lead some to assume that language, unlike any other product of natural selection, must be so. Kinsella claims that the MP literature confuses the terms perfect and optimal and, out of necessity, carefully describes the differences between a perfect system and an optimal one. No natural system or structure is perfect, and redundancies, vestiges, and imperfections abound. She explains how natural selection drifts toward optimal solutions, not perfect solutions, given a set of constraints. The MP fails to explain how language would reach perfection in this manner.

The search for perfectionism in MP drives the assumption that human language is efficient and must have a simple mechanism at its core. It also assumes that FLN arrived as a single saltation, or leap, that gave humans a unique ability to combine and use existing elements of FLB. This leaves little room for progressive adaptation or the existence of a useful, rudimentary form of human language. Among other criticisms, Kinsella argues that this assumption brings MP face-to-face with genetics. She proposes that evidence of a single, critical mutation, an FLN gene, would help align MP with evolutionary theory. However, current research points to several genes underpinning human language, suggesting a trail of adaptive steps rather than a giant leap.

Kinsella concludes her argument about the evolvability of language by citing evidence that language is modular and that modularity gives language a robustness, an ability to withstand damage, that is consistent with an adaptable, evolving system. A unitary, nonmodular, FLN fails to account for the acquisition and use of sign language and the effectiveness of communication when signals are noisy or incomplete.

In their original article, Hauser, Chomsky, and Fitch (2002) stated, “We hypothesize that FLN only includes recursion and is the only uniquely human component of the faculty of language” (p. 1569). Kinsella and others took the authors at their word and started a very fruitful debate. Specific criticism from Pinker and Jackendoff (2005) prompted a lengthy response from the original authors (Fitch, Hauser, & Chomsky, 2005) in which, among other clarifications, they assert “We do not define FLN as recursion by theoretical fiat,” they acknowledge misrepresenting the results of another group of investigators, and they repeatedly attribute much of the debate to misunderstandings. Kinsella might argue that misunderstandings are to be expected when elements of a theory, such as “perfection” are notoriously difficult to pin down. She kindly refers to this article as a rejoinder (p. 158) when it carries many hallmarks of a theory in crisis.
There are many points throughout the book of which specific examples from the literature would have strengthened Kinsella’s case but, unlike Pinker and Jackendoff’s (2005) point-by-point critique of Hauser et al. (2002), her book is a “recipe” (p. xi) for developing evolutionarily plausible syntactic theories. With this guide, she provides a thorough argument for why MP is impossible to reconcile with evolutionary theory. Many supporters would be quick to claim that MP is a process, not a theory, and that it should not necessarily be abandoned just because it fails to explain a few bits of contradictory data. In the case of MP, however, the contradictory data seem to be anything that would make language look like an evolved biological system. Kinsella considers such selective disregard unacceptable. If any syntactic theory is to survive in a useful form, its proponents bear the burden of explaining how it could have emerged beyond the known constraints of evolution.

For many linguists, MP holds the hope of a unified theory of language, and such high hopes are not easily abandoned. Kinsella is aware of this and, as she explains why the empirical case is closed, she warns that an emotional one remains. She is most forceful as she concludes Chapter 4, the chapter dedicated to recursion, and states that it is “a red herring, blinding our view” (p. 159) and warns of the confusion that can arise when theories and subsequent hypotheses are based on intuitions and vague allusions. Her point is well taken, but red herrings often inspire others to expose them as such which, in turn, pushes research into different and more promising avenues. She closes with a firm nudge by warning us that, if the ultimate goal is a complete theory of language, linguists cannot afford to ignore pertinent facts from overlapping fields.

References

Bio
Shawn P. Gallagher (PhD, University of Delaware, 2002) is an associate professor in the Department of Psychology at Millersville University of Pennsylvania. His research interests include the impact of eye disease on visual perception. He teaches courses in perception, cognitive science, and evolutionary psychology.