AJUR

American Journal of
Undergraduate Research

Volume 14 | Issue 2 | June 2017

www.ajuronline.org

Print Edition ISSN 1536-4585
Online Edition ISSN 2375-8732

AJUR
American Journal of
Undergraduate Research

Volume 14 | Issue 2 | June 2017 *www.ajuronline.org*

American Journal of Undergraduate Research (AJUR) is a national, peer-reviewed, open-source, quarterly, multidisciplinary student research journal. It is indexed internationally by EBSCO, and is listed via the Library of Congress under ISSNs of 1536-4585 (for print) and 2375-8732 (for web). The journal was established in 2002.

SPECIAL THANKS

AJUR is made possible through the assistance of our sponsors.

Support for this issue has been provided by Wiley
as well as the Office of the Provost at
the State University of New York at Oswego. Thank you!

WILEY

OSWEGO
STATE UNIVERSITY OF NEW YORK

Interested in supporting quality undergraduate research?
Request sponsorship information at *editor@ajuronline.org*

Blackseed (*Nigella sativa*) Oil and its Active Ingredient, Thymoquinone, Suppress the Aggressive Phenotype of Breast Cancer Cells

*Sabrina Chaudhry, Safia Siddiqui, Tyrnnon K. Steffen, Stacey L. Raimondi**

Department of Biology, Elmhurst College

Students: sabrina.chaudhry@365.elmhurst.edu, safia.siddiqui@365.elmhurst.edu, tyrnnon.k.steffen@365.elmhurst.edu
*Mentor: *raimondis@elmhurst.edu*

ABSTRACT
Breast cancer is a leading cause of cancer deaths in women within the United States. However, current treatment methods for the disease present deleterious side effects themselves. Therefore, there is a move towards finding natural cures in order to mitigate negative side effects while still providing effective treatment for the cancer. Blackseed (*Nigella sativa*) oil is one particular natural remedy, alongside its active ingredient thymoquinone (TQ), which has been successfully tested for suppressing certain types of breast cancer cell proliferation. TQ itself has been seen to be capable of preventing proliferation of both non-aggressive MCF-7 and highly aggressive MDA-MB-231 cancer cells. However, studies which looked at the effects of TQ on MCF-7 cells alone were limiting in their use of high concentrations of the chemical without emphasis on finding a minimum effective dosage. Additionally, a second study which tested the effects of TQ on both MCF-7 and MDA-MB-231 cell lines conducted the experiments in the presence of a lipid-carrier molecule. This, in turn, may have served as a confounding variable in the results. Therefore, it was hypothesized that a minimal effective dosage for both blackseed oil and TQ could be determined, where a significantly greater suppression of MDA-MB-231, in comparison to MCF-7, cell proliferation would be observed. Cell proliferation, cell adhesion, and soft agar assays were used to test the hypothesis of this study. The minimum effective dosage for each substance, characterized by proliferation of the non-aggressive MCF-7 cells to some extent and suppression of the aggressive MDA-MB-231 cells, were determined to be 0.5 μL for blackseed oil and 1.0 μM for TQ. Additionally, TQ's effectiveness was noted to be more time-dependent than blackseed oil. This study supports the use of minimal effective doses for blackseed oil or TQ to naturally treat breast cancer while preventing damage to non-aggressive cells.

KEYWORDS
Breast cancer; Blackseed oil; *Nigella sativa*; Thymoquinone; Effective dose; Natural remedies

INTRODUCTION

Breast cancer currently stands as a leading cause of cancer-related death in women within the United States, second only to lung cancer, and has been estimated to cause 40,890 deaths in the year 2016. In addition, an estimated 249,260 new breast cancer cases are expected to be seen this year.[1] Aside from skin cancers, breast cancer is also the leading form of cancer in US women, responsible for approximately 30% of all new cancer diagnoses.[2] Two forms of breast cancer exist, based on differences in tumor location: invasive and non-invasive. Non-invasive forms of breast cancer are characterized by abnormal cell growth that does not extend beyond the originating layer of cells. Invasive breast cancer cells, however, are those that proliferate beyond the walls of glands or ducts in which they originate, and subsequently invade nearby breast tissue.[2] In the case of non-invasive breast cancers, the 5-year survival rate for females reaches as high as 99%, whereas that for invasive cancers that have infiltrated distant tissues of the body is as low as 26%.[2] These statistics speak to the importance of all attempts to find a cure to breast cancer in general, but more pressingly, to find a means of suppressing aggressive breast cancer cell growth.

Recent studies have found numerous negative side-effects associated with common chemotherapeutic agents used in treating breast cancer. These include such effects as ototoxicity, cardiotoxicity, infertility, premature menopause, osteoporosis, and other similarly deleterious ailments.[3,4] As recent studies expound upon the many negative side effects of current drugs used in treating breast cancer, a shift toward more natural remedies for cancer treatment has been seen. One natural remedy that has recently been making a comeback in the field of medicine for treating various cancers is *Nigella sativa* oil, also known as blackseed or black cumin seed oil. Having been used mostly in areas of the Asian continent for decades, blackseed oil has

been hailed for its position as a medical cure-all.[5]

In recent research, blackseed oil has been seen to suppress cancer cell proliferation and/or viability in a number of cancers, including cancer of the colon, lungs, pancreas and several other cancers.[6–9] Thymoquinone (TQ) has been noted to be the major chemical component of blackseed oil, with thymol, dithymoquinone, and various tocopherols accounting for the other major chemical components of the oil.[10] In particular, TQ is known to have important antioxidant and antimicrobial properties that are thought to contribute to the many health benefits of blackseed oil.[5] Intriguingly, TQ alone has even been seen to display cancer-fighting ability against cancers of the larynx, colorectal area, brain, breast, oral lining, blood, and a number of other cancers.[9,11–17] In the study testing the effects of TQ on the non-aggressive MCF-7 breast cancer line, however, it was noted that very high concentrations of TQ were used as treatments (ranging from 0 to 80 μM, at 10 μM increments).[14] This leads to a lack of knowledge regarding minimal effective dosage of the chemical. In another study using breast cancer cell lines, while both the non-aggressive MCF-7 and highly aggressive MDA-MB-231 cells were tested, these tests were run with the presence of a lipid carrier molecule, which may have affected the results of a study carried out without said component.[15]

Given that very high doses of TQ are capable of causing death of even non-aggressive MCF-7 cells, while the effects of TQ on MDA-MB-231 cells has not, to our knowledge, been tested in the absence of a lipid carrier molecule, it would prove valuable to find the lowest effective doses of TQ and blackseed oil capable of preventing MDA-MB-231 cell proliferation while exhibiting a lesser effect on non-aggressive MCF-7 cells.[14] Thus, this study will serve to provide data regarding the potential effects of blackseed oil and its active ingredient, TQ, on both MCF-7 and MDA-MB-231 breast cancer cells. Considering the previous literature, it is hypothesized that both blackseed oil and TQ will suppress the growth of both MCF-7 and MDA-MB-231 breast cancer cells at high doses, whereas a low dose of each compound will preferentially suppress growth of the aggressive MDA-MB-231 cell line compared to the non-aggressive MCF-7 cell line.

MATERIALS AND METHODS

Cells and culture conditions

MCF-7 and MDA-MB-231 breast cancer cell lines were obtained from ATCC. These cell lines were chosen because they are some of the most commonly used breast cancer cell lines available and were used in previous research so that we can compare our results to others. Furthermore, they show two different sides of breast cancer – a non-aggressive/non-invasive cell line, MCF-7, which maintains virtually all of the characteristics of normal breast epithelia with regard to hormone receptor status, cell morphology, and growth patterns compared to an invasive/metastatic cell line, MDA-MB-231, that has lost its epithelial cell patterns including hormone receptor status, cell morphology, and growth patterns. All cells were cultured at 37°C and 5% CO_2 in complete media containing RPMI 1640 (Life Technologies, Grand Island, NY), 10% FBS (Life Technologies), and 1% Penicillin-Streptomycin (Sigma-Aldrich, St. Louis, MO). Blackseed oil was obtained commercially and directly added to cell cultures in the amounts indicated. Thymoquinone (Sigma-Aldrich) was diluted in DMSO (Sigma-Aldrich) to create a 1 mM stock solution and added to cell cultures in the concentrations indicated. In order to determine minimal effective doses of blackseed oil and TQ to be utilized in the following assays, previous research was used to determine a suitable range for study.[7-9, 11, 12, 14-17] Following this, comparative dose-response experiments for each solution were run to find respective minimal effective concentrations. These were defined by the two lowest solution concentrations at which a majority of cancer cell death was observed, without completely eliminating all cells. For blackseed oil, these values were at 0.5 μL or 1.0 μL of blackseed oil, while for TQ the values were found to be 1.0 μM or 5.0 μM of thymoquinone.

In vitro cell proliferation assay

Cells were plated in a 6-well dish at a density of 1×10^5 cells, in 2 mL of media, per well. All cells were counted and re-plated twice a week for one week. Treatment groups were given 0.5 μL (0.5 μL/2 mL) or 1.0 μL (1.0 μL/2 mL) of blackseed oil, or 1.0 μM or 5.0 μM of thymoquinone with little to no separation out of the solutions from the media. A control without DMSO (data not shown) and with DMSO, using an equal amount of pure DMSO as used to prepare the 1 mM stock solution, were run as well. All controls worked as expected with no difference in growth between the untreated and DMSO-only controls. Experiments were done in triplicate.

In vitro cell adhesion assay

To measure cell adhesion, 1×10^5 cells were plated, in 2 mL of media, in each well of a 6-well dish and allowed to adhere for 30 minutes following treatment with blackseed oil or thymoquinone. Treatment groups were given either 0.5 or 1.0 μL of blackseed oil, or 1.0 or 5.0 μM of thymoquinone with little to no separation out of the solutions from the media. Cells were subsequently washed with PBS and counted to determine adhesive ability. A control without DMSO (data not shown) and

with DMSO, using an equal amount of pure DMSO as used to prepare the 1 mM stock solution, were run as well. As stated above, the controls worked as expected with no difference between DMSO alone and an untreated control. Experiments were done in triplicate.

Soft agar assays

Soft agar assays were performed, as previously described, to measure anchorage-independent growth using 1×10^5 cells in 3 mL of media for each treatment.[18] Treatment groups were given 0.5 or 1.0 μL of blackseed oil, or 1.0 or 5.0 μM of thymoquinone with little to no separation out of the solutions from the media. A DMSO-only control was also utilized following the concentrations used in previous experiments (see above). Experiments were done in triplicate.

RESULTS

Blackseed oil and thymoquinone decrease cell proliferation in aggressive breast cancer cells

In order to determine the minimal effective doses of blackseed oil and TQ to allow suppression of aggressive cell growth, with minimal negative effects on non-aggressive cell growth, a range of concentrations of blackseed oil and TQ used in previous experiments was utilized on MCF-7 and MDA-MB-231 cells (data not shown).[7-9,11,12,14-17] Within this range, the minimal effective dose of each compound was determined to be 0.5 μL and 1.0 μL of blackseed oil and 1.0 μM and 5.0 μM of thymoquinone. In order to determine the effects of these minimal effective doses of blackseed oil and thymoquinone on the cell growth of MCF-7 and MDA-MB-231 breast cancer cell lines, cell proliferation was measured. Cells were allowed to proliferate for a week and counted on the third and seventh day. When grown in the presence of 0.5 and 1.0 μL of blackseed oil, there was no significant difference seen between the MCF-7 control and treatment groups (**Figure 1A**). However, there were significant differences found between the MDA-MB-231 control, 0.5 μL, and 1.0 μL treatment groups for blackseed oil (**Figure 1B**). MCF-7 cells treated with 0.5 and 1.0 μM of TQ also displayed no significant differences from the control (**Figure 1C**). MDA-MB-231 cells treated with 5.0 μM TQ exhibited significant differences when compared to the control group, but 1.0 μM TQ showed no significant differences (**Figure 1D**). In general, MDA-MB-231 cells were affected more greatly than MCF-7 cells by treatment with blackseed oil and thymoquinone.

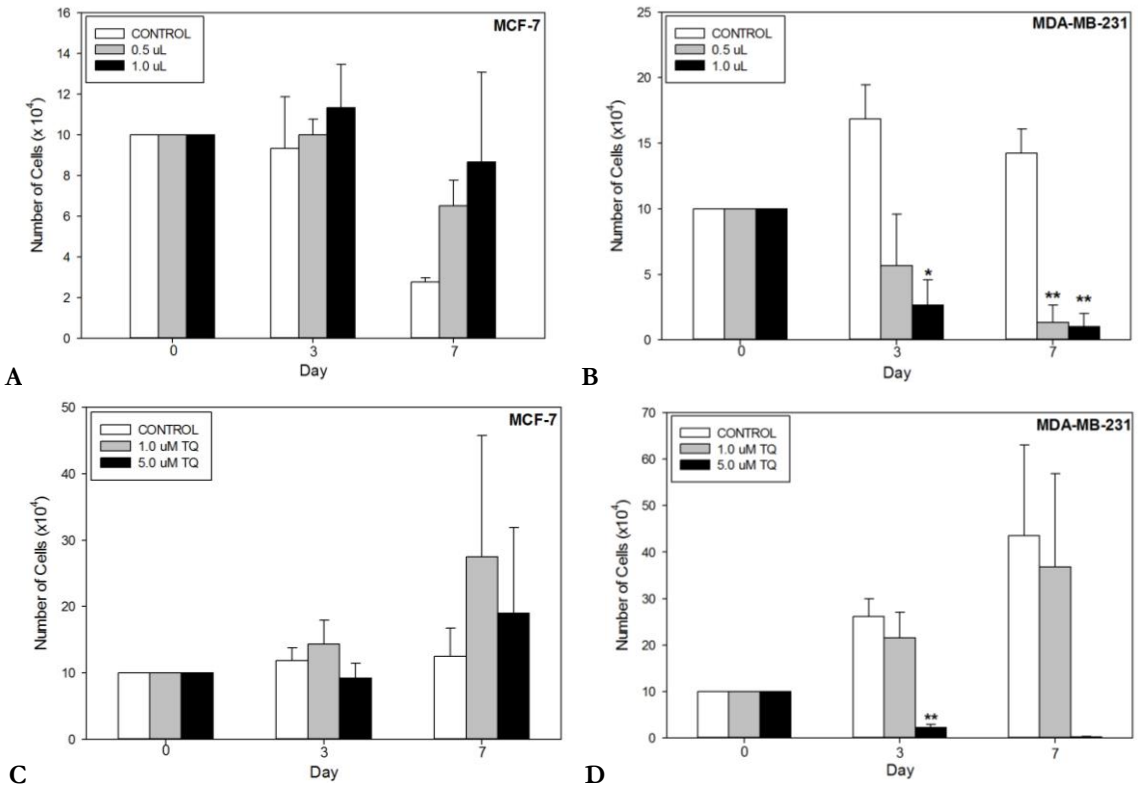

Figure 1. Cell proliferation of MCF-7 (A, C) and MDA-MB-231 (B, D) cell lines in presence of blackseed oil (A-B) and thymoquinone.(C-D) * represents statistical significance $p<0.05$ and ** represents statistical significance $p<0.01$ compared to control.

Blackseed oil and thymoquinone do not affect cell adhesion

More aggressive cell phenotypes are more motile and it is easier for these cells to make and break strong cell adhesions. To test the adhesion abilities of MCF-7 and MDA-MB-231 cells when treated with blackseed oil and thymoquinone, an adhesion assay was run for 30 minutes with and without treatments. Treatment groups for both blackseed oil and thymoquinone showed no significant differences compared to the control groups, indicating that these solutions had no effect on the adhesion ability of MCF-7 and MDA-MB-231 cells (**Figure 2A, 2B**).

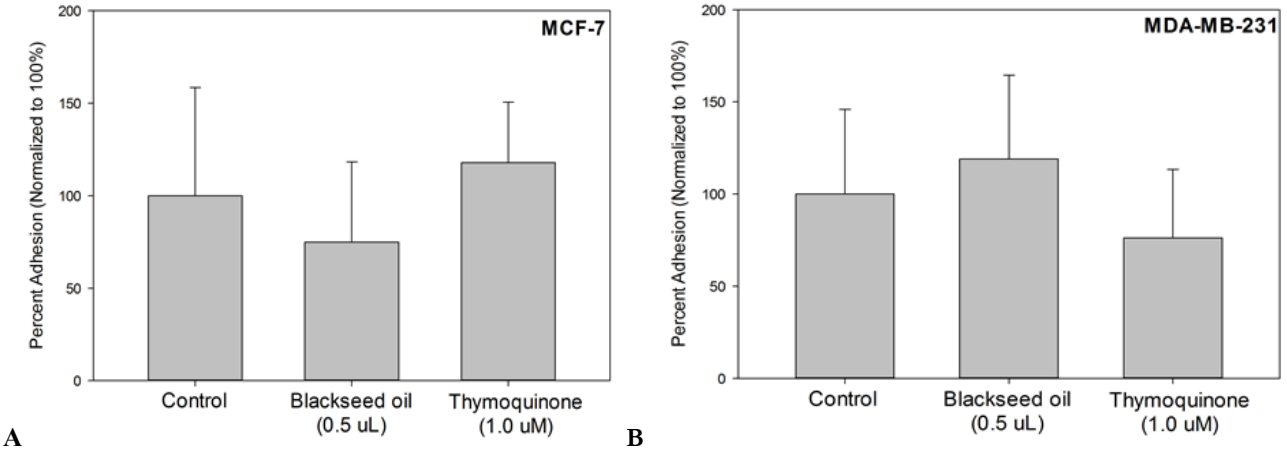

Figure 2. Cell adhesion assay of (A) MCF-7 and (B) MDA-MB-231 cell lines in presence of blackseed oil and thymoquinone**.**

Blackseed oil and thymoquinone cause decreased growth in soft agar

To test the ability of MCF-7 and MDA-MB-231 cells to grow in an anchorage-independent manner, a common sign of tumor progression in cancer cells, a soft agar assay was done in the presence of blackseed oil and thymoquinone. Both MCF-7 and MDA-MB-231 cells demonstrated significantly lower growth in soft agar with treatment of blackseed oil and thymoquinone compared to the control (**Figure 3**). Furthermore, growth of aggressive MDA-MB-231 cells was significantly inhibited compared to non-aggressive MCF-7 cells in the presence of thymoquinone. This indicates that treatment with blackseed oil and thymoquinone decreases the aggressive phenotype of MCF-7 and MDA-MB-231 and that invasive/metastatic MDA-MB-231 cells are more susceptible to thymoquinone than their non-aggressive counterparts.

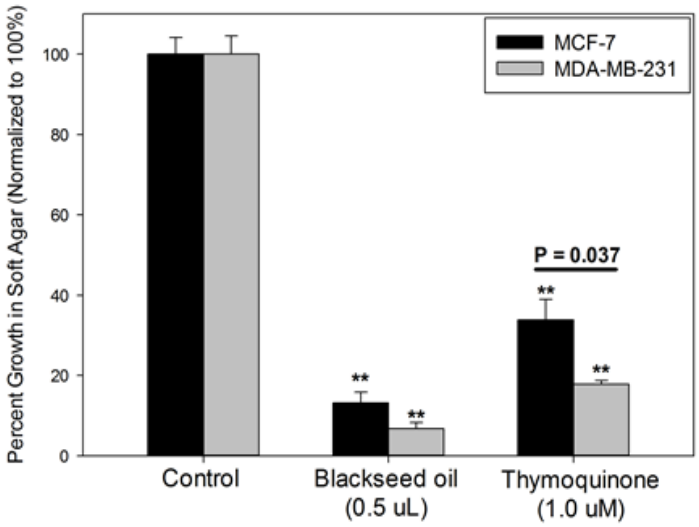

Figure 3. Percent cell growth in soft agar for MCF-7 and MDA-MB-231 cell lines in presence of blackseed oil and thymoquinone.** indicates $p<0.001$ compared to control. P=0.037 (shown on graph) indicates colony growth of MDA-MB-231 compared to MCF-7.

DISCUSSION

Given the unfavorable prognosis of individuals diagnosed with breast cancer—with the 5-year survival rate for females diagnosed with invasive breast cancer found to be as low as 26%—it is imperative to find a novel therapeutic method.[2] Therefore, the data presented in this paper demonstrate the effectiveness of blackseed oil and its active ingredient, thymoquinone, on limiting the aggressive phenotype of breast cancer cells. Although some of the results presented here do not reach significance, as noted in the above section, the general trends observed in the results of this study do provide valuable information regarding the effects of blackseed oil and TQ on both non-aggressive MCF-7 and aggressive MDA-MB-231 breast cancer cells.

In **Figure 1**, cell proliferation assay results are seen for both MCF-7 and MDA-MB-231 cell lines in the presence of varying concentrations of blackseed oil or TQ. Overall, the results shown in this figure are all in agreement with previous works exhibiting the effectiveness of blackseed oil and/or its active ingredient, TQ, in suppressing growth of a variety of cancerous cells—including cancerous cells of the colon, lungs, pancreas, and other cancers.[6–9] An analysis of key points of the figure proved to be quite interesting in this study. In **Figure 1A**, a general trend was seen where growth of MCF-7 cells treated with blackseed oil was generally greater than that of the control cells, which instead seemed to be dying. This trend was seen to be more prominent at the higher concentration of the oil (1.0 μL), as compared to the lower concentration (0.5 μL), while control cells showed the highest death rate overall (**Figure 1A**). This is an interesting observation, as it would suggest that these particular, low doses of blackseed oil may actually aid in the growth of non-aggressive cell types, as MCF-7 are intended to model in this work. In **Figure 1B**, however, a completely different set of results was exhibited, where the effect of varying concentrations of blackseed oil on aggressive MDA-MB-231 cells was noted. Here, it was seen that the presence of blackseed oil inhibited MDA-MB-231 cell growth at both 0.5 and 1.0 μL concentrations, while the control cells instead displayed increased growth in the first three days.

The fact that the MCF-7 cells showed increased growth in the presence of blackseed oil (**Figure 1A**), while the more aggressive MDA-MB-231 cells instead displayed significant reductions in cells (**Figure 1B**), proves to be quite interesting. This phenomenon suggests that low doses of blackseed oil may work to not only suppress the growth of aggressive cancer cells, as noted both in this study and a plethora of previous studies, but also to support the growth of non-aggressive cells in the body.[6–9] Because one of the major issues seen with current forms of breast cancer therapy include the unnecessary death of normal cells of the body—leading to such negative side effects as ototoxicity, cardiotoxicity, infertility, premature menopause, osteoporosis, and other effects—treatment with blackseed oil would seem to be an ideal solution to this issue, where only aggressive cells seem to be negatively affected, while non-aggressive cells are positively affected.[3,4] These results give credence to the notion that the clinical use of blackseed oil for cancer treatment may not only work for targeted killing of aggressive cancerous cells, but also for aiding in the growth of non-aggressive cells in the body. Additionally, the results present a dichotomy between the effects of blackseed oil and TQ on estrogen receptor-positive MCF-7 cells versus estrogen receptor-negative MDA-MB-231 cells.[19] As the presence of these chemicals seemed to have a positive effect on non-aggressive growth of MCF-7 cells at these concentrations, while inhibiting growth of aggressive MDA-MB-231, this may suggest that the mechanisms of action of blackseed oil and TQ are estrogen receptor-dependent. Future studies in which mechanistic actions of these solutions are studied, specifically in relation to the estrogen receptor, would help clarify why this phenomenon was observed, and whether or not the estrogen receptor and subsequent pathways are involved.

Previous literature focusing on the active ingredient of blackseed oil, TQ, has also been noted to have shown the effectiveness of the substance in fighting cancers of the larynx, colorectal area, brain, breasts, oral lining, blood, and a number of other cancers.[9,11–17] However, research done on breast cancer cells was limited in the usefulness of its results by either use of excessively high concentrations of TQ on only a single form of breast cancer cells, as seen in one study, or the use of a lipid carrier molecule alongside TQ in another study, which may have affected the results.[14, 15] In order to find a minimal effective dosage of TQ in preventing the growth of aggressive breast cancer cells, as done with blackseed oil as seen above, without the interference of a lipid carrier molecule, this study utilized lower concentrations of TQ on both cell types. **Figure 1C** displays the growth of the MCF-7 cell lines treated with 1.0 μM and 5.0 μM thymoquinone over the span of a week. A general trend of increased cell growth for the treated groups, in comparison to the control, was observed from day three to day seven of the experiment, although decreased cell growth was observed in the 5.0 μM group from day zero to day three. Nevertheless, the results of this figure again support the notion that even the active ingredient of blackseed oil, thymoquinone, seems to help support cell growth of non-aggressive cell lines, while preventing growth of the aggressive cell line.

Figure 1D displayed cell growth of MDA-MB-231 cells treated with TQ over the span of seven days. Here, it can be observed from the growth of the MDA-MB-231 cells treated with 5.0 μM TQ that cell proliferation was significantly decreased at day three, compared to the control cells (**Figure 1D**). However, this figure suggested that the lower

concentration of TQ did not decrease cell growth quite as significantly as the 5.0 μM concentration. These results would support the use of 5.0 μM TQ as a better dosage for destroying aggressive breast cancer cells, although this concentration was also noted to have slightly less positive effects on non-aggressive cells than the lesser, 1.0 μM dosage, as described above.

The second part of this experiment tested for the effects of blackseed oil and TQ on adhesion of MCF-7 and MDA-MB-231 cells. As mentioned earlier, more aggressive phenotypes are capable of making and breaking strong cell adhesions. Thus, given the effectiveness of blackseed oil and TQ on decreasing cell proliferation of the aggressive MDA-MB-231 cells, while encouraging the growth of non-aggressive MCF-7 cells, it was expected that effects on cell adhesion would also be seen. The adhesion assay for MCF-7 cells treated with 0.5 μL of blackseed oil and 1.0 μM of thymoquinone exhibited no significant differences in comparison to each other or the control (**Figure 2A**). The adhesion assay for MDA-MB-231 cells treated with 0.5 μL of blackseed oil and 1.0 μM of thymoquinone revealed no significant differences from the control, either (**Figure 2B**). While this was the case, it was seen that blackseed oil generally seemed to decrease the adhesion of MCF-7 cells, while increasing the adhesion of MDA-MB-231 cells, and TQ tended to have the exact opposite effects. These preliminary data indicate that the mechanism for action of blackseed oil and TQ is not by regulating adhesion formation in cells. While many studies have shown the effect of blackseed oil and TQ on cell proliferation[6-9], none have provided a viable mechanism of action in cells. Therefore, these results, along with previous published data, support the need for further experimentation with greater sample size to determine the mechanism of action of TQ and blackseed oil.

Finally, a soft agar assay was run in order to test each cell line's ability to grow in an anchorage-independent manner upon treatment with 0.5 μL of blackseed oil and 1.0 μM of thymoquinone. In **Figure 3**, it was shown that treatment with both blackseed oil and thymoquinone exhibited a significant decrease in anchorage-independent growth of both the MCF-7 and MDA-MB-231 cells, as indicated by the results of the soft agar assay. These results correlate with the data observed in **Figure 1** and indicate that blackseed oil and TQ appear to regulate a step in tumor progression to a more aggressive state. However, the precise mechanism is still unknown. Additionally, TQ induced a significant decrease in anchorage-independent cell growth of the MDA-MB-231 cells in comparison to the non-aggressive, MCF-7, breast cancer cells, supporting the idea that aggressive breast cancer cells are more susceptible to TQ than non-aggressive cells. These results are especially intriguing because all prior studies utilizing breast cancer cell lines were performed with such high dosages that no difference was observed between aggressive and non-aggressive cancer cell lines.[14] The results presented here are the first, to our knowledge, to indicate that aggressive cells may be more susceptible to TQ than non-aggressive cells which could have important clinical ramifications.

Although some of the results presented here did not reach significance, the general trends depicted certainly suggest the effectiveness of each substance tested – blackseed oil and its active ingredient TQ – at suppressing the growth of aggressive MDA-MB-231 cells, while possibly even supporting the growth of non-aggressive cells treated with the same substances. These data reinforce the results of previous works showing the effectiveness of blackseed oil and/or TQ at suppressing cancerous cell growth. [6–9, 11-17] Moreover, the current tests also supported the use of as low a clinically effective dose of blackseed oil as 0.5 μL, although the 1.0 μL dose seemed to show better results and even earlier significance at killing aggressive breast cancer cells, while still supporting non-aggressive cell growth. The experiments also support the use of as little as a 5.0 μM dosage of TQ, at which point aggressive breast cancer cell growth was again suppressed, while non-aggressive cell growth was supported. These points are especially important to make, as previous literature has been seen to have shown the death of both non-aggressive and aggressive breast cancer cell lines in the presence of TQ.[14] However, the experiments conducted in those cases utilized extremely high dosages of TQ – even as high as 80 μM – and thus did not provide information on the minimal dosage of TQ or blackseed oil that would be effective at suppressing aggressive cancer cell growth, while supporting the growth of non-aggressive cells. The data presented here do provide such suggested dosages, as mentioned above, and the effectiveness of the lower dosages is more clinically relevant. Interestingly, it was noted in the course of this study that thymoquinone generally took more time than blackseed oil to induce changes in proliferation. This may have indicated the presence of some other compound in the blackseed oil, aside from TQ, that would support the effects described here. This may be a point to note for further study.

CONCLUSIONS

Taken together, the data indicate the potential of blackseed oil and its active ingredient, thymoquinone, to combat breast cancer by limiting cancer progression to a less aggressive phenotype through decreased cell proliferation and anchorage-independent growth. This study is the first to provide a clinically relevant, minimal effective dosage of blackseed oil and thymoquinone. Further studies to elucidate the mechanism by which blackseed oil and TQ function in breast cancer cells may lead to novel treatments in the future.

ACKNOWLEDGMENTS

This work was supported by the Ellen Marks Cancer Foundation (SLR).

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: SC SS TKS SLR. Performed the experiments: SC SS TKS. Analyzed the data: SC SS TKS. Contributed reagents/materials/analysis tools: SS SLR. Wrote the paper: SC SS TKS SLR. All student authors contributed equally to this work.

REFERENCES

1. American Cancer Society, Inc. Cancer Facts and Figures. http://www.cancer.org (accessed May 2016)
2. American Cancer Society, Inc. Breast Cancer Facts and Figures. http://www.cancer.org (accessed May 2016
3. Eryilmaz, A., Demirci, B., Gunel, C., Eliyatkin, N., Aktas, S., Omurlu, I. K., Basal, Y., Sagiroglu, M., Ermisler, B., Basak, S. (2016) Evaluation of lapatinib and trastuzumab for ototoxic effects, *J Int Adv Otol* 11, 207–211.
4. Ewertz, M., Jensen, A. B. (2011) Late effects of breast cancer treatment and potentials for rehabilitation, *Acta Oncol* 50, 187–193.
5. Singh, S., Das, S. S., Singh, G., Schuff, C., Lampasona, M. P. D., Catalán, C. A. N. (2014) Composition, *in vitro* antioxidant and antimicrobial activities of essential oil and oleoresins obtained from black cumin seeds (*Nigella sativa* L.), *Biomed Res Int* 2014, 1–10.
6. Salim, E. I., Fukushima, S. (2003) Chemopreventive potential of volatile oil from black cumin (*Nigella sativa* L.) seeds against rat colon carcinogenesis, *Nutr Cancer* 45, 195–202.
7. Al-Sheddi, E. S., Farshori, N. N., Al-Oqail, M. M., Musarrat, J., Al-Khedhairy, A. A., Siddiqui, M. A. (2014) Cytotoxicity of *Nigella sativa* seed oil and extract against human lung cancer cell line, *Asian Pac J Cancer Prev* 15, 983–987.
8. Banerjee, S., Kaseb, A. O., Wang, Z., Kong, D., Mohammad, M., Padhye, S., Sarkar, F. H., Mohammad, R. M. (2009) Antitumor activity of gemcitabine and oxaliplatin is augmented by thymoquinone in pancreatic cancer, *Cancer Res* 69, 5575–5583.
9. Khan, A., Chen, H. C., Tania, M., Zhang, D. Z. (2011) Anticancer activities of *Nigella sativa* (black cumin), *Afr J Tradit Complem* 8, 226–232.
10. Şeleci, D. A., Gümüş, Z. P., Yavuz, M., Şeleci, M., Bongartz, R., Stahl, F., Coşkunol, H., Timur, S., Scheper, T. (2015) A case study on *in vitro* investigations of the potent biological activities of wheat germ and black cumin seed oil, *Turk J Chem* 39, 801–812.
11. Rooney, S., Ryan, M. F. (2005) Effects of alpha-hederin and thymoquinone, constituents of *Nigella sativa*, on human cancer cell lines, *Anticancer Res* 25, 2199–2204.
12. Gali-Muhtasib, H., Diab-Assaf, M., Boltze, C., Al-Hmaira, J., Hartig, R., Roessner, A., Schneider-Stock, R. (2004) Thymoquinone extracted from blackseed triggers apoptotic cell death in human colorectal cancer cells via a p53-dependent mechanism, *Int J Oncol* 25, 857–866.
13. Gurung, R. L., Lim, S. N., Khaw, A. K., Soon, J. F. F., Shenoy, K., Ali, S. M., Jayapal, M., Sethu, S., Baskar, R., Hande, M. P. (2010) Thymoquinone induces telomere shortening, DNA damage and apoptosis in human glioblastoma cells, *PLoS ONE* 5, e12124.
14. Dastjerdi, M., Mehdiabady, E., Iranpour, F., Bahramian, H. (2016) Effect of thymoquinone on P53 gene expression and consequence apoptosis in breast cancer cell line, *Int J Prev Med* 7, 66.
15. Ng, W. K., Yazan, L. S., Yap, L. H., Wan Nor Hafiza, W. A., How, C. W., Abdullah, R. (2015) Thymoquinone-loaded nanostructured lipid carrier exhibited cytotoxicity towards breast cancer cell lines (MDA-MB-231 and MCF-7) and cervical cancer cell lines (HeLa and SiHa), *Biomed Res Int* 2015, 263131.
16. Abdelfadil, E., Cheng, Y. H., Bau, D. T., Ting, W. J., Chen, L. M., Hsu, H. H., Lin, Y. M., Chen, R. J., Tsai, F. J., Tsai, C. H. (2013) Thymoquinone induces apoptosis in oral cancer cells through P38β inhibition, *Am J Chin Med* 41, 683–696.
17. El-Mahdy, M. A., Zhu, Q., Wang, Q. E., Wani, G., Wani, A. A. (2005) Thymoquinone induces apoptosis through activation of caspase-8 and mitochondrial events in p53-null myeloblastic leukemia HL-60 cells, *Int J Cancer* 117, 409–417.
18. Topczewska, J. M., Postovit, L. M., Margaryan, N. V., Sam, A., Hess, A. R., Wheaton, W. W., Nickoloff, B. J., Topczewski, J., Hendrix, M. J. C. (2006) Embryonic and tumorigenic pathways converge via nodal signaling: role in melanoma aggressiveness, *Nat Med* 12, 925–932.
19. Dore-Savard, L., Lee, E., Kakkad, A., Popel, S., Bhujwalla, ZM. (2016) The Angiogenic Secretome in VEGF Overexpressing Breast Cancer Xenografts. *Sci Rep* 6, 39460.

ABOUT THE STUDENT AUTHORS

Sabrina Chaudhry graduated from Elmhurst College in 2016 with her Bachelor of Science degree in Biology and a minor in Chemistry. She is currently applying to medical school.

Safia Siddiqui graduated from Elmhurst College in May of 2017 with a Bachelor of Science degree in Biology and minors in Chemistry and Medical Humanities. Her future goals include pursuing a career as an osteopathic physician.

Tyrnnon Steffen graduated from Elmhurst College in May of 2016 with a Bachelor of Science degree in Biology and a minor in Chemistry. She hopes to attend graduate school in the future.

PRESS SUMMARY

Breast cancer remains as the second leading cause of cancer deaths in women within the United States with limited treatment options that exhibit a variety of harmful side effects. Blackseed oil and its active ingredient, thymoquinone, are natural remedies that have recently demonstrated their effectiveness in limiting proliferation of breast cancer cells. However, previous studies have only exhibited this phenomenon at compound concentration levels that are too high to be clinically relevant. This study illustrates the effectiveness of low concentrations of blackseed oil and thymoquinone on limiting the aggressive phenotype of breast cancer cells through the suppression of cell proliferation and anchorage-independent growth.

Speech-Language Pathologists' Perceptions of the Common Core State Standards: A Multi-State Study

*Nicole Ariza**, *Patrick R. Walden*

Department of Communication Sciences and Disorders, St. John's University, Jamaica, NY

*Student: nicole.ariza11@stjohns.edu**
Mentor: waldenp@stjohns.edu

ABSTRACT
This study investigated the Common Core State Standards (CCSS) from the perspective of Speech-Language Pathologists (SLPs) providing speech and language services to students with communicative disorders in schools. An invitation to participate in an anonymous, online questionnaire with both closed- and open-ended questions was posted to three online communities comprised of SLPs working in schools across the United States of America (U.S.). Eighty-seven SLPs working in states using the CCSS completed the survey. The survey focused on four primary areas—the perceived impact of the CCSS on service delivery, student outcomes, professional workload and continuing professional education. Participants reported consistent incorporation of standards into services, but varied methods of implementation, primarily unchanged student outcomes, increased professional workload and a need for additional training. Overall, the CCSS' intent to create consistent goals may not be accomplished due to variability in approaches in implementation of the standards. Additionally, more resources and trainings for SLPs are needed to fully implement the CCSS into speech-language intervention in the schools.

KEYWORDS
Common Core State Standards; School-Based Services; School-Based Issues; Speech-Language Pathology; Service Delivery; Student Outcomes; Professional Workload; Continuing Professional Education

INTRODUCTION

The Common Core State Standards (CCSS) are the current academic standards for mathematics and English language arts/literacy (ELA) created by the National Governors Association (NGA) and the Council of Chief State School Officers (CCSSO) in the United States of America (U.S.). The standards are not a curriculum, but instead provide benchmarks for what students at each grade level should know by the end of a school-year, all while ensuring college and career readiness among students.[1] Staskowski reported: "The CCSS are intended to update the way schools educate and the way students learn and to ultimately prepare the nation's next generation for the global workplace" (p.95–96).[2] While the standards are uniform across all states, implementation methods are decided at the state and local level, as are the instructional materials used.[3] The adoption of the standards is completely voluntary, however, there are financial incentives to encourage states to implement the standards. For example, states interested in Federal "Race to the Top" funds (financial incentive from the U.S. government) must adopt the CCSS to be considered eligible.[4] The CCSS have been adopted by 42 U.S. States, the District of Columbia, four U.S. territories and the U.S. Department of Defense Education Activity.[1]

Education Standards Movement

The education standards movement is the nation-wide interest in creating more challenging standards for students in academia and is rooted in the concern that the U.S. is not an academic competitor when compared to other countries.[3] The timeline of events within the movement began in 1989 when an educational summit was held under the leadership of George H.W. Bush and the nation's governors to set state-level standards for mathematics, language arts and science. In 2002, The No Child Left Behind (NCLB) Act placed a new emphasis on educator accountability through the use of state assessments used for school accreditation as well as the sanctioning of "low-performance" schools. Concerns for state-to-state standard variability began to emerge as attention was drawn to students who completed their schooling in different U.S. States, becoming prone to a possible achievement gap as they missed or repeated content.[3] In 2009, the development of the CCSS began in order to create consistent standards in English/language arts and mathematics across all states. The standards were created with input from content experts, teachers, parents, school chiefs and administrators and are "research-and-evidence-based".[1] In 2010, the NGA and CCSSO released the CCSS and in 2014, individual state assessments were replaced by the Partnership for Assessment of Readiness for College and Careers (PARCC) and the Smarter Balanced Assessment Consortium (SBAC), used across states who have adopted the standards.[3]

Support & Criticism for the CCSS

The CCSS have garnered both positive attention and heavy criticism.[5] Some supporters of the CCSS believe the standards are the solution to a broken education system. The Program for International Student Assessment (PISA), which measures the performance of 15 year-olds' reading, mathematics and science across the globe indicated that U.S. student performance, which ranked 31st in math and 17th in reading literacy, was below China, Finland, Singapore, Hong Kong, South Korea, Canada and Japan.[2] Other supporters explain that while the U.S. has high unemployment rates, there is a shortage of high-skilled workers, creating a "skills gap" in the areas of leadership, adaptability, literacy and numeracy, creativity and communication. These skills have not been a substantial focus in previous educational systems.[2] Many CCSS supporters also believe that the standards promote equality. In the past, students who struggled in their early years of school due to disabilities or language differences were taught less course material and fell further behind as they moved up grade levels. New educational standards found in the CCSS, however, promote "higher-quality education for all students, regardless of race, economic status, or disability" (p. 97).[2]

Most who argue against the CCSS agree that past standards have inadequately met the needs of students, but explain that the new standards are not a satisfactory solution to the deep-rooted problems within the U.S. educational system. One prominent argument against the CCSS is the ineffective implementation of the standards. The implementation methods used have been inconsistent across states and districts, and many schools are lacking in professional development efforts, leaving teachers unprepared to develop aligned curriculums.[5] Furthermore, while supporters of the CCSS highlight the opportunity for equality, those opposed argue that the standards do not solve the problems posed by factors such as "poverty, isolation and social inequity" that are present before classroom learning begins.[6] Because diverse students learn and develop at different rates, they should not necessarily be held to the same standards. Other arguments against the CCSS include the inappropriate difficulty level of the assessments, failure of standardization efforts in other countries and conflicts of interest related to much of the CCSS cited research.[5, 6]

School-Based SLPs &Incorporation of the CCSS

Speech-Language Pathologists (SLPs) are licensed professionals who prevent, assess, and treat disorders of communication and swallowing. SLPs work in both educational and healthcare settings with the primary goal to improve the communication (speech/language/hearing) competence of students and patients. According to the American Speech-Language-Hearing Association, SLPs play a crucial role in implementing the CCSS for students with communication disorders.[7] SLPs can provide input on curriculum and functional goals while working with an education team to benefit each individual child. To effectively provide this input, SLPs should be familiar with the standards, know the curriculum being used, understand what is expected of typically developing students in relation to listening, speaking, reading and writing, collaborate with other school professionals and understand the needs of diverse students.

Students with disabilities are protected under the Individuals with Disabilities Education Act 2014 (IDEA), which ensures the right to a "free and appropriate education". Of special interest to the SLP working with the CCSS is the *Application to Students with Disabilities* document, which explains how the CCSS should be applied to students with disabilities. The document states that students with disabilities should be "challenged to excel within the general curriculum and be prepared for success in their post-school lives, including college and/or careers".[8]

Given the developing role of the SLP in providing services consistent with the CCSS, it is vital to understand how SLPs have been prepared to meet this challenge. Therefore, the purpose of this survey study was to learn more about how SLPs working with students in public schools perceive the CCSS. The research asked "How do SLPs working in schools perceive the CCSS' impact on:

1. service delivery
2. student outcomes
3. professional workload
4. continuing professional education?"

METHODS AND PROCEDURES

Procedures

An invitation to take an anonymous, online survey and a direct link to the survey were posted on the American Speech-Language-Hearing Association's (the accrediting body for SLPs in the U.S.) *Special Interest Group 1: Language Learning and Education* online community, the *Special Interest Group 16: School-Based Issues* online community as well as the *SLP schools* online community. One subsequent reminder invitation was posted on these communities. St. John's University IRB approval was granted prior to administration of the survey.

Participants

Table 1 presents a breakdown of the backgrounds of the participants. The 106 participants were school-based Speech-Language Pathologists across the United States. Of the 106 participants, a total of 87 participants indicated they worked in a Common Core State and completed the survey. Most (83; 89.25%) participants worked in traditional public schools. Likewise, most (69; 79.31%) reported being over 40 years of age, and the majority (67; 77.91)% indicated having worked in schools for over 10 years. All participants of the survey were female, which is typical of the profession.[9]

Professional experience, school logistics and demographics	Number of respondents	Percentage of respondents
Type of School		
Traditional public school	83	89.25%
Public charter school	2	2.15%
Private school	1	1.08%
Special Education school	7	7.53%
Employment Location		
Urban	19	22.09%
Suburban	54	62.79%
Rural	13	15.12%
Age		
25 and under	0	0.00%
26-40	18	20.69%
41-55	35	40.23%
56 or older	34	39.08%
Gender		
Male	0	0.00%
Female	87	100%
Degree Level		
Bachelor's	0	0.00%
Master's	82	95.35%
Doctoral	4	4.65%
Grade Level Worked With		
K-2	68	79.07%
3-5	65	75.58%
6-8	42	48.84%
9-12	27	31.40%
		*multiple professions selected by each participant
Years of Clinical Experience in a School		
1-2	6	6.98%
3-10	13	15.12%
11-18	17	19.77%
19-26	17	19.77%
26+	33	38.37%
CF-Year	0	0.00%

Table 1. Participant Demographics.
Note: Some participants reported working in more than one setting creating results above 100%.

Instruments

The 58 question online survey was made using *SoGoSurvey (http://www.sogosurvey.com*). The five-part survey consisted of both open and closed-ended questions. The questions were based on position statements from the American Speech-Language-Hearing Association (accrediting body for SLPs) and overall implications of the Common Core State Standards.[7] The first section was a series of questions pertaining to the respondent and workplace demographics. Responses can be seen in **Table 1**. The second section focused on service delivery. SLPs were primarily asked questions regarding how they aligned standards with therapy and how they collaborated with teachers. They were also asked to rate the difficulty of both general and specific aspects of the CCSS. The questions in the third section were based on professional workload. Participants were asked to compare size of therapy groups, the number of referrals and clients, types of cases and changes in workload. The fourth section pertained to student outcomes. Questions focused on completion of Individualized Education Plan (IEP) goals, language learning-impaired students compared to their same-aged peers, motivation of children in therapy, and perceived mastery of standards. The final section centered on continuing professional education for which SLPs answered questions on training, time dedicated to preparation for therapy and overall familiarity with the CCSS.

Data Analysis

Participant responses for most survey items were coded according to the level of agreement, frequency of occurrence, or specific categories indicated for statements concerning the CCSS and service delivery, workload, student outcomes, and continuing education. Results for these items were interpreted using descriptive statistics, specifically, percent participants reporting at each level of agreement, frequency of occurrence, or specific category chosen.

The participants also answered open-ended questions regarding each survey area. The resultant textual data underwent thematic analysis.[10] Each participant response to each open-ended question was coded for its meaning(s). Coding and thematic analysis were facilitated through use of *Atlas.Ti 7.5*™ software package. Codes from each segment were compared to all other segments to determine suitability and to determine if new codes appeared from the data. New codes were added as necessary. After all segments were coded, codes were collapsed into themes or "similar codes aggregated together to form a major idea in the database" (p. 239).[10]

Authenticity and Dependability of Findings

Authenticity was achieved through the thick, rich descriptions that were used to support the reported themes.[11] Descriptions, or quotes, from participants were removed due to space limitations. Reliability was conceived as the stability of findings (codes and themes) across researchers.[11] To increase the dependability of the findings, both authors reviewed all codes and themes, and consensus for each was reached. Both authenticity and consensus were achieved through hours of face-to-face data review and discussion.

RESULTS

An image of the results of this study is found in **Figure 1**. Specifics to each area depicted in the figure follow in narrative form. Tables of extracted codes and code descriptions are presented throughout associated sections.

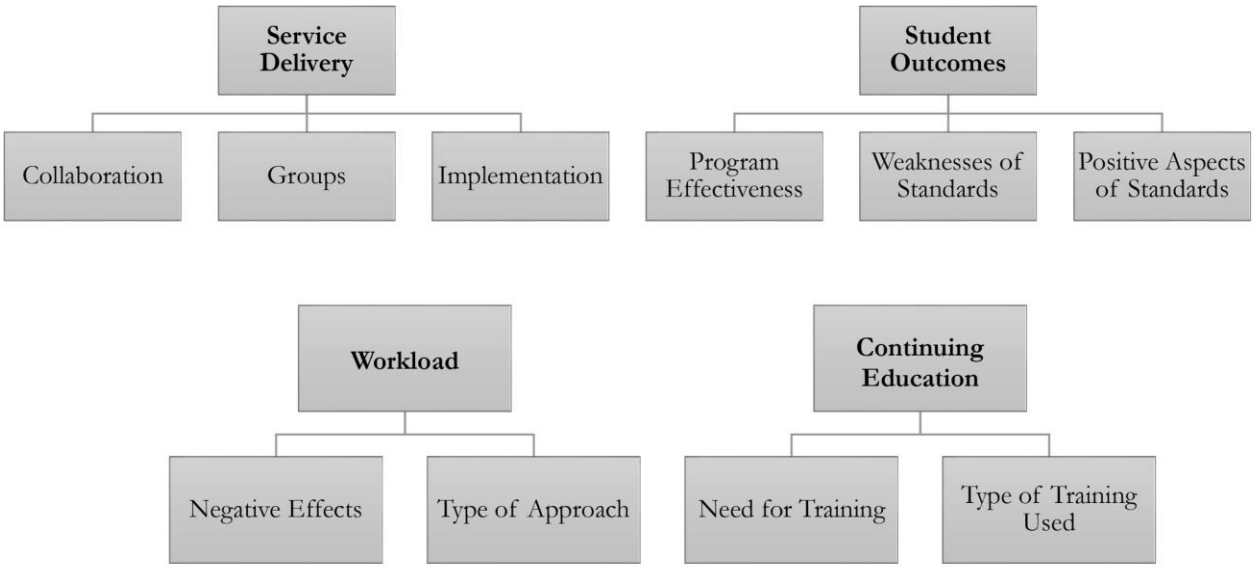

Figure 1. Survey Overview.

Service Delivery

Participant responses regarding the CCSS' effect on delivery of speech-language-hearing services fell into one of four categories: overall service delivery, collaboration, groups, and implementation. Interprofessional practice in the form of SLP-teacher collaboration was the most cited change in service delivery as a result of the CCSS. Characteristics of therapy groups were reported to be largely unchanged as a result of the CCSS (although individual report varied). Further, respondents reported that the implementation of therapy plans was affected by CCSS. Details of each of these follow.

Regarding overall service delivery, 86.05% (74/86) reported incorporating the CCSS into therapy in some or almost all sessions. A total of 44.83% (39/87) of participants indicated that developing a treatment plan that aligns with CCSS was easy or somewhat easy, and a total of 58.14% (50/86) of participants indicated that the CCSS slightly or moderately helped facilitate the generalization of skills focused on in therapy. When asked to compare time spent on the counseling aspect of therapy before and after implementation of the CCSS, 56.63% (47/83) indicated a small amount of time spent before implementation, while 60.98% (50/82) indicated a small amount of time spent after implementation. When asked about the helpfulness of the CCSS *Application to Students with Disabilities* document, 39.76% (33/83) indicated that they were not familiar with the document, while 32.53% (27/83) indicated that the document was somewhat helpful.

On a scale of 0 to 5, 0 being not met and 5 being completely met, participants were asked to rate which resources needed for service delivery were provided. Time to reach goals before CCSS assessments were rated by 49.38% (40/81) as 2 or 3. Sufficient time in therapy was rated by 50% (41/82) as 2 or 3. Sufficient class time was rated by 49.38% (40/81) as 2 or 3. Sufficient time for generalization of topics from therapy was rated by 55.70% (44/79) as 2 or 3. Instructional supports for learning was rated by 62.96% (51/81) as 2 or 3. Instructional accommodations were rated by 44.44% (36/81) as 4 or 5 and 44.44% (36/81) as 2 or 3. Assistive technology devices/services were rated by 48.15% (39/81) as 2 or 3.

Collaboration

Participants were asked to select the professions they collaborated with within the last school year (multiple professions selected by each participant). Of the 85 participants who responded, 92.94% (79/85) selected School Psychologist, 91.76% (78/85) selected another SLP, 87.06% (74/85) selected School Occupational Therapist, 62.35% (53/85) selected other professional, 57.65% (49/85) selected School Social Worker and 54.12% (46/85) selected School Physical Therapist. Participants were also asked about their collaborations with teachers, and 64.29% (54/84) reported the frequency of collaborating as very or extremely often. In addition, 57.83% (48/83) indicated that their collaboration with teachers was slightly to moderately difficult and 59.76% (49/82) indicated that working with teachers to design lesson plans was slightly to moderately difficult. Finally, 53.66% (44/82) of participants rated their overall experience collaborating with teachers as an average experience.

Collaboration-related segments (62 segments total)	Description
Building rapport with the classroom teacher (25/62 collaboration-related segments)	Collaborative success dependent on the development of respect and friendship with classroom teachers. Note: Of the 25 segments assigned in this category, reference was made to supporting or befriending teachers (13/25), initiating collaboration with one teacher at a time (8/25), and explaining the professional role of an SLP (4/25).
Adapting to teacher's lesson plans (10/62 collaboration-related segments)	SLP works around the teacher's set lesson plans and curriculum.
Creating an agreed upon form of communication (7/62 collaboration-related segments)	SLPs and teachers decide on how they will keep each other up-to-date on lesson plans and student progress.
Collaboration as the key factor to students' success (6/62 collaboration-related segments)	Collaboration is the most effective way to aid in the success of students with communication disorders receiving speech and language services.
Offering suggestions to teachers (4/62 collaboration-related segments)	SLPs offer suggestions to teachers regarding how they can help students with communication disorders.
Importance of classroom observations (4/62 collaboration-related segments)	Time spent observing students in the classroom important for collaboration.
Push-in as a service delivery model (3/62 collaboration-related segments)	Preference to keep students with communication disorders in the classroom.
No difficulty collaborating (2/62 collaboration-related segments)	Ease of use of collaboration model in service delivery.
Success of collaboration contingent on teachers' willingness to collaborate (1/62 collaboration-related segments)	Collaboration outcome dependent on teacher's willingness to work with the SLP.

Table 2. Collaboration-related segments.

Sixty-two participant open-ended responses were coded as statements referring to collaboration with teachers. These fell into nine categories: building rapport with the classroom teachers (25/62), adapting to the teacher's lesson plans (10/62), creating an agreed upon form of communication (7/62), collaboration as the key factor to students' success (6/62), offering suggestions to teachers (4/62), importance of classroom observations (4/62), push-in as a service delivery model (3/62), no difficulty collaborating (2/62) and success of collaboration contingent on teachers' willingness to collaborate (1/62).

Building rapport with the classroom teacher (25/62 collaboration-related segments) was assigned to responses that indicated that collaborative success was dependent on the development of respect and friendship with the teachers. Of the 25 segments assigned in this category, reference was made to supporting or befriending teachers (13/25), initiating collaboration with one teacher at a time (8/25), and explaining the professional role of an SLP (4/25). Adapting to the teacher's lesson plans (10/62 collaboration-related segments) referred to participant responses that suggested the SLP should work around the teacher's set lesson plans and curriculum. Creating an agreed upon form of communication (7/62 collaboration-related segments) referred to responses where the participant explained that SLPs and teachers should decide on how they will keep each other up-to-date on lesson plans and student progress. Collaboration as the key factor to students' success (6/62 collaboration-related segments) referred to responses that suggested that collaboration was the most effective way to aid in the success of students with communication disorders receiving speech and language services. Offering suggestions to teachers (4/62 collaboration-related segments) referred to participants who made suggestions to teachers regarding how they can help the students with communication disorders. The importance of classroom observations (4/62 collaboration-related codes) code was assigned to segments that suggested time spent observing students in the classroom was important for collaboration. Push-in as a service delivery model (3/62 collaboration-related codes) was assigned to responses that indicated a preference to keep students with communication disorders in the classroom. No difficulty collaborating (2/62 collaboration-related codes) referred to responses that indicated an ease of use of collaboration in service delivery. Success of collaboration contingent on teachers' willingness to collaborate (1/62 collaboration-related segments) was the code assigned to participants who suggested collaboration is only possible if the teacher is open to working with the SLP.

Groups

When asked about the number of students per group before implementation of the CCSS, 85.71% (72/84) reported small or moderate sized groups. When asked about the number of students per group after implementation of the CCSS, 78.57% (66/84) reported small or moderate sized groups (a decrease of 8.33% after implementation of the CCSS). When asked about the number of group sessions before the implementation of the CCSS, 82.14% (69/84) reported a moderate or large number of sessions. When asked about the number of group sessions after the implementation of the CCSS, 89.29% (75/84) reported a moderate or large number of sessions (an increase of 8.7% after implementation of CCSS).

Group-related segments (12 segments total)	Description
Larger group size (8/12 group-related segments) *part of group size (9/12) category	Increase of group therapy size after the implementation of the CCSS.
Unchanged group size (1/12 group-related segments) *part of group size (9/12) category	Unchanged group therapy size after the implementation of the CCSS.
Group diversity (3/12 group-related segments)	Groups consisted of students with communication disorders with disparate therapy needs.

Table 3. Group-related segments.

Twelve participant open-ended responses were coded as statements regarding characteristics of group sessions after the implementation of the CCSS. These fell into two broad categories: group size (9/12) and group diversity (3/12).

Group size consisted of larger group size (8/12) and unchanged group size (1/12). Larger group size (8/12 group-related segments) was the code assigned to participants who indicated that the size of their group therapy sessions increased after the implementation of the CCSS. Unchanged group size (1/12 group-related segments) was the code assigned to participants who believed group therapy size remained unchanged after CCSS implementation. Group diversity (3/12 group-related segments) was assigned to responses in which SLPs reported groups consisting of students with communication disorders with disparate therapy needs (questionable grouping of students).

Implementation

Implementation-related segments (175 segments total)	Description
Reference to standards (83/175 implementation-related segments)	Speech and language goals included use of Common Core standards. Note: Of the 83 total codes in the reference to standards category, specific mention was made to reference to specific standards (17/83), indirect application of standards into therapy (10/83), and direct application of standards into therapy (6/83).
Difficulty incorporating standards into therapy (24/175 implementation-related segments)	Difficulty aligning standards to speech and language therapy goals.
Compromised service (23/175 implementation-related segments)	Decline of effective speech and language services due to demands related to incorporation of the CCSS.
Focus on functional goals (13/175 implementation-related segments)	Focus on foundational skills as opposed to Common Core standards.
Reference to curricular material (12/175 implementation-related segments)	Speech and language goals centered on teacher's curricular material.
Need for prealigned therapy goals (7/175 implementation-related segments)	Need for pre-made therapy goals that align with standards to adequately implement the CCSS into speech and language services.
Implementation problems (6/175 implementation-related segments)	Issues related to implementation of CCSS.
Need for administrative support (6/175 implementation-related segments)	Administration involvement needed to adequately implement the CCSS into speech and language services.
No incorporation of the CCSS (1/175 implementation-related segments)	No use of CCSS in speech and language services.

Table 4. Implementation-related segments.

One hundred seventy-five participant open-ended responses were coded as statements regarding the implementation of the CCSS (no forced choice items on survey). These fell into nine categories: reference to standards (83/175), difficulty incorporating standards into therapy (24/175), compromised service (23/175), focus on functional goals (13/175), reference to curricular material (12/175), need for prealigned therapy goals (7/175), implementation problems (6/175), need for administrative support (6/175), and no incorporation of the CCSS (1/175).

The reference to standards (83/175 implementation-related segments) category referred to participants who included use of the Common Core standards into speech and language services at some capacity. The 83 segments assigned to the reference to standards included reference to specific standards (17/83), indirect application of standards into therapy (10/83), and direct application of standards into therapy (6/83). The difficulty incorporating standards into therapy (24/175 implementation-related segments) was assigned to participant responses that expressed difficulty in aligning standards with therapy goals. Compromised service (23/175 implementation-related segments) was assigned to responses that indicated a decrease in effective speech and language services due to demands related to incorporation of the CCSS. Focus on functional goals (13/175 implementation-related segments) referred to participants who indicated a focus on foundational skills as opposed to standards. The reference to curricular material (12/175 implementation-related segments) category was assigned to responses where participants explained basing speech and language services off of teacher's curricular material. Need for prealigned therapy goals (7/175 implementation-related segments) was assigned to participant responses that expressed a need for pre-made therapy goals and standard alignment to adequately implement the CCSS into speech and language services. Implementation problems (6/175 implementation-related segments) referred to problems in the implementation of the CCSS. Need for administrative support (6/175 implementation-related segments) arose from participant responses that expressed a need for administration's involvement to adequately implement the CCSS into speech and language services. No incorporation of the CCSS (1/175 implementation-related segments) referred to one participant who did not use the CCSS in speech and language services.

Student Outcomes

Participants' responses fell into one of three categories related to student outcomes. Participants reported data specific to overall program and therapy effectiveness, weaknesses of the CCSS as well as positive aspects of the CCSS. Responses indicated that student outcomes remained primarily unchanged; however, many participant responses drew attention to perceived weaknesses of the standards. Details for each category are described below.

Overall Program/Therapy Effectiveness

When asked if students who needed therapy services received them, 47.06% (40/85) indicated that they disagreed or strongly disagreed. When asked if students with language impairment fulfilled IEP goals for the year, 44.83% (39/87) responded that they agreed or strongly agreed. A total of 88.37% (76/86) indicated that they disagreed or strongly disagreed when asked if students with language impairment receiving speech and language services reached the grade-level of their same-aged peers. When asked if student outcomes changed since the implementation of the CCSS, 48.19% (40/83) indicated that they disagreed or strongly disagreed. When asked how outcomes had changed since implementation of the CCSS, 60.71% (51/84) responded that outcomes were unchanged. A total of 64.71% (55/85) indicated that the CCSS did not adequately address special needs students.

When asked their perception of their students' motivation to meet IEP goals, 78.57% (66/84) indicated students were slightly or moderately motivated. When asked how prepared they believed students were to master CCSS, 76.19% (64/84) indicated students were slightly to moderately prepared. A total of 80% (68/85) of participants reported that students with language-impairment were, from an academic standpoint, moderately distant or very distant from their same-aged peers before therapy, while 78.31% (65/83) reported that students with language-impairment were slightly or moderately distant from their same-aged peers after therapy.

Weaknesses of Standards

Weakness-related segments (44 segments total)	Description
Disregard of continuum of students (34/44 weakness-related segments)	Lack of consideration towards diverse students who are not on the same academic level as their same-aged peers.
Developmentally inappropriate standards (7/44 weakness-related segments)	Common Core Standards considered unrealistically advanced for the corresponding grade-level of students.
Preference to measure students' own progress (3/44 weakness-related segments)	Success of students with communicative disorders should not be measured by standards, but by their own progress throughout the school year.

Table 5. Weakness-related segments.

Forty-four participant open-ended responses were coded as statements regarding the weaknesses found in the CCSS. These fell into three categories: disregard of continuum of students (34/44), developmentally inappropriate standards (7/44) and preference to measure students' own progress (3/44).

Disregard of continuum of students (34/44 weakness-related segments) referred to participants who indicated that the CCSS do not take into account the diversity of students who are not on the same academic level as their same-aged peers. Developmentally inappropriate standards (7/44 weakness-related standards) referred to the CCSS being unrealistically advanced for the corresponding grade-level of students. Preference to measure students' own progress (3/44 weakness-related segments) was the code assigned to participants who believed the success of students with communicative disorders should not be measured by standards, but by their own progress throughout the school year.

Positive Aspects of Standards

Positive-related segments (2 segments total)	Description
Differentiated instruction (2/2 positive-related segments)	Opportunity for school staff to learn how to address diverse learners.

Table 6. Positive-related segments.

Two participant open-ended responses were coded as statements that indicated a positive aspect of the CCSS. These responses fell into a single category, differentiated instruction education (2/2). Differentiated instruction education (2/2 positive-related segments) referred to the opportunity for school staff to learn how to address diverse learners.

Workload

Participant responses regarding the CCSS' effect on workload fell into two categories: negative effects and type of approach. Insufficient amounts of time was the most cited difficulty. Most participants indicated a preference to a workload approach. Details of each category are specified below.

A total of 29.89% (26/87) of participants indicated that the number of assessment referrals increased after implementation of the CCSS. When asked if there was a change in overall caseload after implementation of the CCSS, 65.88% (56/85) indicated that caseload remained the same, however, 23.53% (20/85) indicated that there was an increase.

Negative Effects

A total of 79.52% (66/83) of participants either agreed or strongly agreed when asked if workload impacted service delivery. A majority, 86.36% (57/66) reported a negative impact on service delivery, while 4.55% (3/66) reported a positive impact.

Negative effect-related segments (117 segments total)	Description
Time restraints (50/117 negative effect-related segments)	Inadequate amount of time provided for workload.
Understaffed schools (29/117 negative-effect related segments)	Disproportionately large workload compared to the number of Speech-Language Pathologists employed in a given school.
Non-student related tasks (20/117 negative effect-related segments)	Participant responsibilities not directly related to working with students with communicative disorders. Note: Within this category, specific mention was made of paperwork (17/20) and meetings (6/20).
SLP profession not understood (12/117 negative effect-related segments)	Necessary to explain the role and importance of an SLP in the academic success of children with communicative disorders.
Personal time used (5/117 negative-effect related segments)	Completion of job-related tasks during non-work hours.
Fewer referrals (1/117 negative-effect related segments)	Decrease in speech/language referrals after the implementation of the CCSS.

Table 7. Negative effect-related segments.

One hundred seventeen participant open-ended responses were coded as statements regarding the negative workload effects of the CCSS. They fell into six categories: time restraints (50/117), understaffed schools (29/117), non-student related tasks (20/117), SLP profession not understood (12/117), personal time used (5/117) and fewer referrals (1/117).

Time restraints (50/117 negative effect-related segments) was the code that arose from responses that indicated a lack of time for the amount of work that needed to be completed. Understaffed schools (29/117 negative-effect related segments) referred to a disproportionately large workload compared to the number of Speech-Language Pathologists employed. Non-student related tasks (20/117 negative effect-related segments) referred to participant's responsibilities that were not directly related to working with communicatively disordered students. Within this category, specific mention was made of paperwork (17/20) and meetings (6/20). SLP profession not understood (12/117 negative effect-related segments) was the category that arose from participant responses that described having to explain the role and importance of an SLP in the academic success of children with communication disorders. Personal time used (5/117 negative-effect related segments) referred to participants who completed job-related tasks during their non-work hours. Fewer referrals (1/117 negative-effect related segments) was assigned to responses that indicated a decrease in referrals after the implementation of the CCSS.

Type of Approach

Type of approach-related segments (8 segments total)	Description
Caseload approach (6/8 type of approach-related segments)	Administrations that consider the number of clients SLPs treat for speech and language services.
Workload approach (2/8 approach-related segments)	Administrations that consider both number of clients seen and responsibilities outside of therapy, including meetings, paperwork and professional development.

Table 8. Type of approach-related segments.

Only open-ended questions were used to determine the type of caseload/workload approach used in schools. Eight participant open-ended responses were coded as descriptions of the type of approach used within their district, which fell into two distinct categories/themes: caseload approach (6/8) and workload approach (2/8).

Caseload approach (6/8 type of approach-related segments) referred to administrations who consider the number of clients SLPs treat for speech and language services, while the workload approach (2/8 type of approach-related segments) referred to administrations who took into account both clients seen and responsibilities outside of therapy, including meetings, paperwork and professional development.

Continuing Education

Major themes for continuing education fell into two categories: Need for training and the type of training in which participants engaged. Most participants reported needing more training to familiarize themselves with the CCSS. The training in which they reported engaging included self-directed learning, district/administration led training, teacher training/resources, ASHA training/resources and general online resources among others. The details for each of these categories follow.

Need for training

A total of 98.84% of participants (85/86) reported being at least somewhat familiar with the Common Core State Standards (CCSS). Yet, far fewer, 61.63% (53/86), reported being somewhat or very familiar with the *Application to Students with Disabilities* document.[8] Overall, 82.14% (69/84) of participants somewhat or strongly agreed that more training, similar to teacher continuing education focused on implementation of CCSS into curriculum, is necessary. However, 67.57% (25/37) of participants agreed that the time spent on CCSS training compromises time planning therapy sessions. A total of 88.89% (72/81) of participants indicated that time spent familiarizing with the CCSS was moderately or very time consuming.

Twenty-four participant open-ended responses were coded as statements regarding the nature of the type of additional training needed. These fell into nine categories: general conferences and workshops (7/24), online learning opportunities (4/24), training geared specifically toward SLPs (3/24), school district-specific training (2/24), training in collaboration with teachers (2/24), mentorship from more experienced SLPs (2/24), American Speech-Language-Hearing Association led training (2/24), training specifically on how to apply CCSS to students with disabilities (1/24) and training for administrators (1/24). Categories for type of training needed are self-explanatory. Table/definitions are not provided for this section.

Type of training used

Type of training-related segments (138 segments total)	Description
Self- directed learning (45/138 type of training-related segments)	Self-familiarization of Common Core Standards. Note: Of the 45 total codes in the self-directed learning category, specific reference was made to the "Common Core Standards" App by MasteryConnect (10/45), as well as self-direction by simply reading through the standards (10/45). All other responses were coded as unspecified self-directed learning.
District/administration-led training (34/138 type of training-related segments)	Formal SLP training provided by school districts and administrators. Note: Reference to Boards of Cooperative Educational Services (BOCES) Programming (1/34) and Professional Learning Community (PLC) implementation (1/34) were mentioned in the district/administration-led training.
ASHA trainings/resources (12/138 type of training-related segments)	Resources and trainings provided by ASHA. Note: Of the 12 codes in the ASHA trainings/resources category, specific reference was made to guidance in the form of course offerings and workshops (5/12), ASHA convention workshops (2/12), ASHA's public school conferences (2/12), general/unspecified ASHA resources (2/12), and Continuing Education Units (1/12).
Online resources (12/138 type of training-related segments)	Use of web-based resources to familiarize with the CCSS. Note: The online resources category consisted of online research (6/12) and online training (6/12). The online research category (6/12) referred to resources such as articles and blogs, while online training (6/12) was reserved for participants who used some form of web-based training to familiarize with the CCSS.
Collaboration as a means to familiarization with the CCSS (10/138 type of training-related segments)	SLPs working with teachers to familiarize with the CCSS.
SLP-specific training (9/138 type of training-related segments)	Trainings geared directly towards SLPs for incorporation of the Common Core standards into therapy sessions.
Teacher training/ resources (7/138 type of training-related segments)	Trainings or resources used by SLPs created specifically for classroom teachers. Note: Specific reference to the Teachers Pay Teachers website (*https://www.teacherspayteachers.com)* was made in the teacher trainings/resources category (5/7). The other codes in the category (2/7) did not identify a specific resource.
Unspecified training category (6/138 type of training-related segments)	No specific indication of type of training used.
Name-specific resources (3/138 type of training-related segments)	Inclusion of distinct names of individuals who have helped in the education of the CCSS.

Table 9. Type of training-related segments.

Only open-ended questions were used to determine the type of training used to familiarize oneself with the CCSS. Analysis of narrative responses resulted in 138 type of training-related segments. The resultant segments included nine categories of reported resources used: self-directed learning (45/138), district/administration-led training (34/138), American Speech-Language-Hearing Association (ASHA) trainings/resources (12/138), online resources (12/138), collaboration as a means to familiarization with the CCSS (10/138), SLP-specific training (9/138), teacher trainings/resources (7/138), unspecified training (6/138) and name-specific resources (3/138).

The self- directed learning category (45/138 type of training-related codes) referred to participant remarks that indicated self-familiarization with the standards. Of the 45 total codes in the self-directed learning category, specific reference was made to the "Common Core Standards" App by MasteryConnect (10/45), as well as self-direction by simply reading through the standards (10/45). All other responses were coded as unspecified self-directed learning; for instance, one participant simply reported, "self-familiarization". The district/administration-led training (34/138 type of training-related codes) category referred to formal training that school districts and administrators provided to the SLPs. Reference to Boards of Cooperative Educational Services (BOCES) Programming (1/34) and Professional Learning Community (PLC) implementation (1/34) were mentioned in the district/administration-led training. ASHA trainings/resources (12/138 type of training-related codes) referred to guidance in the form of course offerings and workshops (5/12), ASHA convention workshops (2/12), ASHA's public school conferences (2/12), general/unspecified ASHA resources (2/12), and Continuing Education Units (1/12). The online resources (12/138 type of training-related codes) code referred to participant responses that indicated use of web-based resources to familiarize with the CCSS. The online resources category consisted of online research (6/12) and online training (6/12). The online research category (6/12) referred to resources such as articles and blogs, while online training (6/12) was used for participants who used some form of web-based training to become familiar with the CCSS. Collaboration as a means to familiarization with the CCSS (10/138 type of training-related codes) was assigned to participants who reported working together with teachers to familiarize themselves with the CCSS. SLP-specific training (9/138 type of training-related codes) referred to trainings geared directly towards SLPs for incorporating the CCSS into therapy sessions. Teacher training/ resources (7/138 type of training-related codes) were trainings or resources used by SLPs that were created specifically for classroom teachers. Specific reference to the Teachers Pay Teachers website (*https://www.teacherspayteachers.com)* was made in the teacher trainings/resources category (5/7). The other codes in the category (2/7) did not identify a specific resource. The unspecified training category (6/138 type of training-related codes) indicated that participants received some form of training, but did not detail the type. The name-specific resources category (3/138 type of training-related codes) included distinct names of individuals who have helped in the education of the CCSS.

DISCUSSION

CCSS' Perceived Impact on Service Delivery

In terms of the participant SLPs' perception of the CCSS' impact on delivery of speech/language services, most (86.05%) SLPs reported incorporating the CCSS into speech-language sessions. However, less than half indicated that incorporating the standards was an easy task. Despite the perceived difficulty incorporating the standards, over half of the SLPs reported that the standards were at least somewhat helpful in the facilitation of the generalization of skills.

Around half of the SLPs also reported that time in therapy, class time, time for generalization of skills, as well as assistive technology were resource needs that were only somewhat met since implementation of the CCSS.

It was apparent from the data that school-based SLPs often collaborate with a range of other professionals. The most frequently cited strategy for successful collaboration with teachers was rapport building. For example, participants reported the need to develop both respect and friendship with teachers in order to maximize interprofessional collaborations. Further, participants also reported that collaboration may take the form of the SLP adapting to the teachers' lesson plans versus asking the teacher to adapt to the SLP's needs.

The participant SLPs reported that group sizes, for most, did not change as a result of the CCSS. However, there were some participants who reported fewer small-to-moderate group sizes after implementation of the CCSS as well as some who reported that number of group sessions increased after the implementation of the CCSS. These changes in group size and group session frequency were only reported by a minority of participants. This variability highlights differences across practice settings and suggests that it is unlikely that a "common" effect on service delivery will be experienced across the U.S.

Along the lines of implementation variability, the participants reported direct implementation of the CCSS into speech-language sessions while others reported more indirect methods. Again, a "common" experience across participants was not reported and it appeared that individual SLPs varied in how much difficulty was reported in implementing the standards into sessions.

Interestingly, a small minority of participants reported that implementation of the standards into speech-language sessions has decreased the quality of services.

CCSS' Perceived Impact on Student Outcomes

Despite some participant reports of a decrease in speech-language service quality as a result of the CCSS, more than half of those surveyed reported that there was no change in student therapy outcomes as a result of the CCSS. At the same time, over half also reported that the CCSS do not adequately address special needs students. The CCSS' perceived lack of application to special needs students may explain the perceived lack of change in student speech-language outcomes after CCSS implementation.

Along those lines, a theme based on the weaknesses of the CCSS emerged from open-ended responses to the survey questions. Specifically, some SLPs reported that the CCSS disregard a continuum in student ability. Others reported that the standards are developmentally inappropriate. These perceptions mirror those of critics of the CCSS.[5, 6]

Far fewer positive aspects of the CCSS arose from the participant open-ended response. Only two segments were identified which indicated support for the CCSS. Specifically, both of those segments highlighted how the implementation of the CCSS has added to staff education in terms of recognizing how to better address diverse learners. Identified weaknesses were far more robust in terms of the number of mentions in the dataset.

CCSS' Perceived Impact on Workload

Similar to the greater number of CCSS weaknesses versus strengths that the participants reported, workload was largely perceived to be negatively affected by the implementation of the CCSS. Workload for about a quarter of the participants was reported to have increased as a result of CCSS implementation. At the same time, over half of the participants reported that caseload had not changed since implementation of the CCSS.

Despite the largely unchanged caseload, almost 80% of the participants agreed that their overall workload was impacting service delivery. Over 80% specified that the impact on service delivery was negative.

Inspection of the open-ended data specified that time restraints was the major negative impact of the CCSS. Specifically, increased referrals coupled with understaffed schools and time spent on non-therapy tasks negatively affected the quality of speech-language services.

A small number of participants expressed frustration in the manner in which school districts measured productivity. For example, those SLPs reported that a workload approach would more adequately address the time issues that arose from the data.

CCSS' Perceived Impact on Continuing Education

A majority of participants indicated a desire for more training in order to properly implement the CCSS into speech and language services. Most participants indicated that time spent becoming familiar with the CCSS compromised time spent planning therapy sessions, which was possibly due to a lack of structured training. The most cited response to need for training was a need for more general conferences and workshops, followed by a need for online learning opportunities and need for training geared specifically toward SLPs. Most participants indicated that their current means of familiarizing themselves with the CCSS was through self-directed learning, and some indicated that they received district/administration-led training, although it was not always geared specifically to the role of the SLP. Other means of familiarization included ASHA trainings/resources and online resources. Many participants reported use of more than one approach to familiarize themselves with the CCSS. It is apparent from this research that continuing education/training specific to SLPs' implementation of the CCSS into speech-language therapy is needed. Based on these participants' self-directed learning efforts, online learning could be used to reach the most professionals in the most cost efficient and time-sensitive manner.

CONCLUSIONS

Major Outcomes and Significance

Prior to the completion of this study, we found no research regarding the effect of the CCSS on speech-language services in schools. Our survey highlighted four areas–service delivery, student outcomes, professional workload and continuing professional education. Negative aspects of the CCSS far exceeded positive aspects reported by participants, specifically through open ended questions. Many SLPs cited difficulties with implementation, much of which stemmed from time restraints and variability in implementation methods. Although a heavy majority of SLPs incorporated standards into therapy, some reported loosely using the standards as a guide, while others outlined direct methods of aligning standards to therapy goals. Closed-ended responses, many of which asked participants to compare student outcomes before and after CCSS, revealed that student outcomes after implementation of the CCSS remained primarily unchanged. Although student outcomes were comparable, there was a reported increase in workload responsibilities that negatively impacted speech-language services. Time restraints and paperwork were both cited as sources that compromised planning time for therapy. Many participants indicated familiarization of standards through self-direction and reported a need for additional training to aid in uniform implementation of CCSS into therapy. Our findings call for more communication in how CCSS are implemented across states to identify individual differences between states' efforts and additional training geared specifically to SLPs to increase the likelihood of effective incorporation of the CCSS into therapy.

Limitations

The current study aimed to gather perceptions of the CCSS across the U.S.; however, not every state was represented in the survey responses. The survey had a relatively small sample size of 106, which was reduced to roughly 87 after participants from non-CCSS states were subtracted from the total number of participants. There are approximately 72,359 ASHA-certified SLPs working in school settings across the U.S.[9] The sample of 87 respondents represents 0.1% of the total population working in schools. All respondents were female and a majority was over the age of 40. The invitation to complete the survey was only posted on three American Speech-Language-Hearing Association online communities and participants were self-selected, limiting the diversity of respondents to those who are active on the various online communities. While the results of the survey are helpful in understanding SLPs' perceptions of the CCSS, the strengths and weaknesses of the CCSS for children receiving speech-language services, and the ongoing training needs in educational settings, these results may not be representative of the school-based SLP population at large.

Future Studies

Workload concerns emerged as one of the most prominent barriers identified in the workplace. More research into how administrative support effects the perceived workplace satisfaction in light of the new standards is one possible area to be investigated. The focus can also be turned to ways administration can ease workload impact to promote the most effective speech and language services. In addition, a future study could compare workload and caseload approaches to measure both the job satisfaction and student outcomes that result.

Another area open to investigation is the collaborative relationships between SLPs and teachers. Collaboration was a subsection of the service delivery portion of this study, but can be investigated in much more depth. It would be interesting to consider the perceptions of both SLPs and classroom teachers in their support for collaboration and perception of importance to students with communicative disorders.

ACKNOWLEDGMENT

The authors thank the St. John's University's Graduate Admission Assistance Program for the funding of the research reported in this article.

REFERENCES

1. National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010a) About the Standards. Washington, DC: National Governors Association for Best Practices, Council of Chief State School Officers. Retrieved from *http://www.corestandards.org/about-the-standards/*
2. Staskowski, M. (2012) Overview of the common core state standard initiative and educational reform movement from the vantage of speech-language pathologists. *Semin Speech Lang*, 33(02), 095–101. doi: 10.1055/s-0032-1310310
3. Power-deFur, L. A. (2016) *Common Core State Standards and the speech-language pathologist: Standards-based intervention for special populations*. San Diego, CA: Plural Publishing.
4. McLaughlin, M., and Overturf, B. J. (2012) The common core: Insights into the K-5 standards. *The Reading Teacher*, 66(2), 153–164. doi: 10.1002/trtr.01115
5. VanTassel-Baska, J. (2015) Arguments for and against the common core state standards. *Gifted Child Today*, 38(1), 60–62. doi: 10.1177/1076217514556535
6. Doyle, K. (2012) Powerful alignment: Building consensus around the common core state standards. *Language and Literacy Spectrum*, 22, 7–23.
7. American Speech-Language-Hearing Association (n.d.) Common Core State Standards: A resource for SLPs. Retrieved May 1, 2016, from http*://www.asha.org/SLP/schools/Common-Core-State-Standards/*
8. National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010b) Application to Students with Disabilities. Washington, DC: National Governors Association for Best Practices, Council of Chief State School Officers. Retrieved from *http://www.corestandards.org/wp-content/uploads/Application-to-Students-with-Disabilities-again-for-merge1.pdf*
9. American Speech-Language-Hearing Association. (2017). ASHA summary membership and affiliation counts, year-end 2016. Available from www.asha.org.
10. Creswell, J. W. (2004) *Educational research: Planning, conducting, and evaluating quantitative and qualitative research,* 2nd ed. Upper Saddle River, NJ: Prentice Hall.
11. Miles, M.B., and Huberman, A.M. (1994) *Qualitative research analysis*, 2nd ed. Thousand Oaks, CA: Sage Publications.

ABOUT THE STUDENT AUTHOR

Nicole Ariza is a recent graduate of St. John's University. She majored in Speech-Language Pathology and Audiology and minored in Spanish. With a special interest in bilingualism and professional issues in Speech-Language Pathology, she looks forward to completing her graduate studies at St. John's University in Spring 2018.

PRESS SUMMARY

The implementation of the Common Core State Standards has been received with a combination of both support and opposition. The current study aimed to gather insight from the perspective of school-based Speech-Language Pathologists (SLPs) across the United States to measure the CCSS' impact within the profession. Results indicated that student outcomes have remained consistent before and after implementation, however, workload has heavily increased. Results also showed that there may be flaws and inconsistencies in the implementation of the standards, making it difficult for SLPs to properly incorporate the standards into therapy. The results indicated a need for continued professional education on the subject matter.

A Comparative Study of All-atom Molecular Dynamics Simulation and Coarse-grained Normal Mode Analysis in Identifying Pre-existing Residue Interaction Networks that Promote Coupled-Domain Dynamics in *Escherichia coli* Methionyl-tRNA Synthetase

Samuel C. Fehling, Alexander M. Strom, Brent P. Lehman, Ryan J. Andrews, Sudeep Bhattacharyya, and Sanchita Hati**
Department of Chemistry, University of Wisconsin-Eau Claire, Wisconsin 54702

*Student: andrewrj@uwec.edu**
Mentor: hatis@uwec.edu

ABSTRACT
Inter-domain communication plays a key role in the function of modular proteins. Earlier studies have demonstrated that the coupling of domain motions is important in mediating site-to-site communications in modular proteins. In the present study, bioinformatics and molecular simulations were used to trace "pre-existing" residue-residue interaction networks that mediate coupled-domain dynamics in multi-domain *Escherichia coli* methionyl-tRNA synthetase (Ec MetRS). In particular, a comparative study was carried out to evaluate the effectiveness of coarse-grained normal mode analysis and all-atom molecular dynamic simulation in predicting pre-existing pathways of inter-domain communications in this enzyme. Integration of dynamic information of residues with their evolutionary features (conserved and coevolved) demonstrated that multiple residue-residue interaction networks exist in Ec MetRS that promote dynamic coupling between the anticodon binding domain and the connective polypeptide I domain, which are > 50Å apart, through correlated motions. Mutation of residues on these pathways have distinct impact on the dynamics and function of this enzyme. Moreover, the present study revealed that the dynamic information obtained from the coarse-grained normal mode analysis is comparable to the atomistic molecular dynamics simulations in predicting the interaction networks that are essential for promoting coupled-domain dynamics in Ec MetRS.

KEYWORDS
Domain-domain Communication; Molecular Dynamics; Methionyl-tRNA Synthetase; Normal Mode Analysis; Coupled-domain Dynamics; Course-grained Normal Mode Analysis; Aminoacyl tRNA Synthetases; Statistical Coupling Analysis

INTRODUCTION

Studies on modular enzymes have suggested that coupling of dynamics between domains is critical for coordinating biological events occurring at distant sites.[1-7] There are two well-known models[8-9] for long-range allosteric communications - the "induced-fit" model (substrate-induced conformational change propagated through a single residue-residue interaction pathway) and the "population-shift" model (a perturbation at a distant site that alters the conformational equilibrium through "pre-existing" multiple pathways of residue-residue interactions).[8,9] Additionally, an updated model suggested that "pre-existing" multiple pathways of long-range site-to-site communications are present in the protein even in the absence of allosteric effectors.[8] After 50 years of debate, there is now growing consensus of opinion that conformational selection followed by conformational adjustment ("population-shift") is the mechanism of ligand binding and allostery. [7, 10-13]Also, theoretical and experimental studies have shown that the site-to-site communication is propagated by networks of coupled residues and regulated by enthalpic (conformational) and/or entropic (dynamic) changes.[14-17]

For single domain polypeptides, Ranganathan and coworkers have shown that one or more contiguous networks of residues that are evolutionarily conserved or coevolved facilitate communications between distant functional sites.[18-20] However, for multi-domain proteins, the molecular-level picture of the inter-domain communication becomes quite complex as each domain functions more like an independent entity in terms of backbone flexibility and structural stability. As earlier studies have demonstrated that coupling of domain dynamics is prerequisite in coordinating biological events occurring at distant sites, inter-domain communication pathways in modular proteins could be identified by tracing the evolutionarily constrained residues that could mediate coupled domain dynamics.

Molecular dynamic (MD) simulations and normal mode analysis (NMA) have emerged as two important tools for studying protein dynamics. Although MD simulations provide invaluable insight into the atomic-level details of protein dynamics, they are computationally expensive to sample atomic motions that are relevant for biological functions (microsecond to millisecond time-scale motions). The alternative approach to determine proteins' slow dynamics is NMA[21]. NMA could be used to sample wide-range of protein dynamics, including low-frequency collective motions, as well as high frequency local fluctuations. Interestingly,

Skjaerven et al. have shown that long-time-scale (200 ns) all-atom MD simulations, all-atom NMA, and coarse-grained NMA produce comparable results in depicting protein dynamics.[22] These observations suggest that the computationally inexpensive NMA method could be used to probe the site-to-site communication pathways in large modular proteins.

In this study, an attempt has been made to employ atomistic and coarse-grained simulation methods to trace the pre-existing pathways of residue interaction networks that could facilitate coupled-domain dynamics in the *Escherichia coli* methionyl-tRNA synthetase (Ec MetRS), a member of AARS family. Ec MetRS catalyzes covalent attachment of methionine to the $tRNA^{Met}$. The accurate synthesis of methionyl-$tRNA^{Met}$ requires coordination of several events occurring in distant domains (**Figure 1**) of the Ec MetRS.[23,24] Presence of inter-domain communication in Ec MetRS is evident from various experimental studies. For example, the methionylation of $tRNA^{Met}$ at the catalytic site is triggered by tRNA binding at the anticodon domain, which is located ~ 50 Å away from the aminoacylation site. A 10^6-fold decrease in aminoacylation activity (k_{cat}/K_M) was observed for tRNAs with truncated anticodon sequence, but identical acceptor stem sequence.[25-27] Moreover, mutations of tRNA anticodon nucleotides or the highly conserved C-terminal residues of MetRS (namely, W461, N452, and R395, which are essential for the anticodon base recognition) have significant (~10^5-fold) impact on the efficiency of tRNA aminoacylation.[28-31] Although, separate values of these kinetic parameters were not deciphered, the existence of an inter-domain communication, stretching from the anticodon binding domain to the CP domain (**Figure 1**), is believed to be crucial in shaping a functional Ec MetRS molecule.

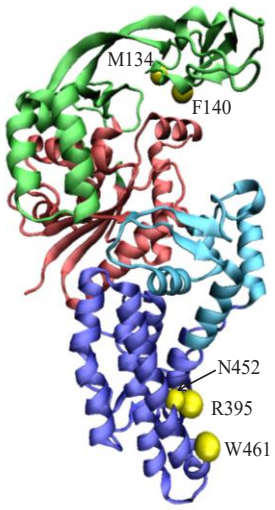

Figure 1. **Cartoon representation of the 3-dimensional structure of the monomeric form of Ec MetRS (residues 3-548, PDB entry: 1QQT).** The structural domains are colored as follows: red, Rossmann-fold (residues 1-96, 252-323); green, connective polypeptide (CP) domain (residues 97-251); cyan, KMSKS domain (residues 324-384 and 536-547) and blue, the C-terminal α-helix bundle domain (residues 385-535). The three residues, W461, N452, and R395, in the C-terminal domain and M134 and F140 of the CP domain are shown in yellow beads.

The long-range communication between the anticodon domain and the catalytic domain in Ec MetRS has been previously studied by atomistic MD simulations in the presence of tRNA.[32,33] Ghosh *et al.* have observed that the predicted communication pathways do not involve any residues of the tRNA. This poses an important question − if long-range inter-domain communications propagate through only protein, then the protein sequence, its structure/folding, and the intrinsic dynamics (arising out of the folding) have all necessary information that can promote such communications. In fact, this is evident from our combined bioinformatics and molecular simulation studies on *Thermus thermophilus* leucyl-tRNA synthetase (Tt LeuRS) and Ec prolyl-tRNA synthetase (ProRS), where we observed that the thermally coupled and evolutionarily constrained (coevolved and conserved) residues facilitate coupled-domain dynamics.[34,35] Therefore, in the present work we have employed the recently developed bioinformatics-based Statistical Thermal Coupling Analysis (STCA) method[35] to explore the molecular mechanism of inter-domain communications in Ec MetRS. In addition, we aimed to explore if the dynamical features obtained from the coarse-grained simulations and all-atom molecular dynamics simulations could produce comparable results. Finally, to validate the STCA results, selected residues on the predicted pathways were mutated and the impact of mutations of pathways residues on the collective dynamics of a distant domain were also probed using atomistic simulations.

METHODS AND PROCEDURES

Based on reported experimental data, W461 (C-terminal domain) and M134/F140 (CP domain) were selected as the two termini between which the interaction networks were mapped.[31] Especially, the M134 region of the CP domain and W461of the C-

terminal are known to be important for aminoacylation.[24] Visualizations and mutations were performed using VMD software[36]. Statistical coupling analysis (SCA) was carried out using a MATLAB script obtained from Ranganathan lab. Normal mode analysis (NMA) was carried using the coarse-grained Anisotropic Network Model (ANM)[21,37,38] using the bioinformatics server http://ignmtest.ccbb.pitt.edu/cgi-bin/anm/anm1.cgi. MD simulations were carried out using NAMD[39] and the all-atom CHARMM22[40] force field. Principal component analysis was carried out using CARMA.[41] The NMA and SCA plots and all data processing were carried out using MATLAB R2006b (The MathWorks Inc., Natick, MA).

Statistical Thermal Coupling Analysis

STCA was carried out in four discrete steps as described by Johnsons et al.[35] First, SCA[18,19] was performed to identify conserved and coevolved residues in the MetRS family. Next, the collective motions of various domains were studied by performing coarse-grained NMA,[21,37,38,42,43] as well as all-atom MD simulations. In the third step, the evolutionary dependence of the coupled-domain dynamics was explored by integrating the results of SCA and NMA/MD, which resulted in a subset of residues that are simultaneously coupled through evolution and thermal motion. In the final step, networks of interacting residues between the two distant sites were identified from the pool of dynamically and evolutionarily coupled residues using Dijkstra's algorithm.[44]

Statistical Coupling Analysis

SCA is based upon the assumption that the "coupling of two sites in a protein, whether for structural or functional reasons should cause those two sites to coevolve".[18,19] SCA was carried out using the protocol described elsewhere.[18,20] Briefly, this method uses a multiple sequence alignment (MSA) of a protein family as an input file and quantifies the extent of residue conservation, as well as coevolution between two residues by calculating the change in the amino acid distribution at one position with respect to a perturbation at another position. In the present work, the MSA of 478 protein sequences of the MetRS family was constructed using PSI-BLAST.[45] Only MetRS sequences that bear significant sequence identity (> 75 %) with Ec MetRS were included in this study. The conservation constant ΔG_i^{stat} and the coupling constant $\Delta\Delta G_i^{\text{stat}}$ were obtained using standard procedures described previously.[34,35]

Normal mode analysis

Low-frequency (large-amplitude) motions are important for the biological function of a protein. NMA has been shown to be reliable in describing the large-scale conformation changes in biomolecules.[27,46-48] In the present study, coarse-grained NMA was used in which a protein is simplified to a string of beads[49] and each bead represents a Cα atom.[21,37,38] The NMA was performed using the Anisotropic Network Model (ANM), where fluctuations are anisotropic and the overall potential energy of the protein system is expressed as the sum of harmonic potentials between the interacting Cα atoms.[49]

$$V_{ANM} = \frac{\gamma}{2}\left[\sum_{j,i\neq j} \Gamma_{ij}\left(r_{ij} - r_{ij}^{\circ}\right)^2\right] \qquad \textbf{Equation 1.}$$

In eq. 1, γ represents the uniform spring constant, r_{ij}° and r_{ij} are the equilibrium and instantaneous distance between residues i and j, respectively, and Γ_{ij} is the ij-th element of the binary connection matrix of inter-residue (Cα – Cα) contacts. Based on an interaction cutoff distance of r_c, Γ_{ij} is equal to 1 if $r_{ij}^{\circ} < r_c$ and zero otherwise.[37] Previous studies by Eyal *et al.* demonstrated that the 18 Å interaction cutoff resulted in a better correlation between experimental and calculated fluctuations.[38] Therefore, a 18 Å cutoff has been used in the present study. The substrate-unbound enzyme; the crystal structure of monomeric Ec MetRS (residues 3-549, PDB entry: 1QQT) was used.

The correlated or anti-correlated motions between Cα atoms were analyzed by computing the dynamic cross-correlation matrix **C**, where the ij-th element C_{ij} represents the cross-correlation coefficient between fluctuations of residues at sites i and j:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\sqrt{\langle \Delta r_i^2 \rangle \cdot \langle \Delta r_j^2 \rangle}} \qquad \textbf{Equation 2.}$$

The atomic (Cα) displacements of residues *i* and *j* are represented by Δr_i and Δr_j, respectively, and the angular brackets represent an ensemble average calculated over structures for combined normal modes.

Molecular Dynamics Simulation

MD simulations of Ec MetRS were performed using the X-ray crystal structure of Ec MetRS (PDB entry: 1QQT). All mutants were generated with the 'Mutate Residue' plug-in (version 1.3) of Visual Molecular Dynamics (VMD) version 1.9.1.[36] Simulations were performed in water (TIP3P model)[50] with substrate-free enzymes using the all-atom CHARMM27 force field[40] within the NAMD package.[39] Nonbonded interactions were truncated using a switching function between 10 and 12 Å, and the dielectric

constant was set to unity. The SHAKE algorithm[51] was used to constrain bond lengths and bond angles of water molecules and bonds involving a hydrogen atom. The MD simulations were performed using isothermal−isobaric (NPT) conditions. Periodic boundary conditions and particle mesh Ewald methods[52] were used to account for the long-range electrostatic interactions. In all MD simulations, a time step of 2 fs was used. The pressure of the system was controlled by the implementation of the Berendsen pressure bath coupling[53] as the temperature of the system was slowly increased from 100 to 300 K. During the simulations at 300 K, the pressure was kept constant by applying the Langevin piston method.[54,55] The WT and mutant proteins were solvated with water in a periodic rectangular box with water padding of 12 Å between the walls of the box and the nearest protein atom. The charge neutralization (with sodium ions) of the solvated system was performed with the VMD autoionize extension.[36] The resultant systems were equilibrated by slightly modifying previously described procedures.[56,57] All simulations were carried out with a 20 ps equilibration to minimize the amino acid side chain interactions followed by a 30 ns production MD run. The details of the MD simulation protocol were as described previously.[58] To evaluate the statistical significance of the MD simulation analysis, three replicates were generated for each protein system, as described in the protein simulation studies by Roy and Laughton.[59]

The correlated motions between residue pairs were studied by principal component analysis (PCA) of collective motions[60] as described earlier.[35,58,61] Following the method described in Johnson et al.,[35] the last 25 ns of the MD simulation data were used to generate principal components of atomic (backbone Cα atoms) fluctuations by Carma.[41] The first three principal components, which represent the low-frequency (high-amplitude) collective motions, were used to perform cluster analysis. This analysis produced a new trajectory of conformations representing the predominant conformational fluctuations and were used for generating dynamic cross-correlation matrix **C**. The *ij*-th element, C_{ij} in matrix **C** represents the cross-correlation coefficient between residue fluctuations at sites *i* and *j* during the simulation:

$$C_{ij} = \frac{\langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle}{\sigma_{x_i} \sigma_{x_j}} \qquad \textbf{Equation 3.}$$

The atomic (Cα) displacements of residues *i* and *j* are represented by x_i and x_j, respectively; the angular brackets represent ensemble averages, and σ_{x_i} and σ_{x_j} represent the standard deviations of these displacements.

Integration of Evolutionary and Dynamic Information

To trace the residue-residue interaction networks that play key role in facilitating coupled-dynamics between domains, we constructed a subset of residues that are both, evolutionarily and dynamically coupled.[35] The motional coupling information, obtained from NMA, was integrated with the evolutionary conservation and coevolution dataset from the SCA. The conserved and coevolved residues were treated separately. The conserved and dynamically coupled residues were obtained by selecting only those residues that exhibit significant conservation ($\Delta G_i^{\text{stat}} \geq 0.5$) as well as motional coupling ($C_{ij} \geq 0.8$) with each other. The value of C_{ij} was set to ≥ 0.8 in order to obtain statistically relevant size of the evolutionarily conserved residues. The coevolved and dynamically coupled residues were obtained from the truncated **C** matrix, which was created by including only those columns that are present in the normalized SCA-derived **G** matrix. Next, a new matrix, the coevolutionary dynamic coupling (**CDC)** matrix, was created by multiplying each *ij*-th element of the **G** matrix with the corresponding element of the truncated **C** matrix:

$$CDC_{ij} = \Delta\Delta G_{ij}^{stat} \times C_{ij} \qquad \textbf{Equation 4.}$$

Identification of interaction networks using dijkstra's algorithm

From the shortlisted residues, residue-residue interaction networks between W461 (C-terminal domain) and M134/F140 (CP domain) of Ec MetRS were identified using Dijkstra's algorithm as described earlier.[35,44] In this method, each Cα atom of the protein backbone represents a node. The connectivity between two adjacent nodes [inter-residue (Cα – Cα) contacts] is described by a binary connection matrix **P**.[35] The Cα – Cα distance matrix, **D** was created from the PDB file containing the Cartesian coordinates of all Cα atoms of the Ec MetRS (PDB entry: 1QQT). Based on a Cα – Cα cutoff distance D_{ij}°, P_{ij} is equal to 1 if $D_{ij} < D_{ij}^{\circ}$ and zero otherwise.[35,44] The interaction networks (pathways) between two functional sites were identified by varying the distance parameter, D_{ij} and the statistical parameter, ΔG_i^{stat}.

Root-mean-square fluctuations

To determine if the motion of CP domain had undergone significant change upon mutation along the interaction networks, the slow dynamics of the CP domain due to a specific mutation was studied by performing PCA for the WT and mutant proteins. Root-Mean-Square fluctuations (RMSF) of Cα atoms, averaged over three replicate simulations, were obtained for the WT and mutants. In these calculations, the last 25 ns of MD simulation data were used, each comprising an ensemble of 250,000

conformations.[35] Similar to our previous work, an ensemble of 750,000 conformations, obtained by combining three replicate trajectories, was used to perform PCA for each of these protein systems. The first three clusters, which represent the predominant conformational fluctuations, were used in this study.

RESULTS

Conserved and coevolved residues

To identify the conserved residues, the ΔG_i^{stat} value of each residue of the Ec MetRS sequence was calculated using SCA. ΔG_i^{stat}, a quantitative measure of the conservation of a residue at the *i*th position of the sequence was obtained as a one-dimensional vector normalized to 1 kT^* (**Figure 2**).

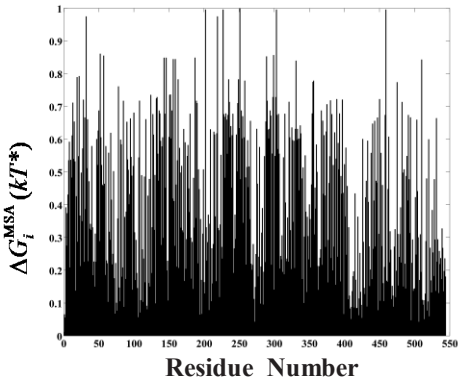

Figure 2. A normalized plot of ΔG_i^{stat} vs. position (residue number) to identify conserved residues from SCA. The extent of evolutionary conservation, ΔG_i^{stat}, of a residue at the *i*th position of the Ec MetRS sequence was determined by using the following equation:

$$\Delta G_i^{stat} = kT^* \sqrt{\sum_x [\ln(P_i^x / P_{MSA}^x)]^2}$$

where kT^* is an arbitrary energy unit, P_i^x is the binomial probability of observing amino acid x at site i, P_{MSA}^x is the probability of observing amino acid x in the overall MSA, and the summation is over all 20 amino acids.

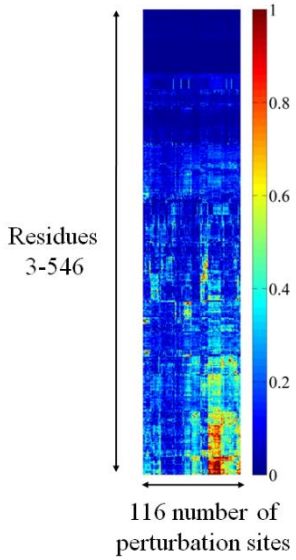

Figure 3. Statistical coupling analysis of the MetRS family. The normalized statistical coupling matrix with one dimensional hierarchical clustering along the perturbation axis is shown. The highly coevolved residues are clustered at the bottom right of the diagram. The color gradient, as indicated in the color bar, is as follows: blue squares represent the lowest (0 kT^*) and red squares represent the highest (1 kT^*) statistical coupling energies.

In addition, the evolutionary coupling indices ($\Delta\Delta G_i^{stat}$) of a total 544 residues for 116 perturbation sites were obtained from a 544×116 coupling matrix (**Figure 3**). It is evident that only a small fraction of Ec MetRS residues have high $\Delta\Delta G_i^{stat}$ values and those highly coupled residues were found to be located farther apart from each other in the three-dimensional structure of Ec

MetRS (**Figure 4**). To explore the role of these conserved and coevolved residues in the long-range interactions, the dynamics of the protein segments and their motional coupling was examined.

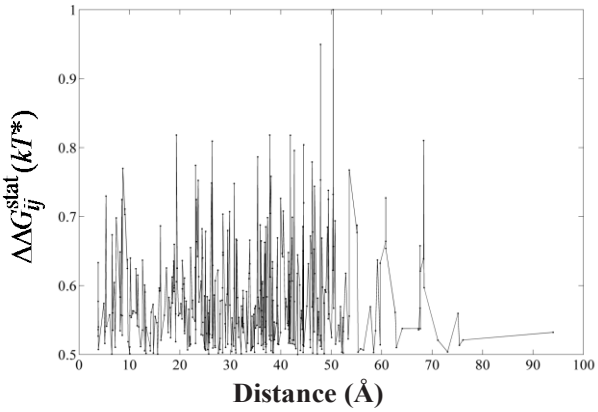

Figure 4. The statistical coupling, $\Delta\Delta G_i^{stat}$ for all residues plotted against their contact (C_α - C_α) distances. The extent of statistical coupling between two sites, *i* and *j*, of the MetRS protein sequence was estimated from a perturbation analysis, where a sub-alignment of the MSA was created using ~ 63% of the total number of sequences and the change in the extent of conservation was calculated. Mathematically, this is represented as a normalized statistical coupling matrix **G**, in which each *ij*-th element, $\Delta\Delta G_{i,j}^{stat}$, measures the perturbation in the conservation of residue *i* due to residue *j*. $\Delta\Delta G_{i,j}^{stat}$ refers to the evolutionary (statistical) coupling between residues at two functional sites, *i* and *j* and is expressed as:

$$\Delta\Delta G_{i,j}^{stat} = kT^* \sqrt{\sum_x \left[\ln\left(\frac{P^x_{i|\delta_j}}{P^x_{\mathrm{MSA}|\delta_j}} \right) - \ln\left(\frac{P^x_i}{P^x_{\mathrm{MSA}}} \right) \right]^2}$$

where $P^x_{i|\delta_j}$ is the probability of *x* at site *i* being dependent on a perturbation at site *j*. A large statistical coupling value, $\Delta\Delta G_i^{stat}$, indicates strong energetic coupling between residues at the two sites *i* and *j* [1; 2]

Normal mode analysis and flexible protein segments

The dynamic correlation between residues of the Ec MetRS was extracted from the NMA using a combination of the first 10 lowest-frequency normal modes (modes 1-10), which are usually biologically more relevant. The identical flexible regions in the protein structure are evident from the plot of computed and the experimental (crystallographic)[23] *B*-factors of C_α atoms (**Figure 5**). The study also revealed that the CP domain, especially residues 100-200, is the most mobile. Residues 56-71 in the Rossmann-fold and 454-476 of the anticodon binding domain also exhibit significant mobility, which is consistent to the atomistic simulation study reported earlier.[32,33]

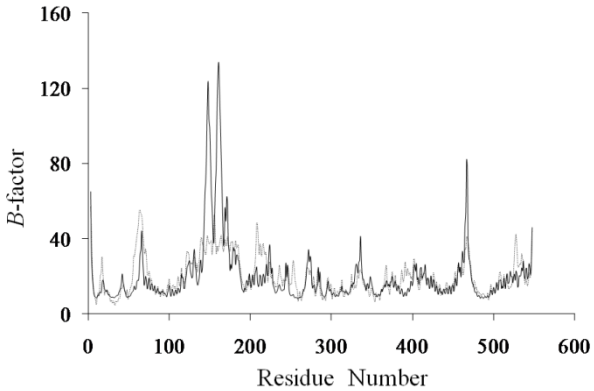

Figure 5. C_α *B*-factor analysis of Ec MetRS. Comparison of the C_α *B*-factor obtained from the crystal structure (gray, dotted line) (PDB entry: 1QQT) and the calculated one (black, solid line) using ANM.

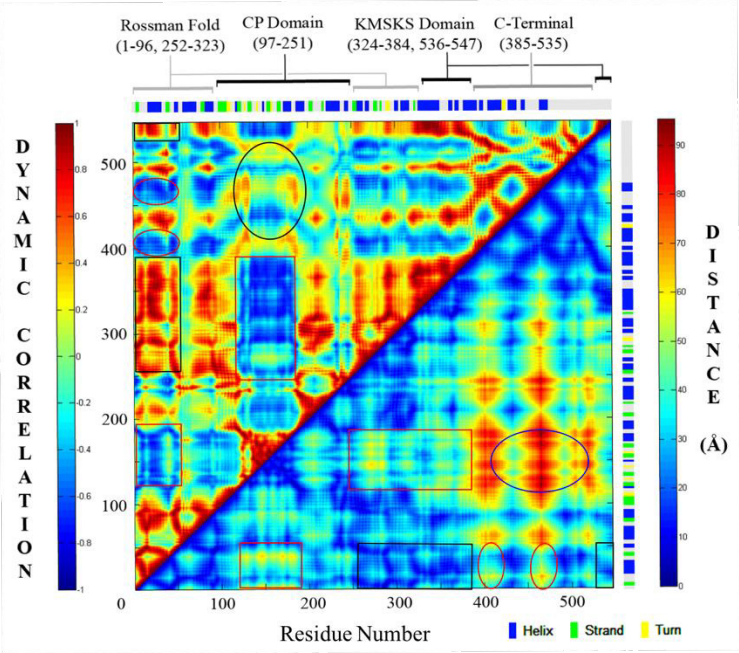

Figure 6. Correlations in fluctuations between residues (above diagonal) and their contact ($C_\alpha - C_\alpha$) distances (below diagonal). Dynamic cross-correlations between the C_α atoms of the Ec MetRS (PDB entry: 1QQT) have been calculated based on the lowest 10 normal modes. A value of +1.0 was set for strongly correlated motion and is colored red, whereas -1.0 was used for the strongly anticorrelated motion and is colored blue. The axis values correspond to residue number and the color scale for contact distances is set between 0 – 100 Å; blue for residues in close contact and red for residues located farther away.

Analysis of the dynamic cross-correlation matrix (DCCM) obtained from the combined modes (1-10) revealed that various structural elements within the catalytic domain are engaged in correlated motions (**Figure 6**). The Rossmann-fold domain (residues 1-96 and 245-323) is engaged in strong correlated motion with the KMSKS domain (residues 324-384 and 536-547, **Figure 6**, black rectangles). In contrast, the CP domain (residues 125-200) moves in an anticorrelated pattern with respect to the Rossmann-fold (residues 1-96 and 252-323) and the KMSKS (residues 324–384) domains (**Figure 6**, red rectangles). The observed anti-correlated motion between the CP domain and the main body is significant as this anti-correlated motion provide an adequate space for the 3'-end of tRNA to enter into the synthetic active site for aminoacylation. Similar observations have been made in other class I aminoacyl-tRNA synthetases.[34,62-64] Furthermore, the anticodon binding domain (residues 385-535) and the catalytic site (residues 1-96) exhibit anticorrelated motion (**Figure 6**, red ovals). The CP and the C-terminal domains (separated by a distance > 70 Å, Fig. 4, blue oval) are engaged in partly anticorrelated motion (**Figure 6**, black ovals). These observations are quite consistent with the earlier reported long time-scale (10 ns) MD simulation results,[32,33] as well as the present MD simulation study (*vide infra*). Therefore, it is evident that the approximations in the coarse-grained NMA are capable of reproducing the same

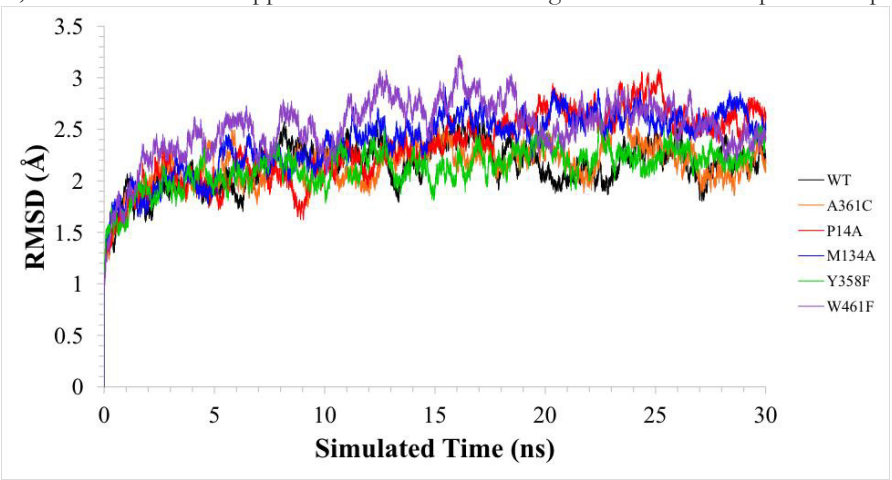

Figure 7. RMSD of the C_α atoms from their initial coordinate as a function of time for WT and the five mutants of Ec MetRS. Calculations of RMSDs were performed using 30-ns MD simulation data.

low-frequency dynamics of a large protein like Ec MetRS as obtained from the use of all-atom MD simulation.

MD simulations

30 ns MD simulation was carried out for the Ec MetRS to generate the dynamic cross-correlation matrix **C**. The quality of the simulation was first tested by computing the root-mean-square-deviation (RMSD) of the C_α atoms from their initial coordinates. The RMSD obtained from the 30 ns MD simulation trajectory is shown in **Figure 7** The C_α RMSD values were observed to fluctuate with a mean value of 1.0-1.5 Å during the production period (30 ns) simulations. The dynamic cross-correlation matrix **C** of the WT enzyme (**Figure 8**) was generated using the first three principal components. Analysis of the cross-correlation of fluctuations of residues revealed both inter- and intra-domain dynamic correlations, which are very similar to those obtained from NMA.

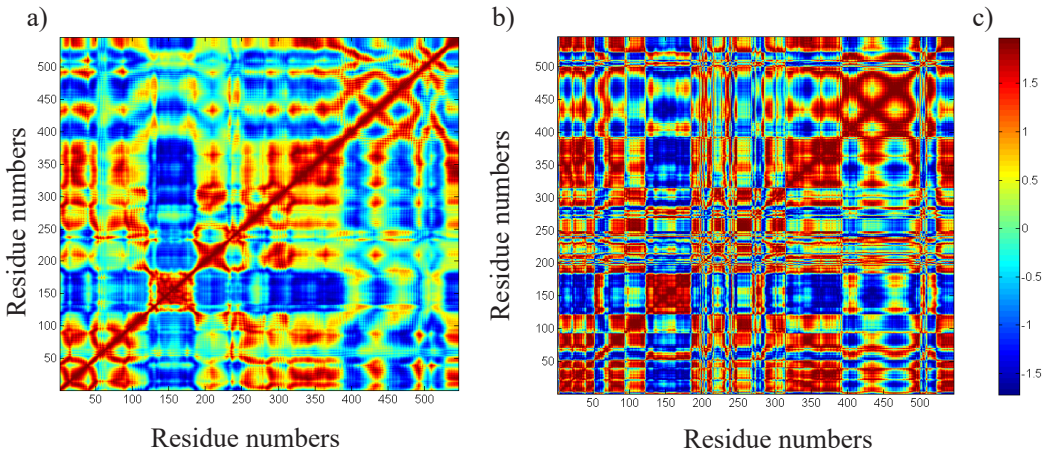

Figure 8. Dynamic cross-correlation between the C_α atoms of Ec MetRS as obtained from the a) combined first 10 low-frequency modes of NMA; b) 30 ns MD simulation; c) A color scale; a value of +1.0 represents a strongly correlated motion whereas -1.0 represents a strongly anticorrelated motion.

Generating residue pools that exhibit statistical and thermal coupling

Results of thermal fluctuations were combined with those of the evolutionary constraints (conservation and coevolution) of residues. Conserved and coevolved residues were treated differently. In order to obtain a reliable set of residues that were dynamically coupled as well as highly conserved in the MetRS family, a controlled experiment was carried out by varying ΔG_i^{stat} cutoff values and monitoring their distributions in various domains. For a specific ΔG_i^{stat} cutoff value, the number of conserved residues varies in each domain of Ec MetRS (**Table 1**). The study revealed that only ~ 15% of C-terminal residues are moderately conserved ($\Delta G_i^{stat} \geq 0.5$) and exhibit strong motional coupling. In contrast, the Rossmann-fold domain, which is the catalytic domain, contains higher number of conserved residues (~ 40%) that are also dynamically correlated (**Table 1**). In the present study, subsets of conserved and dynamically correlated residues were generated by using a high dynamic coupling constant ($C_{ij} \geq$ 0.8) and varying the conservation cutoff (ΔG_i^{stat}) between 0.50-0.65 (**Table 1**). The cutoffs of these parameters were selected in order to obtain a reasonable size of residue pool consisting of statistically and thermally coupled residues

Domain	ΔG_i^{stat}			
	≥ 0.5	≥ 0.55	≥ 0.60	≥ 0.65
Rossmann-fold (residues 1-96 and 252-323)	72	56	44	32
Connective polypeptide (residues 97-251)	66	57	49	36
KMSKS (residues 324-384 and 536-547)	26	24	19	9
C-terminal (residues 385-535)	29	22	16	12
Total	193	159	128	89

Table 1. Number of conserved residues in different domains of Ec MetRS obtained using various cutoff values for ΔG_i^{stat} . The dynamic coupling cutoff (C_{ij}) was set to greater than or equal to 0.8.

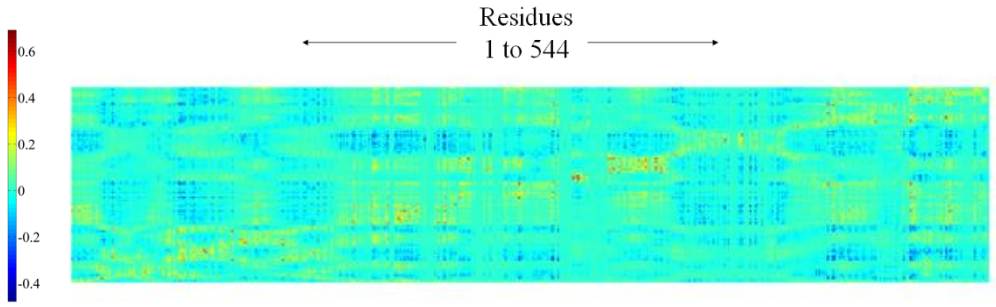

Figure 9. The SCA•NMA plot obtained from the **CDC** matrix by multiplying individual elements of the SCA matrix with the corresponding elements of the truncated NMA matrix. Values range from +0.67 (coevolved and thermally correlated) to −0.47 (coevolved and thermally anticorrelated).

The coevolving and dynamically correlated residues were extracted from the coevolutionary dynamic coupling matrix **CDC** (**eq. 4**), where CDC_{ij} values range between –0.48 to 0.69 for NMA (**Figure 9**) and -0.74 to 0.70 for MD simulation (data not shown). In order to obtain a statistically significant set of residues, a systematic study was conducted using variable CDC_{ij} cutoffs. This study resulted in subsets of residues with various degrees of statistical-thermal coupling and their distributions in different domains are reported in **Table 2**. In the present study, we have chosen the CDC_{ij} cutoff value of 0.4, which resulted in a pool of 76 residues.

Domain	ΔG_i^{stat}			
	≥ 0.3	≥ 0.35	≥ 0.40	≥ 0.45
Rossmann-fold (residues 1-96 and 252-323)(168 Total)	25%	19%	15%	11%
Connective polypeptide (residues 97-251) (155 Total)	23%	20%	15%	12%
KMSKS (residues 324-384 and 536-547) (73 Total)	33%	29%	20%	9%
C-terminal (residues 385-535) (151 Total)	28%	20%	10%	5%
Total	27%	21%	14%	9%

Table 2. Percentage of coevolved residues within different domains of Ec MMetRS obtained using various cutoff values for CDC_{ii} (eq. 4).

Interaction networks across domains

To map interaction networks between C-terminal domain (W461) and CP domain (M134/F140), Dijkstra's algorithm[44] was `van der Waal's radius of an amino acid C_α is ~ 3 Å and the acceptable distance of non-covalent interaction between two interacting atoms is 2.0−3.0 Å, a strong non-covalent interaction occurs between two C_α atoms of a folded protein when they are within a distance of 8.0-9.0 Å. Therefore, the distance cutoff, D°_{ij} was varied between 8.0-9.0 Å to include only those neighboring C_α atoms that are engaged in strong non-covalent interactions and therefore could propagate site-to-site communications between distant domains.[35]

Pathways	Parameters	Residue Network (MD)	Residue Network (NMA)
I	$G_i^{stat} \geq 0.5$, $D_{ij} \leq 8.0$ Å	W461→P460→C477→**G480** → F484 → L487→L491→P493→R36→V35→**F47**→ **C49** →A50→D52→H54→F238→P236→ M134	W461→P460→C477→**G480**→ F484 → L487 →L491→P493→R36→V35→**F47**→C49→ A50→D52→H54→ F238 →P236→M134
II	$G_i^{stat} \geq 0.5$, $D_{ij} \leq 8.5$ Å	W461 →P460→C477→**G480**→F484→ L487→L491→P493→V35→**F47**→**C49** →A50→D52→H54→Y237→P236→M134	W461→P460→C477→**G480**→F484→L487 →L491→P493→R36→V35→**F47**→C49→ A50→D52→H54→F238→P236→M134
III	$G_i^{stat} \geq 0.5$, $D_{ij} \leq 9.0$ Å	W461→P460→R395→N391→V386→ V381 →Y358→**Y359**→**L26**→H24→P14→ D52 →H54→Y237→M134	W461 → P460 → R395 → N391→V386→ V381 →Y358→**Y359**→**L26**→H24→P14→ D52 →H54→Y237→M134
IV	$G_i^{stat} \geq 0.55$, $D_{ij} \leq 8.0$ Å	W461 →P460→**V455**→N452→K388→ D384→ Y357→R380→Y358→**Y359**→I29→**H28→C11** →C49→D51→**A53**→H54→F238→P236→ M134	W461→P460→C477→G480→**A451**→D449 → **A446** →R442→**A361**→**Y359**→I29→**I89→Y91** →**Y94**→D51→**A53**→H54→F238→P236→ M134
V	$G_i^{stat} \geq 0.55$, $D_{ij} \leq 8.5$ Å	W461→P460→**V455**→N452→K388→ D384→ Y357→**R380**→Y358→Y356→F87→F84→H80 → D51→A53→H54→Y237→P236→ M134	W461→P460→C477→R395→N391→K388 →D384→Y357→**R380**→Y358→**Y359→L26→** H24 →P14→ D52 →H54→Y237→P236→M134
VI	$G_i^{stat} \geq 0.55$, $D_{ij} \leq 9.0$ Å	W461 →P460→**V455**→N452→K388 →D384 →Y357→**R380**→Y358→**Y359**→**L26**→ H24 →P14→D52→H54→Y237→M134	W461 → P460 → R395→N391→K388→ D384→ Y357→**R380**→Y358→**Y359**→**L26**→H24→ P14 →D52→H54→Y237→M134

Table 3. Probable pathways of communication between W461 and M134 in the Ec MetRS. Residues in bold are coevolved and the rest are evolutionarily conserved. The residues that are different between MD and NMA predicted pathways are underlined. Values of the two parameters, C_{ij} and CDC_{ij}, are set to greater than or equal to 0.8 and 0.4, respectively.

Pathways	Parameters	Residue Network (MD)	Residue Network (NMA)
I	$G_i^{stat} \geq 0.5$, $D_{ij} \leq 8.0$ Å	W461→P460→C477→**G480**→F484→L487 →L491→P493→D32→**A31**→T10→Y260 →A256→V252→Y251→R233→D234→ F135→L136→ F140	W461→P460→C477→**G480**→F484→L487 →L491→P493→R36→V35→V9→T10→Y260 →A256→V252→Y251→R233→D234→F135 →L136→F140
II	$G_i^{stat} \geq 0.5$, $D_{ij} \leq 8.5$ Å	W461→P460→C477→**G480**→F484→L487 →L491→P493→V35→V9→T10→Y260 →A256→Y251→R233→D234→F135→ L136→F140	W461→P460→C477→**G480**→F484→L487 →L491→P493→R36→V35→V9→T10→Y260 →A256→V252→Y251→R233→D234→F135 →L136→F140
III	$G_i^{stat} \geq 0.5$, $D_{ij} \leq 9.0$ Å	W461→P460→R395→N391→V386→ V381→Y358→**Y359**→**L26**→H24→P14→ D52→G55→T56→P137→F140	W461→P460→R395→N391→V386→V381→ Y358→**Y359**→**L26**→H24→P14→D52→H54→ Y237→M134→L136→F140
IV	$G_i^{stat} \geq 0.55$, $D_{ij} \leq 8.0$ Å	W461 → P460 →**V455** →N452→ K388→ D384→Y357→R380→Y358→**Y359**→Q30→ A31 →T10→ Y260 →A256→V252→Y251→ R233 →D234→F135→L136→F140	W461→P460→C477→G480→**A451**→D449→ A446 →R442→A361→Y359→I29→**I89**→Y91 →**C49**→G259→D255→Y251→R233→D234→ F135→L136→F140
V	$G_i^{stat} \geq 0.55$, $D_{ij} \leq 8.5$ Å	W461 → P460 →**V455** →N452→ K388→ D384→Y357→**R380**→Y358→**Y359**→Q30→ A31→T10→ Y260 →A256→Y251→R233→ D234 →F135→L136→F140	W461 → P460 →C477 →R395 → N391→ K388→D384→ Y357→**R380**→Y358→**Y359**→ **L26**→ H24→P14→D52→H54→Y237→P236 →M134 →L136→ F140
VI	$G_i^{stat} \geq 0.55$, $D_{ij} \leq 9.0$ Å	W461 →P460→**V455**→N452→K388→ D384 → Y357→**R380**→Y358→**Y359**→**L26**→ H24→P14→D52→H54→Y237 →M134→ L136→F140	W461→P460→R395→N391→ K388→ D384→ Y357→**R380**→Y358→**Y359**→**L26**→H24→ P14 →D52→H54→Y237→M134→L136→ F140

Table 4. Probable pathways of communication between W461 and F140 in the Ec MetRS. Residues in bold are coevolved and the rest are evolutionarily conserved. The residues that are different between MD and NMA predicted pathways are underlined. Values of the two parameters, C_{ij} and CDC_{ij}, are set to greater than or equal to 0.8 and 0.4, respectively.

Several interaction networks were identified between W461 and M134/F140 by using various cutoff values for ΔG_i^{stat} and D°_{ij}. Six contiguous interaction networks between C-terminal domain and CP domain are listed in **Tables 3 and 4**. These pathways (residue interaction networks) between W461 to M134/140 either pass through the α-helix bundle (residues 385 to 535; pathways I and II) or the helix-loop-strand-helix motif (residues 352 to 385; pathways III - VI) (**Figure 10**). Very similar results were almost identical (**Figure 10 and Tables 3 and 4**). Moreover, analysis of SCA results revealed that conserved residues are dominant over coevolved residues in these predicted pathways (**Table 3 and 4**).

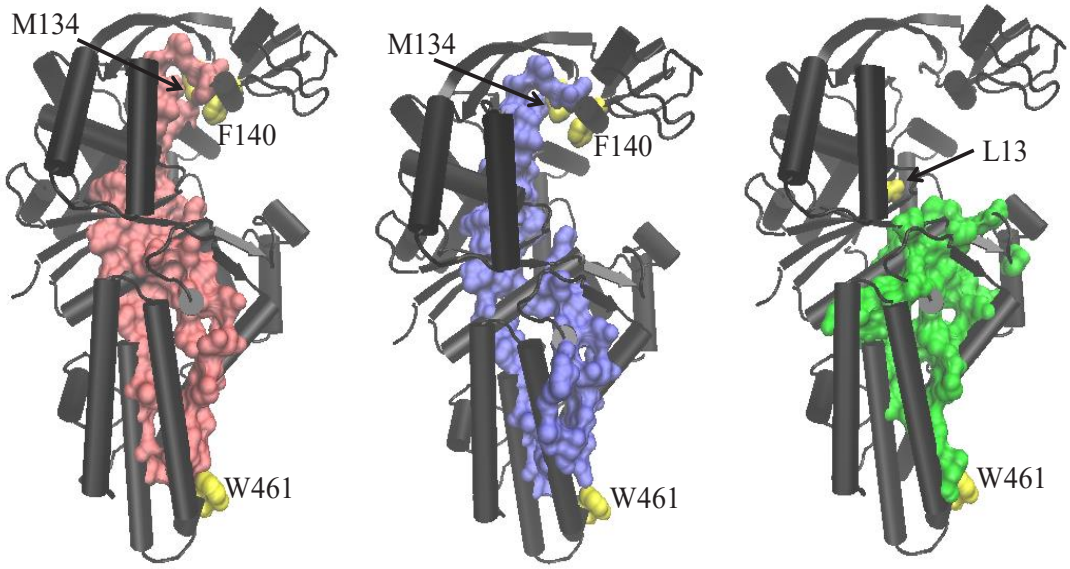

Figure 10. Representation of residue-residue interaction networks between the anticodon binding domain (W461) and the CP domain identified in this study. The two terminuses, W461 and M134/F140, are shown in yellow space-filling surface representation. The secondary structure is shown in cartoon representation. From left to right, contiguous pathways as predicted by a) NMA-SCA (12 pathways, Tables 3 and 4), b) MD-SCA (12 pathways, Tables 3 and 4) and C) MD-PSN (4 pathways, reference (*33*)) are shown in space-filling surface representations.

Essential dynamics analysis and role of pathway residues

Recent studies have shown that the principal components obtained from long duration simulations (200ns) provide very similar description of the conformational change as that obtained from short duration simulations (10 ns).[22] In this study, we have performed 30 ns simulations for the WT and various mutants to see the impact of mutation of pathway residues on the distant domain dynamics. Essential dynamics analysis of the WT protein was carried out using last 25 ns of 30 ns MD simulation data. The RMSF analysis reveals a flexible CP domain (residues 97-251, **Fig. 11**), similar to what was observed in coarse-grained (NMA) analysis (**Figure 4**). The CP domain, which is adjacent to the catalytic domain, is known to undergo conformational change upon

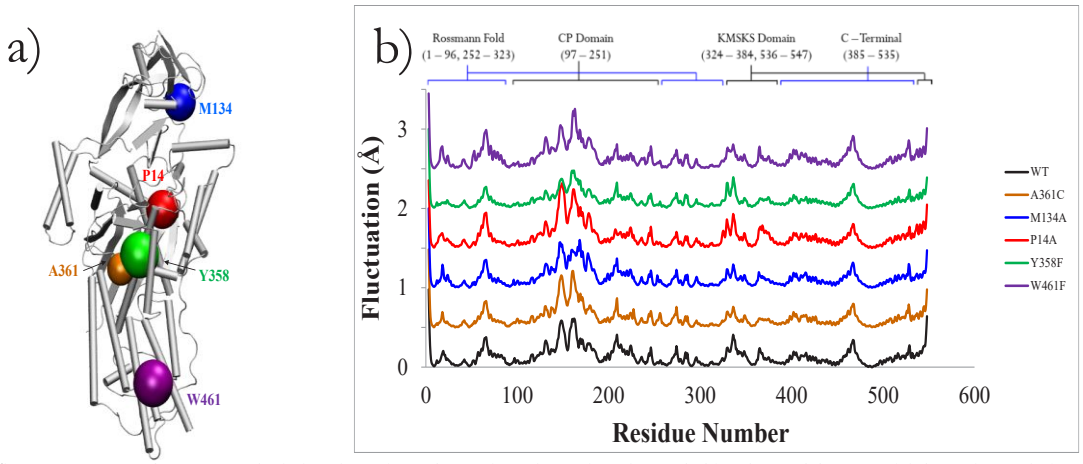

Figure 11. a) The 3D structure of Ec MetRS depicting the 5 sites of mutation, shown in color-coded beads, used for essential dynamics analysis. **b)** The replica-averaged root-mean-square fluctuations (RMSF) of individual amino acids of WT MetRS and the five variants. The RMSF of C_α atoms calculated from the time-averaged structures over the last 25 ns of MD trajectories are shown. The calculated propagated uncertainties are 0.034 Å for WT, 0.049 Å for P14A, 0.046 Å for M134A, 0.037 Å for Y358F, 0.046 Å for A361C, and 0.055 Å for W461F.

tRNA binding at the anticodon domain and plays an important role in tRNA aminoacylation by guiding the acceptor stem towards the active site.[24,33] Therefore, we hypothesized that a mutation of residues in the predicted pathways would cause a change in the CP domain dynamics. In the subsequent analysis, therefore, we investigated the dynamics of the CP domain (residues 97-251, shown in green in **Figure 1** in response to mutations of a few selected residues in the predicted paths using MD-PCA analysis.

MD simulations were performed in triplicate for the WT and 5 mutants, namely, P14A, M134A, Y358F, A361C, and W461F for a period of 30 ns. These mutational sites are evenly distributed along the communication pathways (**Figure 11a**). The four mutants - P14A, M134A, Y358F, and W461F, were experimentally found to have significant impact in MetRS function[24,65] (personal communication with Dr. R. W. Alexander). In addition, A361, which is only present in the NMA-derived pathway (Pathway 3) and was experimentally observed to have slight impact on the catalysis, was also chosen for this study.[66] The stability of the dynamics was first evaluated by computing the RMSD of C_α atoms along the simulated time (**Figure 7**). A sharp change of RMSD was observed in the initial 500-700 ps. Essential dynamics analysis were conducted on the 5-30 ns data, where the RMSDs were within 1.0 Å. The last 25 ns of MD simulation data were used to assess the quality of simulations by computing the RMSF of each amino acid from the time-averaged structure. The RMSF for each replica, as well as the replica-averaged fluctuations of the WT and the five MetRS variants, are reported in **Figure 12** and **Figure 11b**, respectively. The RMSF data demonstrates that the backbone flexibilities are quite reproducible for each of these protein systems, with only a propagated uncertainty of 0.03−0.06 Å for the three replica simulations (**Figure 11b**). These results indicate that all simulations have reached equilibrated states.

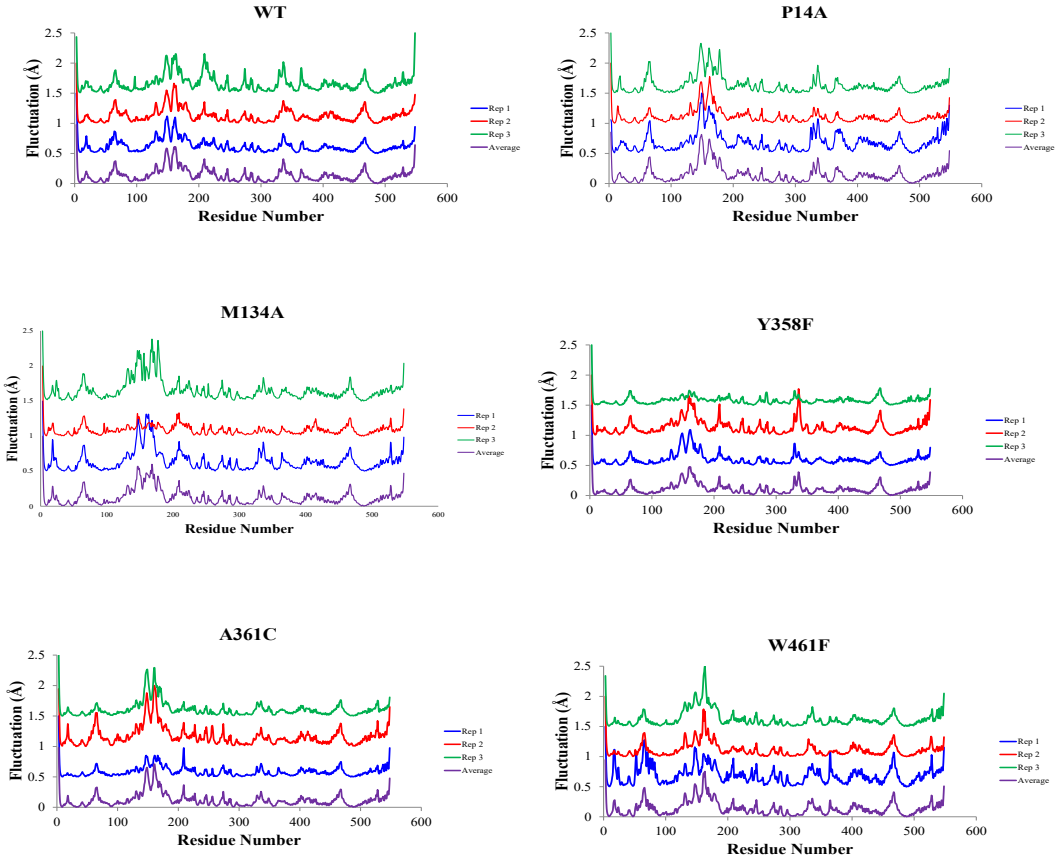

Figure 12. Root-mean-square fluctuations of individual amino acids for WT MetRS and five variants. In each stacked plot, the rms fluctuations of C_α atoms calculated from the time-averaged structures over the last 25 ns of MD trajectories are shown. Fluctuations for the three replicas are separated by 0.5 Å and color-coded for the sake of clarity: blue for replica 1, red for replica 2, and green for replica 3. In each case, the bottom plot (purple) represents the replica-averaged rms fluctuations. The calculated propagated uncertainties are 0.034 Å for WT, 0.049 Å for P14A, 0.046 Å for M134A, 0.037 Å for Y358F, 0.046 Å for A361C, and 0.055 Å for W461F.

Next, we probed the impact of these mutations on the collective dynamics of the Ec MetRS. Following the method described in our earlier work,[35] PCA was conducted using the combined 75 ns trajectory (last 25 ns of three replicates) for each protein system. The first three clusters representing the predominant collective dynamics were extracted. The RMSF of C_α atoms were computed from their respective average structures, normalized, and averaged over the three clusters.[35] The impact of mutation of on-pathway residues on the protein flexibility was examined by computing the difference of these cluster-averaged RMSF between the WT and a specific MetRS variant (**Figure 13**). The RMSF analysis demonstrates that the overall flexibility of the protein backbone was altered due to the mutation of on-pathway residues.

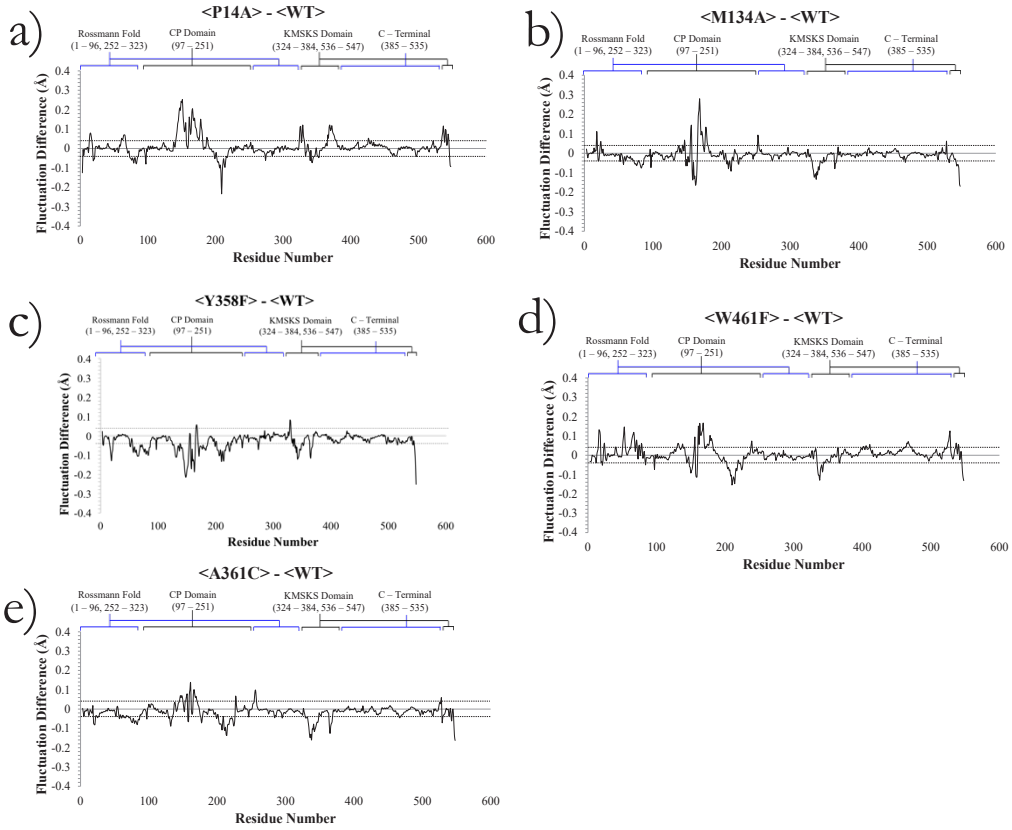

Figure 13. Changes in the normalized RMSF of C_α atoms observed in the collective dynamics of the five mutants with respect to WT Ec MetRS. The angular bracket indicates that the RMSF are averaged over the first three clusters representing the predominant collective protein motions. The propagated uncertainties for each of these plots are within 0.04 Å and are shown with two parallel dotted lines.

The RMSF analysis revealed that mutations of these "on-pathways" residues have altered the dynamics of the CP domain and other structural elements of this multi-domain protein. For example, alanine substitution of P14 resulted in increased dynamics of residues 324 to 336 and residues 364 to 378 of KMSKS domain (**Figure 13a**); residues of KMSKS domain are important for stabilizing methionyl-adenylate. Also, an increase in the CP domain dynamics was observed. However a sharp decrease in the fluctuation of the residue 209 of the CP domain was noticed in P14A variant.

Mutation of M134 to alanine also demonstrated an increase in CP domain dynamics (**Figure 13b**). The RMSF analysis indicates decreased dynamics of the catalytically important Rossmann-fold domain. However, a small increase in flexibility of W253 of the methionine binding pocket was noticed. Decreased dynamics of the KMSKS domain and C-terminal domain was also observed. In the case of Y358F, an overall decrease in the dynamics of virtually every domain was observed (**Figure 13c**). On the other hand, both increased and decreased fluctuations of various domains were noticed in the case of W461F (**Figure 13d**). Lastly, in the case of A361C mutant, where small impact in catalysis was experimentally observed, the fluctuation of the protein backbone was also altered by the alanine substitution (**Figure 13e**). Taken together, significant alternations in the dynamics of various structural elements were observed for these MetRS variants. These computational analyses demonstrated that mutations along the predicted pathways between the C-terminal and CP domain of MetRS do indeed alter the backbone fluctuations of the distant domains and secondary elements, which might resulted in reduced catalytic efficiency.

DISCUSSION

Protein dynamism and evolvability

Dynamism is an intrinsic property of a protein, an inseparable element from its function, and encrypted in its primary structure. Therefore, amino acid residues that are critical for maintaining the protein's intrinsic functional dynamics should either remain conserved or their mutations will be correlated. The present study revealed the existence of multiple interaction networks, involving dynamically and evolutionarily coupled residues, between the CP domain and the C-terminal domain. Scrutiny of these networks of interactions showed that the involved residues are predominantly conserved in this protein family (**Table 3**). Only a small fraction of residues (10-26%, **Tables 3 and 4**) were found to be coevolving and these residues appeared to be important for forming a contiguous network of residue interactions between W461 and M134/F140. Thus, the present findings demonstrated that conserved residues are the key players in regulating distant domain dynamics, while the role of coevolved residues is only complementary. This is consistent with the observations in myosin motor protein[67] and other protein systems[19,35] indicating that integration of evolutionary information with dynamic coupling data is an important criterion for identifying interaction networks between distant functional sites.

Cooperative dynamics between the two functional sites

Previous biochemical and computational studies have established that the cognate anticodon binding at the C-terminal triggers the conformational change of the CP domain, which facilitates the binding of tRNA acceptor stem at the active site.[24,33] Moreover, it has also been reported that the anticodon-triggered conformational change, which is important for efficient $tRNA^{Met}$ aminoacylation, propagates through the enzyme, not through the tRNA.[33] Therefore, if the identified residue networks facilitate coupled-domain dynamics, then a mutation along these predicted pathways is expected to have an observable impact on the dynamics of the CP domain. This hypothesis was tested by the combined essential dynamics analysis performed on various mutant protein systems. The four mutants (P14A, M134A, Y358F, and W461F), which are resided on the residue interaction networks between C-terminal and CP domains, were found to have significant impact on catalysis. Also, mutation of A361 to cysteine has impact on the distant domain dynamics. The pair-wise RMSF comparisons of the WT and mutant variants portray the notable effects of discrete mutations along the interaction networks on the distant CP domain dynamics (**Figure 13**). The variation in the CP domain dynamics is somewhat related to the distance between the site of mutation and the CP domain. Taken together, the *in silico* mutational study illustrates the role of predicted interaction networks in maintaining the distant domain dynamics.

Existing theoretical and experimental results

The residue interaction networks (pathways I –VI) identified in the present bioinformatics study also bears a close similarity to previously reported communication pathways (between W461 to L13) obtained from the atomistic simulations and protein structure networks (PSN) analysis.[33,68] However, unlike the previously reported pathways, pathways identified in the present study are contiguous. The involvement of the α-helix bundle, which encompasses pathways I and II, as well as the helix-loop-strand-helix motif (pathways III – VI) in domain-domain communication of Ec MetRS has also been supported by earlier studies.[66] Therefore, the revelation of similar residue-residue interaction networks, as obtained from the use of two different strategies – i) MD and PSN[33] and ii) MD/NMA and SCA, strongly suggests that the bioinformatics-based STCA method could be used as an alternative, fast yet robust method of predicting long-range communication pathways in multi-domain proteins.

Results obtained in the present study also enable us to explain some of the previously reported experimental mutational results. Previous mutational studies have showed that the mutation of P460, N452, R395, N391 and R233 cause functional defects.[24] These residues actually belong to the predicted interaction networks and appeared to be important for promoting coupled-domain dynamics, which are essential for enzymatic function. Also, it has been reported that the mutation of N452 and N387 to alanine resulted in functionally defective mutants.[69] Our studies show that these two residues are within the range of H-bonding interaction with K388, the residue present in three of the six pathways.

Coarse-grained NMA and all-atom MD simulations

Coarse-grained simulations have emerged as valuable tools for studying conformational changes in large biomolecules. Despite the fact that NMA implies infinitesimal displacements near the local energy minimum, recent studies have demonstrated that micro- to millisecond global motions can be modeled successfully with coarse-grained model.[70] We have also demonstrated that coarse-grained NMA is comparable to the all-atom MD simulation in depicting the intrinsic global dynamics of AARSs including Ec MetRS.[64,71] However, local fluctuations upon substrate binding were poorly captured by the coarse-grained simulations.[71] Although coarse-grained NMA failed to capture time-dependent fluctuations, it has been well-documented that NMA can provide relevant information regarding functional motions and allosteric mechanisms.[22,72-75] Surprisingly, very similar results were obtained from both all-atoms MD simulation and coarse-grained NMA in the present study; four out of six pathways identified by MD and NMA are almost identical (**Tables 3 and 4**). The close similarities between the all-atom and coarse-grained simulations results

suggested that for large biomolecules like AARSs, computationally less intensive coarse-grained NMA could be used to trace the conformational transition ("population-shift") pathways i.e. residue interaction networks between functional sites.

CONCLUSIONS

In the present study, pathways of inter-domain communication in Ec MetRS were identified by integrating evolutionary information with that of dynamic coupling. Several pathways (residue interaction networks) were identified through which local perturbation could propagate between two functional sites, 53 Å apart. Residues identified in these pathways are predominantly conserved and are also physically proximate in the structure. These residues are engaged in strong correlated dynamics. Long-duration MD simulation study followed by essential dynamics analysis provided evidence that these residues are important in maintaining protein dynamics and their mutations are capable of altering the dynamics of protein segments over great distances, even across domain interfaces. Therefore, the present STCA method shows promise in identifying and exploring residues that mediate long-range inter-domain communications in large protein systems.

The domain dynamics occur in micro- to millisecond timescale. Therefore, the atomistic simulation of a large protein is quite challenging. Although several hundred nanoseconds of simulations can be performed using the more advance computing systems, it remains quite challenging to simulate functionally important long-timescale collective dynamics of a large protein like Ec MetRS. Therefore, the use of coarse-grained method has substantially reduced the time for obtaining the information of thermal coupling from the collective dynamics of a large protein system like Ec MetRS. However, the coarse-grained simulation does not provide any information about the type of residues involved in coupled motions and statistical coupling analysis complements this by providing the data of residue-residue compatibility. Therefore integration of evolutionary information with dynamic features of residues could enable one to identify long-range interaction networks that are contiguous and critical for maintaining coupled dynamics in modular proteins.

Taken together, the present STCA study has enabled us to identify a set of residues that are potentially involved in maintaining the coupled dynamics among domains in Ec MetRS. The present study also suggested that in order to facilitate long-range communication, multi-domain proteins like Ec MetRS use parallel pathways of residue interactions. This information can be used as a guide to explore more about these interaction networks and the role of intrinsic dynamics in the function of Ec MetRS through spectroscopic, mutational, and theoretical studies. In addition, the STCA method can be extended to other multi-domain proteins to gain molecular-level understanding of the domain-domain communications. Moreover, the present study suggested that the correlated motions derived from NMA can also provide insight into long-range inter-domain communication.

ACKNOWLEDGEMENT

Authors thank Dr. Ranganathan (University of Texas, Southwestern Medical Center, Dallas) for the MATLAB scripts used in the SCA. We would also like to thank Mr. Alexander Jerome Greene for his initial contributions in developing the STCA script.

REFERENCES

1. Bu, Z., Biehl, R., Monkenbusch, M., Richter, D., and Callaway, D. J. (2005) Coupled protein domain motion in Taq polymerase revealed by neutron spin-echo spectroscopy, *Proc. Natl. Acad. Sci.* U. S. A. 102, 17646-17651.
2. Yu, H., Ma, L., Yang, Y., and Cui, Q. (2007) Mechanochemical coupling in the myosin motor domain. II. Analysis of critical residues, *PLoS Comput. Biol.* 3, e23.
3. Chennubhotla, C., Yang, Z., and Bahar, I. (2008) Coupling between global dynamics and signal transduction pathways: a mechanism of allostery for chaperonin GroEL, *Mol. Biosyst.* 4, 287-292.
4. Daily, M. D., and Gray, J. J. (2009) Allosteric communication occurs via networks of tertiary and quaternary motions in proteins, *PLoS Comput. Biol.* 5, e1000293.
5. Fidelak, J., Ferrer, S., Oberlin, M., Moras, D., Dejaegere, A., and Stote, R. H. (2010) Dynamic correlation networks in human peroxisome proliferator-activated receptor-gamma nuclear receptor protein, *Eur. Biophys. J.* 39, 1503-1512.
6. Zheng, W., Liao, J. C., Brooks, B. R., and Doniach, S. (2007) Toward the mechanism of dynamical couplings and translocation in hepatitis C virus NS3 helicase using elastic network model, Proteins 67, 886-896.
7. Weinkam, P., Pons, J., and Sali, A. (2012) Structure-based model of allostery predicts coupling between distant sites, *Proc. Natl. Acad. Sci.* U. S. A. 109, 4875-4880.
8. del Sol, A., Tsai, C. J., Ma, B., and Nussinov, R. (2009) The origin of allosteric functional modulation: multiple pre-existing pathways, *Structure* 17, 1042-1050.
9. Tsai, C. J., del Sol, A., and Nussinov, R. (2009) Protein allostery, signal transmission and dynamics: a classification scheme of allosteric mechanisms, *Mol. Biosyst.* 5, 207-216.
10. Changeux, J. P., and Edelstein, S. (2011) Conformational selection or induced fit? 50 years of debate resolved, F1000 *Biol. Rep.* 3, 19.

11. Lee, Y., Mick, J., Furdui, C., and Beamer, L. J. (2012) A coevolutionary residue network at the site of a functionally important conformational change in a phosphohexomutase enzyme family, *PLoS One* 7, e38114.
12. Nussinov, R., Tsai, C. J., and Ma, B. (2013) The underappreciated role of allostery in the cellular network, Annu. Rev. Biophys. 42, 169-189.
13. Tsai, C. J., and Nussinov, R. (2014) A unified view of "how allostery works", *PLoS Comput. Biol.* 10, e1003394.
14. Gunasekaran, K., Ma, B., and Nussinov, R. (2004) Is allostery an intrinsic property of all dynamic proteins?, *Proteins* 57, 433-443.
15. Tsai, C. J., del Sol, A., and Nussinov, R. (2008) Allostery: absence of a change in shape does not imply that allostery is not at play, *J. Mol. Biol.* 378, 1-11.
16. Popovych, N., Sun, S., Ebright, R. H., and Kalodimos, C. G. (2006) Dynamically driven protein allostery, *Nat. Struct. Mol. Biol.* 13, 831-838.
17. Daily, M. D., and Gray, J. J. (2007) Local motions is a benchmark of allosteric proteins, *Proteins* 67, 385-399.
18. Lockless, S. W., and Ranganathan, R. (1999) Evolutionarily conserved pathways of energetic connectivity in protein families, *Science* 286, 295-299.
19. Suel, G. M., Lockless, S. W., Wall, M. A., and Ranganathan, R. (2003) Evolutionarily conserved networks of residues mediate allosteric communication in proteins, *Nat. Struct. Biol.* 10, 59-69.
20. Hatley, M. E., Lockless, S. W., Gibson, S. K., Gilman, A. G., and Ranganathan, R. (2003) Allosteric determinants in guanine nucleotide-binding proteins, *Proc. Natl. Acad. Sci.* U. S. A. 100, 14445-14450.
21. Bahar, I., and Rader, A. J. (2005) Coarse-grained normal mode analysis in structural biology, *Curr. Opin. Struct. Biol.* 15, 586-592.
22. Skjaerven, L., Martinez, A., and Reuter, N. (2011) Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit, *Proteins* 79, 232-243.
23. Mechulam, Y., Schmitt, E., Maveyraud, L., Zelwer, C., Nureki, O., Yokoyama, S., Konno, M., and Blanquet, S. (1999) Crystal structure of Escherichia coli methionyl-tRNA synthetase highlights species-specific features, *J. Mol. Biol.* 294, 1287-1297.
24. Blanquet, S., Crepin, T., Mechulam, Y., and Schmitt, E. (2005) Methionyl-tRNA Synthetase, *Landes Biosciences/Eurekah.com*, Georgetown, TX
25. Martinis, S. A., and Schimmel, P. (1992) Enzymatic aminoacylation of sequence-specific RNA minihelices and hybrid duplexes with methionine, *Proc Natl Acad Sci* U S A 89, 65-69.
26. Martinis, S. A., and Schimmel, P. (1993) Microhelix aminoacylation by a class I tRNA synthetase. Non-conserved base pairs required for specificity, *J Biol Chem* 268, 6069-6072.
27. Alexander, R. W., Nordin, B. E., and Schimmel, P. (1998) Activation of microhelix charging by localized helix destabilization, *Proc. Natl. Acad. Sci.* U. S. A. 95, 12214-12219.
28. Meinnel, T., Mechulam, Y., Blanquet, S., and Fayat, G. (1991) Binding of the anticodon domain of tRNA(fMet) to Escherichia coli methionyl-tRNA synthetase, *J. Mol. Biol.* 220, 205-208.
29. Ghosh, G., Kim, H. Y., Demaret, J. P., Brunie, S., and Schulman, L. H. (1991) Arginine-395 is required for efficient in vivo and in vitro aminoacylation of tRNAs by Escherichia coli methionyl-tRNA synthetase, *Biochemistry* 30, 11767-11774.
30. Ghosh, G., Pelka, H., and Schulman, L. H. (1990) Identification of the tRNA anticodon recognition site of Escherichia coli methionyl-tRNA synthetase, *Biochemistry* 29, 2220-2225.
31. Meinnel, T., Mechulam, Y., Le Corre, D., Panvert, M., Blanquet, S., and Fayat, G. (1991) Selection of suppressor methionyl-tRNA synthetases: mapping the tRNA anticodon binding site, *Proc. Natl. Acad. Sci.* U. S. A. 88, 291-295.
32. Budiman, M. E., Knaggs, M. H., Fetrow, J. S., and Alexander, R. W. (2007) Using molecular dynamics to map interaction networks in an aminoacyl-tRNA synthetase, *Proteins* 68, 670-689.
33. Ghosh, A., and Vishveshwara, S. (2007) A study of communication pathways in methionyl- tRNA synthetase by molecular dynamics simulations and structure network analysis, *Proc. Natl. Acad. Sci.* U. S. A. 104, 15711-15716.
34. Weimer, K. M., Shane, B. L., Brunetto, M., Bhattacharyya, S., and Hati, S. (2009) Evolutionary basis for the coupled-domain motions in Thermus thermophilus leucyl-tRNA synthetase, *J. Biol. Chem.* 284, 10088-10099.
35. Johnson, J. M., Sanford, B. L., Strom, A. M., Tadayon, S. N., Lehman, B. P., Zirbes, A. M., Bhattacharyya, S., Musier-Forsyth, K., and Hati, S. (2013) Multiple pathways promote dynamical coupling between catalytic domains in Escherichia coli prolyl-tRNA synthetase, *Biochemistry* 52, 4399-4412.
36. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, J. Mol. Graph. 14, 33-38, 27-38.
37. Bahar, I., Atilgan, A. R., and Erman, B. (1997) Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential, *Fold Des.* 2, 173-181.
38. Eyal, E., Yang, L. W., and Bahar, I. (2006) Anisotropic network model: systematic evaluation and a new web interface, *Bioinformatics* 22, 2619-2627.
39. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., Kale, L., and Schulten, K. (2005) Scalable molecular dynamics with NAMD, *J. Comput. Chem.* 26, 1781-1802.

40. MacKerell, A. D. J., Bashford, D., Bellott, M., Dunbrack, R. L. J., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Gou, J., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E. I., Roux, B., Schelenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanbe, M., Wiórkiewicz-Kuczera, J., Yin, D., and Karplus, M. (1998) Allatom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* 102, 3586.

41. Glykos, N. M. (2006) Software news and updates. Carma: a molecular dynamics analysis program, *J. Comput. Chem.* 27, 1765-1768.

42. Brooks, B., and Karplus, M. (1983) Harmonic dynamics of proteins: normal modes and fluctuations in bovine pancreatic trypsin inhibitor, *Proc. Natl. Acad. Sci.* U. S. A. 80, 6571-6575.

43. Tama, F., and Sanejouand, Y. H. (2001) Conformational change of proteins arising from normal mode calculations, *Protein Eng.* 14, 1-6.

44. Dijkstra, E. W. (1959) A note on two problems in connexion with graphs, *Numerische Mathematik* 1, 269-271.

45. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* 25, 3389-3402.

46. Yang, L. W., and Chng, C. P. (2008) Coarse-grained models reveal functional dynamics--I. Elastic network models--theories, comparisons and perspectives, *Bioinform Biol Insights* 2, 25-45.

47. Skjaerven, L., Hollup, S., and Reuter, N. (2009) Normal mode naalysis for protein, J. Mol. Chem. THEOCHEM 898, 42-48.

48. Ma, J. (2005) Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes, *Structure* 13, 373-380.

49. Atilgan, A. R., Durell, S. R., Jernigan, R. L., Demirel, M. C., Keskin, O., and Bahar, I. (2001) Anisotropy of fluctuation dynamics of proteins with an elastic network model, *Biophys J* 80, 505-515.

50. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79, 926.

51. Ryckaert, J. P., Ciotti, G., and Berensden, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.* 23, 327-341.

52. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald: An N. Log(N) Method for Ewald Sums in Large Systems, *J. Chem. Phys.* 98, 10089-10092.

53. Berendsen, H. J. C., Postma, J. P., van Gunsteren, M. W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81, 3684-3690.

54. Feller, S. E., Zhang, Y., Pastor, R. W., and Brooks, B. R. (1995) Constant pressure molecular dynamics simulation: The Langevin piston method, *J. Chem. Phys.* 103, 4613-4621.

55. Martyna, G. J., Tobias, D. J., and Klein, M. L. (1994) Constant pressure molecular dynamics algorithms, *J. Chem. Phys.* 101, 4177-4189.

56. Bhattacharyya, S., Ma, S., Stankovich, M. T., Truhlar, D. G., and Gao, J. (2005) Potential of mean force calculation for the proton and hydride transfer reactions catalyzed by medium-chain acyl-CoA dehydrogenase: effect of mutations on enzyme catalysis, *Biochemistry* 44, 16549-16562.

57. Rauschnot, J. C., Yang, C., Yang, V., and Bhattacharyya, S. (2009) Theoretical determination of the redox potentials of NRH:quinone oxidoreductase 2 using quantum mechanical/molecular mechanical simulations, *J. Phys. Chem. B* 113, 8149-8157.

58. Sanford, B., Cao, B. V., Johnson, J. M., Zimmerman, K., Strom, A. M., Mueller, R. M., Bhattacharyya, S., Musier-Forsyth, K., and Hati, S. (2012) Role of coupled-dynamics in the catalytic activity of prokaryotic-like prolyl-tRNA synthetases *Biochemistry* 51, 2146-2156.

59. Roy, J., and Laughton, C. A. (2010) Long-timescale molecular-dynamics simulations of the major urinary protein provide atomistic interpretations of the unusual thermodynamics of ligand binding, *Biophys. J.* 99, 218-226.

60. van Aalten, D. M., Amadei, A., Linssen, A. B., Eijsink, V. G., Vriend, G., and Berendsen, H. J. (1995) The essential dynamics of thermolysin: confirmation of the hinge-bending motion and comparison of simulations in vacuum and water, *Proteins* 22, 45-54.

61. Mueller, R. M., North, M. A., Yang, C., Hati, S., and Bhattacharyya, S. (2011) Interplay of flavin's redox states and protein dynamics: an insight from QM/MM simulations of dihydronicotinamide riboside quinone oxidoreductase 2, *J. Phys. Chem. B* 115, 3632-3641.

62. Silvian, L. F., Wang, J., and Steitz, T. A. (1999) Insights into editing from an ile-tRNA synthetase structure with tRNAile and mupirocin, *Science* 285, 1074-1077.

63. Tukalo, M., Yaremchuk, A., Fukunaga, R., Yokoyama, S., and Cusack, S. (2005) The crystal structure of leucyl-tRNA synthetase complexed with tRNALeu in the post-transfer-editing conformation, *Nat. Struct. Mol. Biol.* 12, 923-930.

64. Warren, N., Strom, A., Nicolet, B., Albin, K., Albrecht, J., Bausch, B., Dobbe, M., Dudek, M., Firgens, S., Fritsche, C., Gunderson, A., Heimann, J., Her, C., Hurt, J., Konorev, D., Lively, M., Meacham, S., Rodriguez, V., Tadayon, S., Trcka, D.,

Yang, Y., Bhattacharyya, S., and Hati, S. (2014) Comparison of the Intrinsic Dynamics of Aminoacyl-tRNA Synthetases, *Protein J.* 33, 184-198.
65. Alexander, R. W., and Schimmel, P. (2001) Domain-domain communication in aminoacyl-tRNA synthetases, Prog. *Nucleic Acid Res. Mol. Biol.* 69, 317-349.
66. Banerjee, P., Warf, M. B., and Alexander, R. (2009) Effect of a domain-spanning disulfide on aminoacyl-tRNA synthetase activity, *Biochemistry* 48, 10113-10119.
67. Tang, S., Liao, J. C., Dunn, A. R., Altman, R. B., Spudich, J. A., and Schmidt, J. P. (2007) Predicting allosteric communication in myosin via a pathway of conserved residues, *J Mol Biol* 373, 1361-1373.
68. Ghosh, A., and Vishveshwara, S. (2008) Variations in clique and community patterns in protein structures during allosteric communication: investigation of dynamically equilibrated structures of methionyl tRNA synthetase complexes, *Biochemistry* 47, 11398-11407.
69. Alexander, R. W., and Schimmel, P. (1999) Evidence for breaking domain-domain functional communication in a synthetase-tRNA complex, *Biochemistry* 38, 16359-16365.
70. Gur, M., Zomot, E., and Bahar, I. Global motions exhibited by proteins in micro- to milliseconds simulations concur with anisotropic network model predictions, *J Chem Phys* 139, 121912.
71. Strom, A. M., Fehling, S. C., Bhattacharyya, S., and Hati, S. (2014) Probing the global and local dynamics of aminoacyl-tRNA synthetases using all-atom and coarse-grained simulations, *J. Mol. Model.* 20, 2245.
72. Cui, Q., Li, G., Ma, J., and Karplus, M. (2004) A normal mode analysis of structural plasticity in the biomolecular motor F(1)-ATPase, *J. Mol. Biol.* 340, 345-372.
73. Mouawad, L., and Perahia, D. (1996) Motions in hemoglobin studied by normal mode analysis and energy minimization: evidence for the existence of tertiary T-like, quaternary R-like intermediate structures, J Mol Biol 258, 393-410.
74. Ma, J., and Karplus, M. (1998) The allosteric mechanism of the chaperonin GroEL: a dynamic analysis, *Proc. Natl. Acad. Sci.* U. S. A. 95, 8502-8507.
75. Zheng, W., Brooks, B. R., and Thirumalai, D. (2007) Allosteric transitions in the chaperonin GroEL are captured by a dominant normal mode that is most robust to sequence variations, *Biophys. J.* 93, 2289-2299.

ABOUT THE STUDENT AUTHOR

Ryan Andrews recently graduated from the University of Wisconsin – Eau Claire and is currently a PhD student at Iowa State University in the department of Biochemistry, Biophysics and Molecular Biology where he continues to do similar research he found a passion for while an undergraduate.

PRESS SUMMARY

Proteins are the machinery of all living cells, and research is conducted every day to further our understanding of how they work. This paper compares two computational methods' ability to study how different parts, or domains, of an individual protein "communicate" with each other, a phenomenon known as inter-domain communication. We found that the less computationally demanding process known as coarse-grained analysis was comparable to the more demanding (though more theoretically accurate) process known as atomistic molecular dynamics in investigating inter-domain communications.

Skewed and Flexible Skewed Distributions: A Modern Look at the Distribution of BMI

Thao Tran, Cara Wiskow, and Mohammad Abdus Aziz*
Department of Mathematics, University of Wisconsin–Eau Claire, Eau Claire, WI

Students: tranp@uwec.edu, wiskowca@uwec.edu*

Mentor: azizm@uwec.edu

ABSTRACT
The purpose of this study is to find distributions that best model body mass index (BMI) data. BMI has become a standard health indicator and numerous studies have been done to examine the distribution of BMI. Due to the skew and bimodal nature, we focus on modeling BMI with flexible skewed distributions. The distributions are fitted to University of Wisconsin–Eau Claire (UWEC) BMI data and to data obtained from National Health and Nutrition Survey (NHANES). The model parameters are obtained using maximum likelihood estimation method. We compare flexible models to more conventional distributions, such as skew-normal and skew-t distributions, using AIC and BIC and Kolmogorov-Smirnov (K-S) goodness-of-fit test. Our results indicate that the skew-t and Alpha-Skew-Laplace distributions are reasonably competitive when describing unimodal BMI data whereas Alpha-Skew-Laplace, finite mixture of scale mixture of skew-normal and skew-t distributions are better alternatives to both unimodal and bimodal conventional distributions. The results we obtained are useful because we believe the models discussed in our study will offer a framework for testing features such as bimodality, asymmetry, and robustness of the BMI data, thus providing a more detailed and accurate understanding of the distribution of BMI.

KEYWORDS
Body Mass Index; Skew-normal distribution; Skew-t distribution; Flexible skewed distributions; Mixture distributions; Scale mixture of skew-normal distribution; K-S test

INTRODUCTION

Obesity has been reaching epidemic proportions in the United States. The rise has implications both on health and health care costs. Body mass index (BMI; kg/m^2) in the overweight (25 to 29.9) or obese category (30 or above) has been linked to cancer, hypertension, heart failure, cardiovascular disease, diabetes, stroke, and more. Because of this, obesity places a burden on the health care system, raising costs for the public. Given these negative impacts, governments and organizations have been actively trying to reduce obesity. In order to better assess obesity risk and address its prevalence, a better understanding of the distribution of BMI is imperative.

BACKGROUND

Since obesity became a major public health concern, many distributions have been applied to BMI data in an effort to find the best way to describe it. Multiple probability distributions have been fitted to Australian athletes' BMI data. Examples for the univariate case include Ma et al.'s generalized skew-normal (GSN) distribution,[1] Canale's extended skew-normal (ESN),[2] and Olivares-Pacheco et al.'s epsilon-skew-slash (ESS) and epsilon-skew-normal (ESN)[3]. For multivariate analysis of the same data set, Tan et al. considered the skew-slash t (SSLT) and skew-slash-normal (SSLN)[4] distributions and Arslan[5] applied generalized hyperbolic (GH) and variance-mean mixture of skew-normal (SN) distributions. When analyzing potential effectiveness of a tax on sweetened beverages in South Africa, Manyema et al.[6] applied log-normal and gamma distributions to describe the BMI distribution in their data set. For a US data set, Miljkovic et al.[7] found the k-component

Gaussian mixture distribution best fit their BMI data when compared to the log-normal, Weibull, logistic, inverse Gaussian, and gamma distributions.

While plenty of work has been done to examine unimodal BMI distributions, little has been done concerning bimodal BMI data. Mixture distributions provide a way to model bimodal data well. Lin et al.[8] investigated the potential for describing a bimodal BMI data set using mixture of normal, mixture of skew-normal, mixture of student's t, and mixture of skew-t distributions.

In this paper, we explore the abilities of various distributions to model both UWEC BMI and NHANES BMI data. Recent developments have been made in distributions which can account for either unimodality or bimodality of skewed data. They comprise alpha skew-normal, a new family of distributions, introduced by Elal-Olivero, that is flexible to both unimodal and bimodal data.[9] In 2012, Harandi et al.[10] presented a new class of skew distributions of Elal-Olivero's family using Laplace distribution, known as Alpha-Skew-Laplace distribution. This distribution can fit unimodal and bimodal shapes with increasing and u-shaped hazard functions for the truncated case at the origin. In 2014, the simple approach was used with logistic functions by Hazarika et al.,[11] who proposed a new distribution called alpha-skew-logistic distribution. This distribution can especially model the data given either positive or negative skewness. Another recently proposed model is the bimodal skew-symmetric normal distribution by Hassan et al.,[12] which can resolve problems of asymmetry, platykurtic/leptokurtic data (exhibit excess negative/positive kurtosis), and different types of bimodality. Kollu et al. also introduced three new mixture models, including Weibull-log-normal, GEV-log-normal, and Weibull-GEV.[13] We apply some of these newly introduced distributions as well as other mixture distributions to two BMI data sets in order to see if they provide a more accurate description of BMI distributions.

Significance

In recent years, there has been considerable interest in skewed distributions,[14] and considering the skew and bimodal features of some of the medical data, it is really important that appropriate distributions be fitted to these kinds of data. The use of statistical procedures is inappropriate if the actual distribution differs from the assumed type. This study mainly focuses on a statistical practice of fitting skew-symmetric distributions to medical data with bimodal characteristics while still considering unimodal and mixture distributions.

The most commonly used techniques for modeling bimodality involve using the mixture distributions. However, the proposed models created computational implementation difficulties. Many authors have also proposed different versions of the bimodal normal distribution to replace mixture distributions but because they fail to take into account the asymmetry, these studies did not materialize in the real world of statistics.

Many medical data sets, including that of BMI, are bimodal, asymmetric and platykurtic, or leptokurtic. We believe that the models we consider in this study will offer a framework for testing these features of the data at hand and overcome some of the complexities of the existing models. If these features are found to be significant, the proposed distribution will provide the user with a parsimonious model that will fit the data adequately. Thus, the user can test symmetry, excess kurtosis, and bimodality in order to adjust the values of model parameters accordingly to ensure a good fit.

The uniqueness and skewness of the data make it difficult to model perfectly with many known distributions, but the flexibility of the skew-symmetric distribution offers an increasingly insightful perspective that could offer a solution in dealing with bimodality of data in the field of medicine.

METHODS AND PROCEDURES

Flexible Skew-Symmetric Distributions

Unimodal Skew-Symmetric Distributions

1. Skew-normal distribution

   A skew-normal distribution has the following probability density function (pdf)

$$f(y;\alpha) = 2\phi(y)\Phi(\alpha y),\ y \in R \qquad \textbf{Equation 1.}$$

   where ϕ represents the density distribution of $N(0,1)$, Φ represents the cumulative distribution of $N(0,1)$, and α represents the shape parameter. We use the notation $SN(\alpha)$ to denote a skew-normal random variable.
   The linear transformation $X = \mu + \alpha Y$ with $\mu \in R$ and $\sigma > 0$ has the density of

$$f(x;\mu,\sigma,\alpha) = \frac{2}{\sigma}\phi\Big(\frac{x-\mu}{\sigma}\Big)\Phi\Big(\alpha\frac{x-\mu}{\sigma}\Big),\ x \in R. \qquad \textbf{Equation 2.}$$

   Then $X \sim SN(\mu,\sigma,\alpha)$, which reduces to the standard skew-normal distribution when $X \sim SN(0,1,\alpha)$. If α is set to 0, the distributions become the pdf of a standard normal distribution.
   The skew-normal cumulative distribution function is given by the following equation:

$$\Phi(z,\lambda) = 2\int_{-\infty}^{z}\int_{-\infty}^{\lambda t}\phi(t)\phi(u)\,dudt. \qquad \textbf{Equation 3.}$$

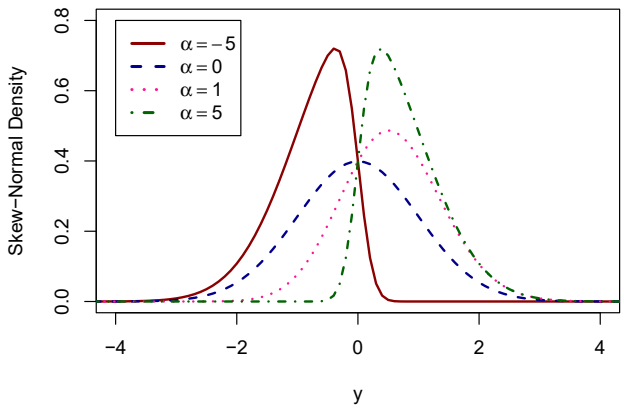

Figure 1. Density plot of skew-normal distribution for some selected values of α.

2. Skew-t distribution

   Let Z be a standard skew-normal random variable and W be a variable with $\chi^2(\nu)$ distribution. Suppose Z and W are independent. Define

$$Y = \frac{Z}{\sqrt{\frac{W}{\nu}}}. \qquad \textbf{Equation 4.}$$

   Then the linear transformation $X = \mu + \sigma Y$ has a skew-t distribution with parameters μ, σ, α, and ν and we introduce the notation $ST(\mu,\sigma,\alpha,\nu)$ to denote the skew-t random variable X.

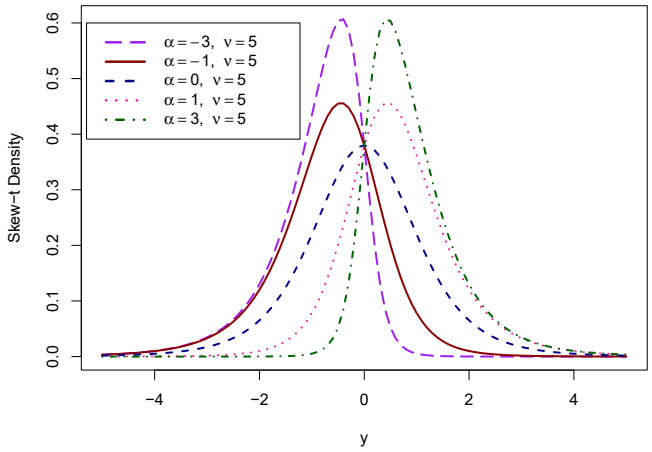

Figure 2. Density plot of skew-t distribution for some selected values of α and ν.

3. Generalized extreme value distribution

   The random variable X is said to have generalized extreme value (GEV) distribution when X has the following density function

$$e(x;\zeta,\delta,l) = \left(\frac{l}{\delta}\right)\left(1+\frac{\zeta(x-1)}{\delta}\right)^{-\frac{1}{\zeta}-1} e^{-\left(1+\frac{\zeta(x-1)}{\delta}\right)^{\frac{1}{\zeta}}} \qquad \textbf{Equation 5.}$$

   where $\zeta \neq 0$. The cumulative distribution of X is given by

$$E(x;\zeta,\delta,l) = e^{-\left(1+\frac{\zeta(x-1)}{\delta}\right)^{\frac{-1}{\delta}}}. \qquad \textbf{Equation 6.}$$

We do not provide the mathematical description of traditional distributions such as Gamma, Weibull, and log-normal. Their pdfs can be found in any standard book on distributions, for example, Balakrishnan and Johnson.[15]

Bimodal Skew-Symmetric Distributions

1. Alpha-skew-normal distribution

   A continuous random variable Y has an alpha-skew-normal distribution with a probability density function

$$f(y;\alpha) = \frac{(1-\alpha y)^2+1}{2+\alpha^2}\phi(y),\ y \in R \qquad \textbf{Equation 7.}$$

   where α represents the shape parameter. We denote this density as $Y \sim ASN(\alpha)$. If we adjust the pdf to include location and scale parameters the density becomes

$$f(y;\mu,\sigma,\alpha) = \frac{\left[1-\alpha\left(\frac{y-\mu}{\sigma}\right)\right]^2+1}{\sigma\left(2+\alpha^2\right)}\phi\left(\frac{y-\mu}{\sigma}\right),\ y \in R. \qquad \textbf{Equation 8.}$$

   Alpha-skew-normal has cumulative distribution function

$$F(y) = \Phi(y) + \alpha\left(\frac{2-\alpha y}{2+\alpha^2}\right)\phi(y). \qquad \textbf{Equation 9.}$$

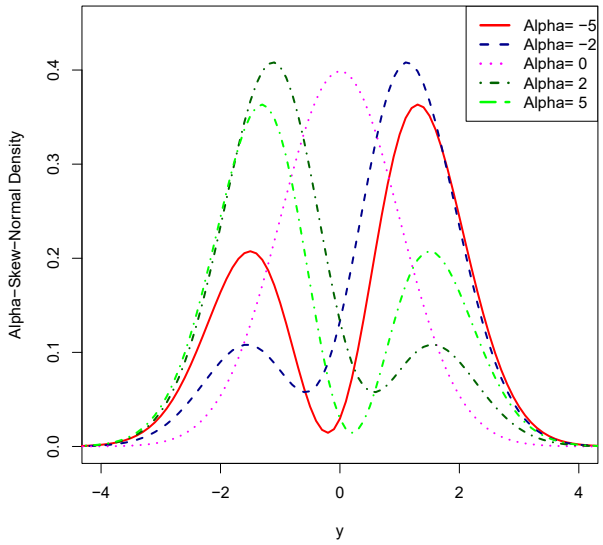

Figure 3. Density plot of alpha-skew-normal distribution for some selected values of α.

2. Alpha-skew-logistic distribution

   Let Y be an alpha-skew-logistic random variable with parameter α then Y has density function

$$f(y;\alpha) = \frac{3((1-\alpha y)^2+1)e^{-y}}{(6+(\alpha^2\pi^2))(1+e^{-y})^2}. \qquad \textbf{Equation 10.}$$

   We denoted it by $y \sim ASLG(\alpha)$. If α equals 0, we get the standard logistic distribution given by

$$f_y(y) = \frac{e^{-y}}{(1+e^{-y})^2}, \quad y \in R. \qquad \textbf{Equation 11.}$$

   The Alpha-skew-logistic cumulative distribution function is given by

$$F(y;\alpha) = \frac{3}{6+\alpha^2\pi^2}\left(\frac{(1-\alpha y)^2+1}{1+e^{-y}} + 2\alpha(1-\alpha y)\log(1+e^{y}) - 2\alpha^2 Li_2(-e^{y})\right). \qquad \textbf{Equation 12.}$$

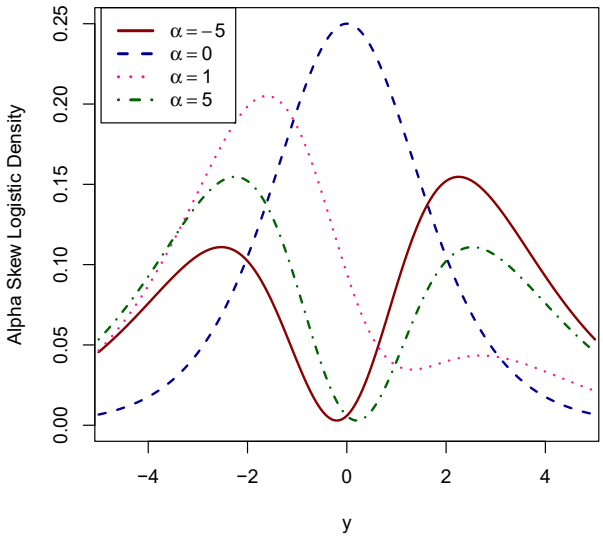

Figure 4. Density plot of Alpha-skew-logistic distribution with chosen α values -5, 0, 1, 5.

3. Alpha-Skew-Laplace distribution

A continuous random variable Y is said to follow an Alpha-Skew-Laplace distribution if its pdf has the form

$$f(y) = \frac{(1-\alpha y)^2+1}{4(1+\alpha^2)} e^{-|y|},\ y \in R \qquad \textbf{Equation 13.}$$

where α represents the shape parameter. An Alpha-Skew-Laplace random variable is denoted by $ASLP(\alpha)$.

Suppose $Y \sim ASLP(\alpha)$. Then ASLP density of location and scale is defined as the distribution of $X = \mu + \sigma Y$ for $\mu \in R$ and $\sigma > 0$. The corresponding density function is given by

$$f(x) = \frac{(1-\alpha\frac{x-\mu}{\sigma})^2+1}{4\sigma(1+\alpha^2)} e^{-\frac{|x-\mu|}{\sigma}},\ x \in R \qquad \textbf{Equation 14.}$$

where $\theta = (\mu, \sigma, \alpha)$.

Alpha-Skew-Laplace cumulative distribution is given by the following.

$$F(t;\alpha) = \frac{1+(1-\alpha t)^2}{4(1+\alpha^2)} e^{t} + \frac{\alpha(1+\alpha(1-t))}{2(1+\alpha^2)} e^{t},\ t < 0 \qquad \textbf{Equation 15.}$$

$$F(t;\alpha) = 1 - \frac{1+(1-\alpha t)^2}{4(1+\alpha^2)} e^{-t} + \frac{\alpha(1+\alpha(1-t))}{2(1+\alpha^2)} e^{-t},\ t \geq 0$$

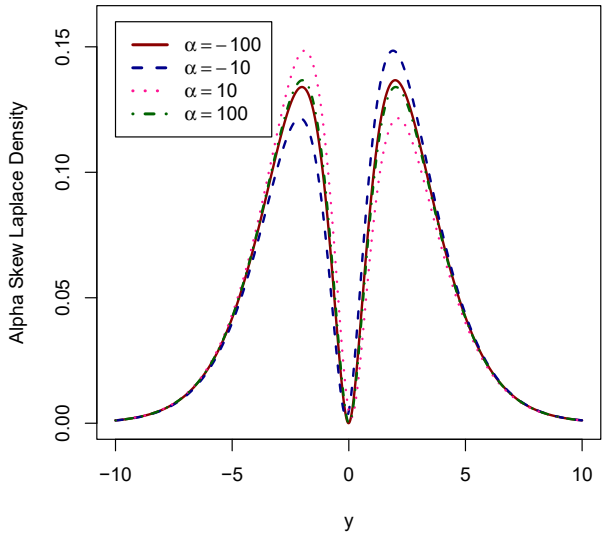

Figure 5. Density plot of Alpha-Skew-Laplace Distribution with chosen α values -5, 0, 1, 5.

4. Bimodal skew-symmetric normal distribution

   The random variable Y is said to have a bimodal skew-symmetric normal distribution when Y has the density distribution as follows

$$\Psi(y) = \Phi(y) - \frac{y + \mu - 2\beta}{1 + 2\psi(\delta + (\beta - \mu)^2))}\phi(y) \qquad \textbf{Equation 16.}$$

   where $\mu, \beta \in R$ represents the location parameter, $\psi > 0$ represents the shape parameter, and $\theta > 0$ represents the bimodality parameter.

Mixture Distributions

In this section we consider several conventional mixture distributions and recently developed scale mixture of skew-normal and skew-t distributions. In the mixture model context the density of x is expressed as a mixture of P parametric densities such that

$$f(x, \psi) = \sum_{i=1}^{p} \pi_i f(x; \theta_i). \qquad \textbf{Equation 17.}$$

The $\pi_i \geq 0, i = 1, 2, \ldots, p$ with $\sum_{i=1}^{p} \pi_i = 1$ are called mixing weights of the ith component of the mixture, which is characterized by parameter θ_i, and $\psi = (\pi_1, \pi_2, \ldots, \pi_{p-1}, \theta_1, \theta_2, \ldots, \theta_p)$ denotes the vector of parameters of the model.

1. Finite mixture of scale mixture of skew-normal distribution

   Suppose $Z \sim SN(0, \sigma^2, \alpha)$ and U be a positive random variable, independent of Z, with distribution function $H(u; \nu)$. Then the random variable $Y = \mu + U^{-1}Z$, where $\mu \in R$ is a location parameter, is said to follow a scale mixture of skew-normal (SMSN) distribution if its pdf is given by

$$f(y) = \int_0^\infty \phi\left(y; \mu, \sigma^2 u^{-1}\right) \Phi\left(u^{1/2}\alpha\left(\frac{y - \mu}{\sigma}\right)\right) dH(u). \qquad \textbf{Equation 18.}$$

   In the definition $H(.; \nu)$ is known as the *mixing scale distribution* and for each choice of this we get different members of the family such as normal, skew-normal, or student-t. A finite mixture of SMSN distributions[7] model is a density

defined as in Equation 17 where the ith component of the mixture is a SMSN density with parameters μ_i, σ_i^2, α_i, and ν_i. For simplicity we assume $\nu_1 = \nu_2 = \ldots = \nu$.

2. Two-component mixture Weibull distribution

   The two-component mixture Weibull distribution has five parameters and its probability distribution function is given as follows:

$$ff(x; k_1, c_1, k_2, c_2, w) = wf(x; k_1, c_1) + (1 - w)f(x; k_2, c_2). \quad \textbf{Equation 19.}$$

   Its cumulative distribution function is given by

$$FF(x; k_1, c_1, k_2, c_2, w) = wF(x; k_1, c_1) + (1 - w)F(x; k_2, c_2). \quad \textbf{Equation 20.}$$

3. Mixture gamma and Weibull distribution

   A random variable X is said to have mixture gamma and Weibull distribution when X has the following probability distribution:

$$h(x; \alpha, \beta, k, c, w) = wg(x, \alpha, \beta) + (1 - w)f(x; k, c). \quad \textbf{Equation 21.}$$

   The mixture gamma and Weibull cumulative distribution function is given by

$$H(x; \alpha, \beta, k, c, w) = wG(x, \alpha, \beta) + (1 - w)F(x; k, c). \quad \textbf{Equation 22.}$$

4. Mixture normal distribution

   A single-truncated normal probability distribution function is given by

$$q(x; \mu, \sigma) = \frac{1}{I(\mu, \sigma)\sigma\sqrt{2\pi}} e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]} \quad \textbf{Equation 23.}$$

   for $x \geq 0$, where

$$I(\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^\infty e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}. \quad \textbf{Equation 24.}$$

   The cumulative distribution function of the single truncated normal distribution is given by

$$Q(x; \mu, \sigma) = \int_0^x \frac{1}{I(\mu, \sigma)\sigma\sqrt{2\pi}} e^{\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]}\, dx. \quad \textbf{Equation 25.}$$

   The mixture of two-component truncated normal distributions has the following probability density function:

$$r(x; \mu_1, \sigma_1, \mu_2, \sigma_2, w) = wq(x; \mu_1, \sigma_1) + (1 - w)q(x; \mu_2, \sigma_2). \quad \textbf{Equation 26.}$$

   Its cumulative distribution function is given by

$$R(x; \mu_1, \sigma_1, \mu_2, \sigma_2, w) = wQ(x; \mu_1, \sigma_1) + (1 - w)Q(x; \mu_2, \sigma_2). \quad \textbf{Equation 27.}$$

5. Mixture normal and Weibull distribution

   The mixture normal and Weibull distribution combines a single-truncated normal and a Weibull distribution, which has the following probability density function

$$s(x; \mu, \sigma, k, c) = wq(x; \mu, \sigma) + (1 - w)f(x; k, c). \quad \textbf{Equation 28.}$$

   The cumulative distribution is given by

$$S(x; \mu, \sigma, k, c) = wQ(x; \mu, \sigma) + (1 - w)F(x; k, c). \quad \textbf{Equation 29.}$$

6. Mixture Weibull and GEV distribution

   The probability density function of a mixture Weibull and GEV distribution is given by the following:

$$t(x;k,c,\zeta,\delta,l) = wf(x;k,c) + (1-w)e(x;\zeta,\delta,l). \quad \textbf{Equation 30.}$$

   The mixture Weibull and GEV cumulative distribution is given by

$$T(x;k,c,\zeta,\delta,l) = wF(x;k,c) + (1-w)E(x;\zeta,\delta,l). \quad \textbf{Equation 31.}$$

7. Mixture Weibull and log-normal distribution

   A random variable X has mixture Weibull and log-normal distribution when it has the following probability density function:

$$u(x;k,c,\lambda,\theta) = wf(x;k,c) + (1-w)l(x;\lambda,\theta). \quad \textbf{Equation 32.}$$

   Its cumulative distribution function is given by

$$U(x;k,c,\lambda,\theta) = wF(x;k,c) + (1-w)L(x;\lambda,\theta). \quad \textbf{Equation 33.}$$

8. Mixture GEV and log-normal distribution The probability density function of the mixture GEV and log-normal distribution is given by

$$v(x;\zeta,\delta,l,\lambda,\theta) = we(x;\zeta,\delta,l) + (1-w)l(x;\lambda,\theta). \quad \textbf{Equation 34.}$$

   The mixture GEV and log-normal distribution has the following cumulative distribution function:

$$V(x;\zeta,\delta,l,\lambda,\theta) = wE(x;\zeta,\delta,l) + (1-w)L(x;\lambda,\theta). \quad \textbf{Equation 35.}$$

Estimation Method

We estimate the parameters of all models using the maximum likelihood estimation method. For all models considered, in the first step, the log likelihood function is written using the probability density functions provided. As an example, the log likelihood function for the mixture of gamma and Weibull distribution can be written as:

$$LL = \sum_{i=1}^{n} \ln\{wg(x,\alpha,\beta) + (1-w)f(x;k,c)\}. \quad \textbf{Equation 36.}$$

In the second step, the GenSA function from R package 'GenSA' is used for optimization.

Selection Criteria

In order to assess the descriptive ability of multiple distributions, we use the following selection criteria: the Akaike Information Criterion, the Bayesian Information Criterion, and the Kolmogorov-Smirnov test.

Akaike Information Criterion (AIC)

AIC is an index to measure the fit of the model and compare the proposed models to other competitive models. It is defined as

$$\text{AIC} = 2k - 2\ln(L) \quad \textbf{Equation 37.}$$

where k is the number of estimable parameters, and $\ln(L)$ is the log-likelihood at its maximum point of the model estimated. While comparing different models, the one with the smallest AIC value is considered the best.

Bayesian Information Criterion (BIC)

BIC is another criterion for model comparison, which is defined as

$$\text{BIC} = \ln(n)k - 2\ln(L) \quad \textbf{Equation 38.}$$

where n is the sample size, k is the number of estimable parameters, and L is the maximum value of the likelihood function. Similar to AIC. the models with smaller BIC value are better than others.

Kolmogorov-Smirnov test (K-S test)

The K-S test is used to check the goodness of fit of a given set of data to a theoretical distribution $F(x)$. Suppose $X_1, X_2, \ldots, X_n$ is a random sample. The null hypothesis, H_o, gives a theoretical distribution function, F_o. The K-S test compares $F_o(x)$, to $S(x)$, the empirical distribution function, where

$$S(x) = \frac{\text{Number of observations with } x_i < x}{n}. \qquad \textbf{Equation } 39.$$

The test statistic is the the maximum (denoted by "sup" for supremum) vertical distance between the two functions and is defined as

$$D = |F_o(x) - S(x)|. \qquad \textbf{Equation } 40.$$

The value of D is compared with a critical value and H_o is rejected for large value of D. For further detail we refer to Conover.[16]

RESULTS AND DISCUSSIONS

In this section, the various models discussed above are applied to two BMI data sets. The first BMI data set was retrieved from the University of Wisconsin–Eau Claire (UWEC)'s Student Health Services. The second BMI data set was collected from National Health and Nutrition Survey (NHANES) and is available in the R package mixsmsn. The model selection criteria, AIC and BIC, and the K-S goodness of fit test p-values are calculated to assess the suitability of the fitted distributions. All computations are conducted using the statistical software R.

UWEC BMI Data

The UWEC BMI data was retrieved from the National College Health Assessment conducted by the UWEC Student Health Service. The analyzed data came from a sample of 630 students attending UWEC during the 2014–2015 academic year. The BMIs of the students who have visited the Student Health Service Center were used for the analysis. **Table 1** shows descriptive statistics for this data set where g_1 and g_2 are the skewness and kurtosis, respectively.

n	$\bar{x}$	s	g_1	g_2
630	24.65	5.07	1.99	8.76

Table 1. Summary Statistics of the UWEC BMI data.

The data clearly shows a highly skewed pattern. Therefore, some recently developed skewed distributions like skew-t and skew-normal distributions may provide competitive fits for this unimodal, skewed data.

We fit all the unimodal and bimodal distributions considered above to this data. The maximum likelihood estimates of the parameters are provided in **Table 3**, found in the Appendix. The AIC and BIC values are presented in **Table 5**. Among the unimodal skew distributions, GEV, skew-t and skew-normal distributions are the best models. When flexible skew distributions are considered, the alpha skew-Laplace distribution is shown to best model the data as it has the smallest AIC and BIC values* (3606.54 and 6315.84, respectively). Weibull, log-normal and gamma distributions are not able to describe BMI characteristics well. This is shown by the small p-value of the K-S test, provided in **Table 5**. Using the estimates of the parameter values from **Table 3**, we plotted the expected densities for all models including the observed data in **Figure 6**.

*See Appendix **Table 5** for complete data.

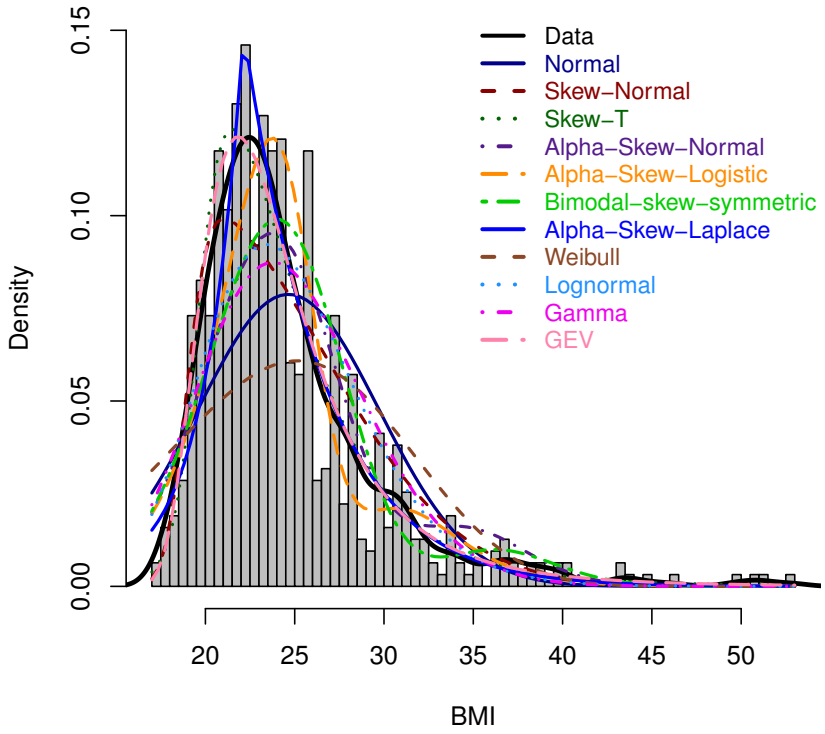

Figure 6. Observed and Expected Density plot of BMI data obtained from 630 UWEC students.

From the observed and expected density plot, it is also confirmed that skew-normal, skew-t, and alpha skew-Laplace models are the best among the skewed and flexible models. Considering the noise at the tail of the observed density, one may argue that flexible distributions, such as bimodal-skew-symmetric distributions and alpha-skew-logistic distributions, also give better fit because they take into account some of that noise. Using the same reasoning, we decided to fit mixture models to this data to see if they provide better fit than the regular models. Using the estimates of the parameter values from **Table 4**, we plotted the expected densities for all mixture models including the observed data in **Figure 7**.

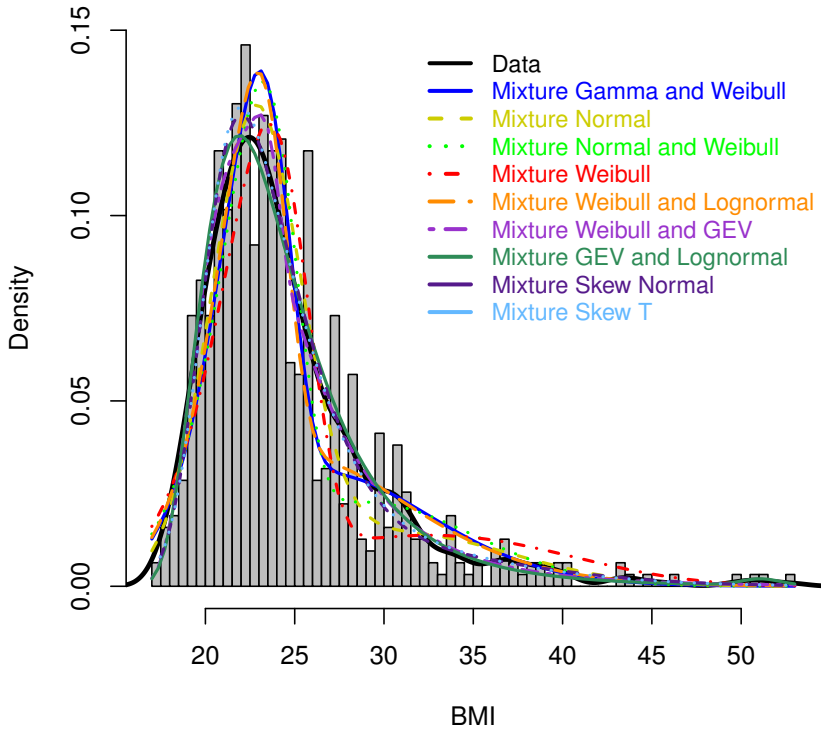

Figure 7. Observed and Expected Density plot of UWEC BMI data using mixture distributions.

Even though univariate distributions and flexible skew distributions show good fits for the UWEC BMI data, the mixture models fit the data better due to their smaller model selection criteria values and larger p-values for the K-S test. According to **Table 6** in the Appendix, mixture of GEV and log-normal provides the closest fits for the observed data followed by mixture of GEV and Weibull, and mixture of scale normal of skew-t .

NHANES BMI Data

We considered the body mass index for men aged between 18 to 80 years. The data set comes from the National Health and Nutrition Examination Survey (NHANES), made by the National Center for Health Statistics (NCHS) of the Center for Disease Control (CDC) in the USA. The data are taken from the NHANES 1999–2000 and NHANES 2001–2002 cycles. Originally, the set had 4579 participants with BMI records. However, to explore the pattern of bi-modality, we consider only those participants who have their weights within [39.50 kg, 70.00 kg] and [95.01 kg, 196.80 kg]. This data set has been used by many authors, for example Lin et al.,[8] and is available in the R package mixsmsn. BMI is defined as the ratio of body weight in kilograms and body height in meters squared. **Table 2** shows descriptive statistics for this data set.

n	$\bar{x}$	s	g_1	g_2
2107	28.19	7.50	0.71	3.30

Table 2. Summary Statistics of the body mass index of 2107 American men.

The mean and standard deviation of the NHANES BMI data are 28.19 and 7.50, respectively. The skewness of the data is 0.71, and the kurtosis is 3.30. The data clearly shows two peaks with some skew pattern. Therefore, some recently developed

flexible skewed distributions such as alpha skew-normal distribution, alpha-skew-logistic distribution, alpha-skew-Laplace distribution may provide more competitive fit for this bimodal skewed data.

We fit all of the skewed and flexible skewed distributions considered above to the NHANES BMI data. Since skew-t and skew-normal distributions most accurately model the UWEC BMI data, we also consider the mixture of these distributions to describe the NHANES BMI data. The maximum likelihood estimates of the parameters are provided in **Table 3**, found in the Appendix. The AIC and BIC values are presented in the **Table 5**. According to AIC and BIC values, among the unimodal skewed distributions, skew-t and skew-normal distributions are the best models while other traditional skewed distributions such as gamma, log-normal, and Weibull has relatively high AIC and BIC values. Using the estimates of the parameter values from **Table 3**, we plotted the expected densities for all models including the observed data in **Figure 8**. Although skew-t and skew-normal are a better fit for the NHANES data than the other univariate distributions, they do not account for the second peak of the distribution as observed from **Figure 8**.

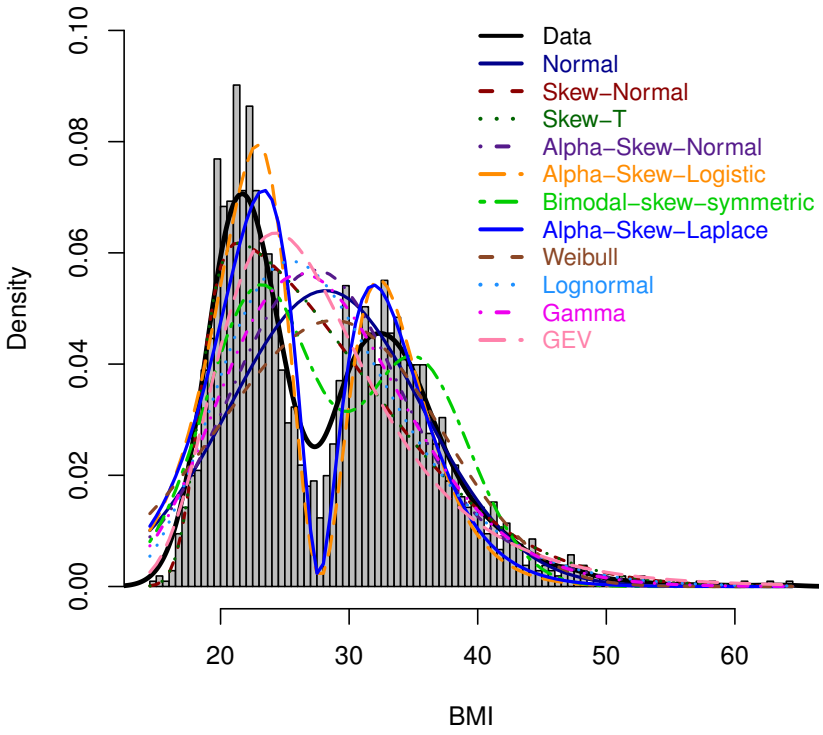

Figure 8. Observed and expected densities of NHANES BMI Data with skewed models.

Among the flexible skewed distributions, we find evidence that alpha-skew-Laplace best describes the NHANES data characteristics—due to its small model selection criteria values (AIC and BIC values)—followed by alpha skew-logistic distribution. These distributions not only have small AIC and BIC values but they also take into account the second peak of the distribution. Since the data are bimodal, we believe that mixture distributions may be suitable to model the data as well. The maximum likelihood estimates for the parameters of the mixture models are presented in **Table 4** in the appendix. The model selection criteria, AIC and BIC values, are given in **Table 6**.

According to the AIC and BIC, mixture of GEV and Log-normal model turns out to be the best fitting model (AIC = 13720.74, BIC = 13724.05). Mixture of skew-normal and mixture of skew-t also describe the data well with small AIC

values of 13764.42 and 13737.61, respectively, and BIC values of 13803.99 and 13777.19, respectively. If we consider large *p*-values from the K-S test with the AIC and BIC, then the mixture of gamma and Weibull also provides a good fit.

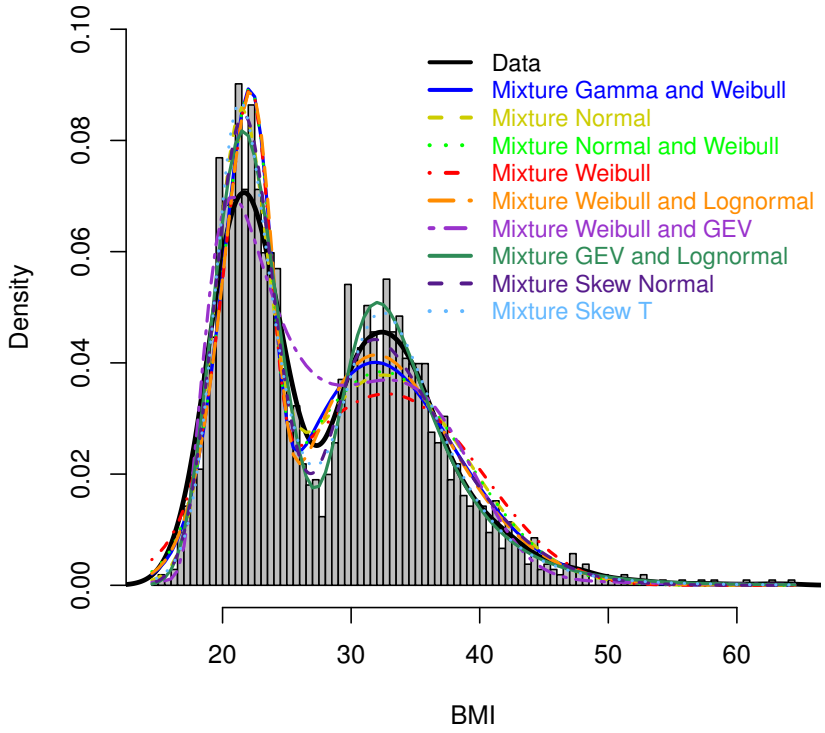

Figure 9. Observed and expected density plot of NHANES BMI Data using mixture distributions

Figure 9 shows the observed and expected density plot of all the mixture models. From the density plot we clearly observed that the scale mixture of skew-normal and skew-t distributions fit the data reasonably well along with the mixture of GEV and log-normal distributions. One of the main advantages of using flexible skewed distributions over mixture distributions is that they are easier to interpret because of the smaller number of parameters. For the mixture models, as the number of components increase, the parameters become very difficult to interpret.

CONCLUSIONS

In this paper, we demonstrated a framework for testing features, such as bimodality, asymmetry, and robustness, of BMI data. We conducted a comparison of skewed, flexible and mixture distribution models using two sets of BMI data. The comparison of the distributions is made using AIC, BIC and the K-S test. We not only looked at conventional unimodal distributions, but also included recently constructed flexible skewed distributions and the models of their mixture. Our findings showed that among the univariate distributions, skew-t and skew-normal proved to provide the best fit for both of the data sets. However, these unimodal skewed distributions fail to account for the second mode of the NHANES BMI data. Among the flexible skewed distributions, alpha skew-Laplace distribution best describes both data sets. Though univariate distributions and flexible skewed distributions may have shown a fairly accurate description of the BMI data, there is evidence that the mixture distributions have considerably good fit as well. For the UWEC BMI data, mixture of GEV and log-normal and the scale mixture of skew-normal and skew-t provided the best fit out of all considered distributions. For the NHANES BMI data, mixture of GEV and log-normal, and the scale mixture of skew-t are the most suitable distributions to represent the data. There are no significant differences between the performance of mixture of GEV and log-normal,

mixture of skew-normal, and mixture of skew-t—though mixture of GEV and log-normal performed slightly better. As a standard health indicator, BMI data still needs extended studies beyond this framework. Given other data sets, we can further analyze the covariance relationship between BMI and other variables such as age and gender using a regression model with error distributions discussed in this paper which covers both skewness and heavy tailed properties of some real data. A more accurate understanding of BMI distribution can assist further research involving topics like obesity and its links to health outcomes.

ACKNOWLEDGMENTS

We want to thank the Office of Research and Sponsored Programs (ORSP) and the Mathematics Department at the University of Wisconsin–Eau Claire for supporting our work, and Student Health Services at UWEC for providing us the data. We also want to thank two anonymous referees for their valuable comments which improved the standard of the paper.

REFERENCES

1. Ma, Y., Genton, M. G., Tsiatis, A. A. (2005). Locally efficient semiparametric estimators for generalized skew-elliptical distributions. *Journal of the American Statistical Association*, **100**(471), 980–989.
2. Canale, A. (2011). Statistical aspects of the scalar extended skew-normal distribution. *International Journal of Statistics*, **69**(3), 279, 295.
3. Olivares-Pacheco, J. F., Salas, E., Gómez, H. W., Bolfarine, H. (2012). An asymmetric extension for the family of elliptical slash distributions. *Communications in Statistics - Theory and Methods*, **41**, 1000–1012.
4. Tan, F., Tang, Y., Peng, H. (2015). The multivariate slash and skew-slash student t distributions. *Journal of Statistical Dirstribution and Applications*, **2**(1), 1–22.
5. Arslan, O. (2015). Variance-mean mixture of the multivariate skew-normal distribution. *Statistical Papers*, **56**(2), 353–378.
6. Manyema, M., Veerman, L. J., Chola, L., Tugendhaft, A., Sartorius, B., Labadarios, D., Hofman, K. J. (2014). The potential impact of a 20% tax on sugar-sweetened beverages on obesity in South African adults: A mathematical model. *PLoS ONE*, **9**(8).
7. Miljkovic, T., Shaik, S., Miljkovic, D. (2016). Redefining standards for body mass index of the US population based on BRFSS data using mixtures. *Journal of Applied Statistics*, 1–15.
8. Lin, T., Lee, J., Hsieh, W. (2007). Robust mixture modeling using the skew-t distribution. *Statistics and Computing*, **17**(2), 81–92.
9. Elal-Olivero, D. (2010). Alpha-skew-normal distribution. *Proyecciones*, **29**, 224–240.
10. Harandi, S. S., Alamatsaz, M. H. (2012). Alpha-Skew-Laplace distribution. *Statistics and Probability Letters*, **83**, 774–782.
11. Hazarika, P. J., Chakraborty, S. (2014). Alpha-Skew-Logistic Distribution. *IOSR Journal of Mathematics*, **10**, 36–46.
12. Hassan, M.Y., El-Bassiouni, M.Y. (2013). Bimodal Skew-Symmetric Normal Distribution. *Communications in Statistics–Theory and Methods*, **45**(5), 1527–1541.
13. Kollu, R., Rayapudi, S. R., Narasimham, S. V. L., Pakkurthi, K.M. (2012). Mixture probability distribution functions to model wind speed distributions. *International Journal of Energy and Environmental Engineering*, **3**(27), 1–10.
14. Aziz, M. (2011). Study of Unified Multivariate skew-normal Distribution with Applications in Finance and Actuarial Science. Unpublished PhD dissertation, Bowling Green State University.
15. Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions*. New York: Wiley.
16. Conover, W. J.,(1999). *Practical Nonparametric Statistics, 3rd edition*, John Wiley & Sons, Inc. New York.
17. Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171–178.
18. Wang, J., Boyer, J., Genton M. G. (2004). A skew-symmetric representation of multivariate distributions. *Statistica Sinica.*, **14**, 1259–1270.
19. Ma, Y., Genton, M. G. (2004). Flexible class of skew-symmetric distributions, *Scandinavian Journal of Statistics*, **31**, 459–468.

20. Basso, R. M., Cabral, C. R. B., Lachos, V. H., Ghosh, P. (2010). Robust mixture modeling based on scale mixtures of skew-normal distributions. *Computational Statistics & Data Analysis*, **54**(12), 2926–2941.

ABOUT THE STUDENT AUTHORS

Thao Tran is an undergraduate student at the University of Wisconsin–Eau Claire with a double major in Comprehensive Finance and Mathematics with a Statistics emphasis. She is also a member of Kappa Mu Epsilon (a mathematics honor society) and the University Honors Program. With her high impact practices, she plans to expand her knowledge and conduct further research in Statistics at the graduate level.

Cara Wiskow is a 3rd year undergraduate student at the University of Wisconsin–Eau Claire and is pursuing a degree in Mathematics with a Statistics emphasis. She is a member of Kappa Mu Epsilon and the University Honors Program. She is also the student representative on the university's Institutional Review Board. After completing her undergraduate coursework, Cara plans on attending graduate school for an MPH in Biostatistics.

PRESS SUMMARY

The normal distribution comes as a first choice while fitting real data, but it may not be suitable to use if the assumed distribution deviates from normality. Flexible skewed distributions have recently drawn considerable attention as alternative models because they are not only capable of including skewness but also flexible enough to take into account multimodality. Using two BMI data sets, we used flexible skewed and mixture distributions to search for the best models. Our results indicate that the skew-t and alpha-skew-Laplace distributions are reasonably competitive when describing unimodal BMI data whereas mixture of normal and finite mixture of scale mixture of skew-normal and skew-t distributions are better alternatives to both unimodal and bimodal conventional distributions. We believe the models discussed in ours study will offer a framework for testing features such as bimodality, asymmetry, and robustness of the BMI data, thus providing a more detailed and accurate understanding of the distribution of BMI.

APPENDIX

Model		NHANES BMI Data	UWEC BMI
Normal	μ	28.188	24.649
	σ	7.499	5.065
Skew-normal	μ	28.151	18.937
	σ	7.546	7.634
	α	0.951	8.048
Skew-t	μ	18.314	19.575
	σ	12.397	5.241
	α	9.454	4.446
	ν	16896.843	4.198
Weibull	σ	31.067	4.296
	α	3.897	26.751
GEV	μ	24.556	0.1631
	σ	0.0453	3.070
	α	5.791	22.294
Log-normal	μ	3.305	3.187
	σ	0.259	0.182
Gamma	α	14.915	28.314
	β	0.529	1.149
Alpha-skew-normal	μ	32.314	27.987
	σ	7.123	4.335
	α	0.754	1.357
BSSN	μ	28.891	28.500
	ψ	0.019	0.024
	β	29.491	32.434
	δ	18.191	5.864
Alpha-skew-logistic	μ	27.310	26.314
	α	4.571	0.857
	β	1.968	1.737
Alpha-skew-Laplace	μ	27.332	22.140
	σ	2.146	2.960
	α	7.307	-0.355

Table 3. Estimated parameter values for skew and flexible models.

Model		NHANES BMI Data	UWEC BMI
MN	μ_1	32.548	30.581
	σ_1	6.418	6.514
	μ_2	21.413	22.809
	σ_2	2.018	2.480
	ω	0.608	0.237
MSN	μ_1	28.457	19.871
	σ_1	62.240	19.215
	α_1	3.667	2.869
	μ_2	25.529	20.221
	σ_2	7.997	94.663
	α_2	0.991	3.661
	ω	0.517	0.848
MST	μ_1	19.877	19.928
	σ_1	9.198	18.240
	α_1	1.444	2.708
	μ_2	29.742	26.078
	σ_2	33.249	83.404
	α_2	2.246	3.296
	ν	6.021	25.501
	ω	0.517	0.859
MW	σ_1	34.427	32.355
	α_1	4.760	4.495
	σ_2	22.203	23.746
	α_2	13.121	10.417
	ω	0.661	0.279
MNW	μ_1	32.377	28.865
	σ_1	6.456	6.158
	σ_2	22.166	23.316
	α	12.487	11.823
	ω	0.622	0.351
MGW	α_1	30.479	23.704
	β	0.925	0.850
	σ	22.169	23.0980
	α_2	12.157	12.640
	ω	0.592	0.429
MWLN	σ_1	22.258	13.138
	α	11.924	22.962
	μ	0.169	0.201
	σ_2	3.493	3.288
	ω	0.430	0.516

Model		NHANES BMI Data	UWEC BMI Data
MWGEV	σ_1	35.131	23.449
	α_1	6.662	30.108
	μ	21.442	0.172
	σ_2	3.253	3.156
	α_2	0.285	22.296
	ω	0.438	0.056
MGEVLN	μ_1	32.400	22.295
	σ_1	3.520	3.038
	α	0.088	0.139
	μ_2	3.087	3.934
	σ_2	0.122	0.021
	ω	0.483	0.995

Table 4. Estimated parameter values for mixture models

Model	NHANES BMI Data				UWEC BMI Data			
	LogL	AIC	BIC	K-S Test	LogL	AIC	BIC	K-S Test
Normal	-7234.19	14472.38	14483.69	2.20E-16	-1915.99	3835.98	3844.87	2.47E-12
Skew-normal	-6995.39	13996.78	14013.74	2.20E-16	-1790.70	3587.41	3594.30	1.27E-6
Skew-t	-6995.40	13982.79	13975.49	1.64E-10	-1763.31	3534.62	3539.51	0.4050
Weibull	-7280.88	14565.76	14577.07	2.20E-16	-1988.89	3981.80	3990.69	2.20E-16
GEV	-7092.59	14191.18	14200.48	2.20E-16	-1759.766	3525.53	3532.423	2.20E-16
Log-normal	-7104.323	14212.65	14223.95	2.20E-16	-1827.00	3657.99	3666.88	2.20E-16
Gamma	-7130.11	14264.22	14275.52	2.20E-16	-1852.21	3708.42	3717.31	2.20E-16
Alpha-skew-normal	-7202.510	14399.02	14427.98	2.20E-16	-1840.63	3687.27	3694.16	2.20E-16
BSSN	-7141.82	14291.64	14298.94	NA †	-1833.82	3675.63	3680.53	NA
Alpha-skew-logistic	-7128.22	14262.44	14271.75	NA	-1825.67	3657.34	3664.23	NA
Alpha skew-Laplace	-7091.47	14188.93	14198.24	2.20E-16	-1800.27	3606.54	3615.84	2.20E-16

Table 5. Model fitting results: skewed and flexible models.

Model	NHANES BMI Data				UWEC BMI Data			
	LogL	AIC	BIC	K-S Test	LogL	AIC	BIC	K-S Test
MN	-6911.67	13833.35	13803.66	2.20E-16	-1779.75	3549.50	3546.61	2.20E-16
MSN	-6875.21	13764.42	13803.99	NA	-1758.29	3534.25	3565.37	NA
MST	-6859.81	13737.61	13777.19	NA	-1758.29	3534.25	3565.37	NA
MW	-6984.54	13979.08	13984.38	3.29E-4	-1814.92	3639.84	3642.7	2.20E-16
MNW	-6926.70	13863.39	13868.70	0.0262	-1801.98	3613.95	3616.84	0.2480
MGW	-6903.83	13817.66	13822.96	0.1390	-1787.58	3585.16	3588.05	0.3490
MWLN	-6897.41	13804.82	13810.12	0.2682	-1781.41	3574.82	3575.71	0.2873
MWGEV	-6938.18	13888.36	13891.67	0.00122	-1758.21	3528.43	3529.32	2.20E-16
MGEVLN	-6854.37	13720.74	13724.05	2.20E-16	-1757.57	3527.13	3528.03	2.20E-16

Table 6. Model fitting results: mixture models.

†Not available

Exploration of the Influence of Smiling on Initial Reactions Across Levels of Facial Attractiveness

Stephanie M. Shields[a], Caitlin E. Morse[a], Paige Arrington[b], and David F. Nichols[a]*

[a]*Department of Psychology, Roanoke College, Salem, VA*
[b]*Graduate Program in Liberal Arts, Hollins University, Roanoke, VA*

Students: smshields@mail.roanoke.edu, cemorse@mail.roanoke.edu, arringtonap@hollins.edu*
Mentor: dnichols@roanoke.edu

ABSTRACT
Both attractiveness and emotionality independently affect perception and interact to influence how a person perceives others. It has previously been shown that expressing positive emotions increases perceived attractiveness in general, but the relative influence of smiling across attractiveness levels and timing of this interaction is unknown. Such an interaction could entail dependent brain processing with interactions between brain areas or independent processing within each brain area. The present studies aimed to investigate this interaction and how it occurs through behavioral, specifically self-report, and physiological, specifically electrophysiological, methods. In each study, female undergraduate participants were shown images of male faces with smiling or neutral expressions. Study 1 used participant ratings to provide insight into the interaction and to establish an image subset of faces of high attractiveness (HA) and low attractiveness (LA). An interaction was found wherein HA faces were rated significantly higher on attractiveness when smiling whereas LA faces were rated similarly attractive regardless of emotional expression. Study 2 used electroencephalography (EEG) to examine the timing of brain responses to attractiveness, emotionality, and their interaction. Though a main effect of attractiveness consistently occurred prior to a main effect of emotional expression across two data sets, the presence of an interaction effect was inconsistent. There was some evidence for independent processing wherein the earliest brain responses are predominantly affected by attractiveness and are influenced by emotional expression, but dependent interactions between modular processing areas cannot be ruled out. Together, these results help to shed light on the interplay of attractiveness and emotionality though additional research could help to clarify the timing of the interaction on a neural level.

ABBREVIATIONS
HA – high attractiveness; MA – medium attractiveness; LA – low attractiveness; FERET – Facial Recognition Technology; ERP – event-related potential; EPN – early posterior negativity

KEYWORDS
Attractiveness, Emotionality, Emotional Expression, Smiling, Electroencephalography, Event Related Potentials

INTRODUCTION

How one appears to others can be important as first impressions can have a significant and lasting influence on how people perceive others.[1] Physical appearance and current emotional expression are readily apparent upon first meeting someone, making these important factors to consider in understanding the initial reactions of others. Attractiveness in particular has been shown to be a major factor that can have an impact on how one is perceived, especially in regards to the age of the respective person.[2] Specifically, and regardless of gender, attractive people are perceived as more socially desirable and are believed to have more successful life outcomes.[3] Raters who either know or do not know the people they are rating judge and treat attractive children and adults more positively than unattractive children and adults.[4] Emotionality, the observable component of emotion, has also been shown to affect the perceptions of others.[5] For example, impressions are impacted by smiling and the display of angry facial expressions.[5] Regardless of the sex of both the observer and the person being observed, smiling people are perceived to be more likable and intelligent[6] and more attractive.[6,7] The apparent influence of emotional expression on judgements of attractiveness implies that there is an interaction at play between the neural processing of facial attractiveness and emotionality. Previous research shows that attractive faces activate the medial orbitofrontal cortex (OFC) and that responses in the medial OFC to attractive faces are enhanced when the face was smiling, suggesting that the medial OFC classifies attractiveness and smiling expressions as rewarding for observers.[8]

To better understand the neural processing of attractiveness and emotionality, electroencephalography (EEG) can be used to measure neural activity through event related potentials (ERPs) evoked by images that are high or low in particular qualities. Prior findings on the timing of differential responses to high and low attractiveness and of positive versus neutral emotional expressions suggest that attractiveness is processed faster than positive emotions. Specifically, differences in ERP signals recorded

after the presentation of attractive vs. nonattractive images appear by at least 250 ms post-stimulus[8] and possibly earlier when participants are required to make judgements of attractiveness.[9] Differences between negative emotional expressions and neutral expressions also appear early, i.e., within 250 ms, but differences between happy expressions versus neutral expressions appear about 100 ms later.[10] However, conclusions on the relative timing of the neural processing are currently based on results from separate studies. To our knowledge, no studies have tested the relative timing of attractiveness and emotional processing using the same set of stimuli nor have any studies reported on the timing of their interaction. The current study is an attempt to fill this gap in the literature.

Behavioral investigations, such as those using self-reported ratings, can provide insight into the potential existence and nature of a neural interaction between attractiveness and emotionality. If an interaction effect does exist, the timing of it needs to be explored with an electrophysiological measure in order to determine how the brain combines information about these two influences on facial perception. Three possibilities for the timing of the interaction exist. If attractiveness and emotionality are processed sequentially but then combined in a dependent manner, an interaction effect would arise only after the occurrence of both main effects. Based on the timing discussed above, this would be about 450 ms or more after the stimulus. If they are processed in parallel but information dependent on both components is required before the interaction can take place, the interaction effect would be expected to occur no earlier than a main effect of emotion, thus about 350 ms or more after the stimulus. If facial expression has a direct, independent effect on attractiveness and if processing of emotional expression is not a prerequisite for the interaction, the interaction effect could happen before the main effect of emotionality and could be observed at the same time as or soon after the main effect of attractiveness. This would be approximately in the time window of 200 ms to 300 ms after the stimulus. This last possibility is most consistent with modular processing of facial attributes which would suggest that the processing of attractiveness and emotionality involve different parts of the brain that interact with one another for some judgments but not others[11]. Prior research has indicated that attractiveness is processed in the medial OFC[8] and emotional expression is processed in the amygdala as well as the OFC.[12]

The current studies aimed to investigate the impact of facial attractiveness and emotionality, specifically smiling expressions, on the perception of others through behavioral and physiological measures. The behavioral measures are largely exploratory regarding any influence of attractiveness level on the effect that emotional expression has on ratings of facial attractiveness and perceived happiness. The physiological measures are in part confirmatory in that we hope to replicate prior research on the relative timing of neurological responses to facial attractiveness and emotional expression and in part exploratory because of the different possibilities discussed above regarding the timing of a neurological interaction effect between facial attractiveness and emotionality. Study 1 had female participants rate the attractiveness and happiness of a wide range of everyday male faces that contained images of each face with a neutral expression and with a smile. Study 2 extended these findings by investigating brain responses to the subset of images that were rated highest and lowest on attractiveness in order to localize in time the differences found among conditions in Study 1.

STUDY ONE

The ultimate goal of Study 1 was to evaluate the existence and nature of an interaction between attractiveness and emotionality in the perception of facial attractiveness. A preliminary step, and additional aim, was to establish a four-category set of face images that clearly varied in attractiveness, i.e., a high attractiveness and a low attractiveness set of images, but that were also different in emotional appearance, i.e., were perceived as happier when smiling versus with a neutral expression. Without clear differences in those dimensions, definitive comparisons could not be made across image categories. The images needed to be rated across enough participants to establish confidently that they would reliably be considered relatively attractive or unattractive and relatively more or less happy. This would allow them to be used in Study 2 with a separate set of participants who were not performing explicit ratings of the stimuli. It was expected that happiness ratings would be influenced by emotional expression across all attractiveness categories. Having a set of stimuli for which the attractiveness ratings change as a function of emotional expression more so in one attractiveness category than another would allow us to unambiguously determine whether an interaction effect in the neural response between attractiveness and happiness is primarily driven by attractiveness or emotionality processing. Stimuli used in studies on attractiveness have come from a wide array of sources, such as photographs taken by the researchers,[13,14] images taken from the internet,[15] or pictures from previously established databases of face images.[7,9,16,17,18] While all image sets have their utilities in answering various questions, they are generally limited in how representative they are of the general population. We sought to utilize face images that had a broader characterization of the general population, including a range of ages and ethnic groups. Such an increased heterogeneity of faces may reveal effects that have previously been masked when researchers use only a narrow range of images. For instance, using only college-aged male faces would not allow us to determine if college-aged females respond differently to smiles of various-aged males, and we would not be able to draw broader conclusions about patterns in the perception of attractiveness and emotionality. Therefore, we decided not to limit the range of face images based on age or ethnicity and instead used images of individuals of a variety of ages and ethnicities.

METHODS

Participants

A total of 29 female undergraduate students enrolled in psychology courses participated in this study. No further information about demographics was recorded. Students were primarily in introductory level psychology courses and were recruited from a participant pool which primarily consisted of young college-aged women of European descent. The study was conducted in accordance with the Roanoke College Institutional Review Board.

Stimuli

Two hundred black and white photographs of men were obtained from the FERET facial stimuli database from the Defense Advanced Research Products Agency.[19] Photographs depict close-ups of faces displaying happy and neutral facial expressions with multiple images per person, some including different backgrounds and partial obstruction of the face. Images taken of individuals from an angle other than their front were excluded. Images were also excluded if the background of the image was anything other than plain white, if the individual was wearing accessories that obscured the face (e.g., hats or sunglasses), or if the individual's eyes were closed. To avoid biases in image selection, the first 60 photos of randomly ordered male faces exhibiting a smiling facial expression that met these criteria and images of the same individuals exhibiting a neutral facial expression were selected for use in the study. Though demographic information was not tied to the images, from visual inspection we believe that the stimuli consisted of individuals of European, African, Middle Eastern, and Asian descent and that the ages ranged from that of the average college student to that of senior citizens (see Figure 1 for representative stimuli).

Figure 1. Example faces from the FERET face database used in Study 1. A set of two images were chosen per person, one with a smiling expression (top row) and one with a neutral expression (bottom row). Images varied in age, race, and ethnicity but were all of the opposite sex to the participant.

The face images were 384 pixels high and 256 pixels wide, embedded in a gray rectangle that was 420 pixels high and 420 pixels wide. The gray rectangle had an RGB value that was the midpoint of the face image RGB value (127) and was centered on a black screen (RGB = 0) which was 900 pixels high by 1600 pixels wide. The rating scale was gray (RGB = 150), 8 pixels high by 730 pixels wide, and located 50 pixels below the stimulus image. The nine tick marks were 6 pixels wide by 16 pixels high and equally spaced along the rating scale, approximately 90 pixels apart. Text for the rating scale was gray (RGB = 150) and centered below the corresponding tick marks.

Equipment

The computer used was a Mac Mini running Mac OS X version 10.6.8 connected to a 20" ViewSonic VX2033wm LCD monitor with a 1600x900 resolution at 60 Hz. Stimuli were presented using MATLAB 2009b with Psychtoolbox-3 code. A chinrest was used to maintain a constant viewing distance of 30 cm.

Procedures

Participants were tested individually in an isolated room. They were given a study information sheet to explain the fact that they would be completing two blocks and an instruction sheet with example images to explain the rating procedure before they began. In one block, participants were asked to rate the attractiveness level of the images on a nine-point scale from Very Unattractive (1) to Very Attractive (9) with intermediate points Somewhat Unattractive (3), Neutral (5), and Somewhat Attractive (7) also labeled. In a separate block, participants were asked to rate the level of happiness displayed in each photograph on a nine-point scale of Very Unhappy (1) to Very Happy (9) with intermediate points Somewhat Unhappy (3), Neutral (5), and Somewhat Happy (7) also labeled. Participants entered their rating for each image by moving the mouse cursor to the tick mark on the scale for their selection.

The 120 images (60 smiling, 60 neutral expression) were displayed in a random order one image at a time. Images displaying smiling or neutral facial expressions were intermixed within a block, and therefore participants rated the same individual twice in each block. The order of the attractiveness and happiness rating blocks was randomized across participants. The face image and scale remained visible until a selection was made. Time expectancy for rating the images was approximately six minutes for each scale though participants were allowed to move at their own pace. The entire experiment, including initial training, took roughly 15 minutes per participant.

Analysis

Three different types of statistical tests were used to analyze separate aspects of the ratings data. The Wilcoxon paired sample signed-rank test was used for analysis of the distribution of individual ratings combined across all participants. A nonparametric test was chosen because the rating values are discrete rather than continuous. This test shows if attractiveness or happiness ratings are consistently higher for one emotional expression (i.e., smiling versus neutral). Correlations between the average ratings of each image across participants were also calculated. This test examines consistencies in attractiveness and happiness ratings across emotional expressions of the same individuals. Repeated Measures (RM) ANOVAs across attractiveness categories were conducted based on the average ratings across all images within each category. This test reveals if differences across categories are consistent across participants. A level of significance of $\alpha = 0.05$ was used throughout.

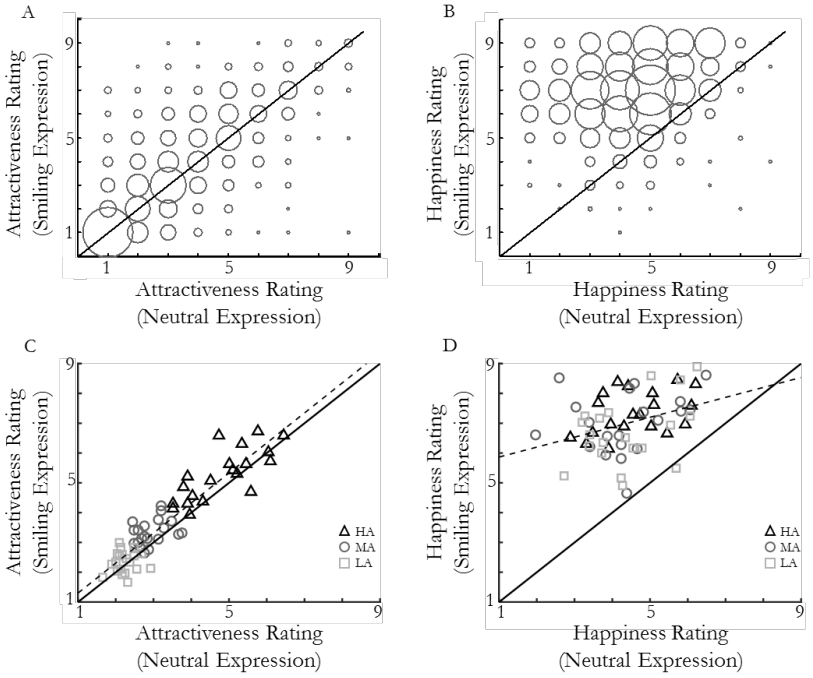

Figure 2. Scatterplots of the ratings across emotional expression for all ratings performed (A&B) and on the basis of each of the 60 photographed individuals split into High Attractiveness (HA), Medium Attractiveness (MA), and Low Attractiveness (LA) categories (C&D). Lines of equality are included as solid lines, and best fit lines are included as dashed lines. A) A bubble plot of the attractiveness ratings shows the relative frequency of particular combinations of ratings across emotion. Data are shown for all 1740 rating pairs obtained from the 29 participants rating 60 individuals each. B) A bubble plot of the 1740 happiness ratings across the same individuals as in A. C) A scatterplot of the attractiveness ratings of each of the 60 individuals averaged across all participants. D) A scatterplot of the happiness ratings of the 60 individuals averaged across all participants.

RESULTS

Ratings

Participants used a variety of attractiveness and happiness values when rating images (see Figure 2A&B). In order to exclude clearly invalid responses, we made sure that no participant used only a single scale value in either rating block and that each participant consistently rated the smiling images as happier than the neutral images. No participants' data needed to be discarded. For the high end of the attractiveness scale, 52% of participants rated at least one image a 9, 94% rated at least one image a 7 or above, and 100% used a rating of 5 or above at least once. This indicates that the image set included at least some images considered attractive by a vast majority of participants. Regarding the range of the attractiveness scale used by each participant, 45% used the entire width of the scale, 88% had a difference of at least 6 places from their highest to lowest ratings, and 100%

had a difference of at least 2 places. For the high end of the happiness scale, 97% of participants rated at least one image a 9 and 100% used a rating of 8 or above at least once. Regarding the range of the happiness scale used by each participant, 66% used the entire width of the scale, 97% had a difference of at least 6 places for their highest to lowest ratings, and 100% had a difference of at least 5 places.

In order to determine attractiveness categories across all participants, the overall attractiveness rating for each individual depicted in the images was determined by averaging across emotional expression and the ratings of all participants (Figure 2C). The average attractiveness rating per individual ranged from 1.8 to 6.5. The 60 individuals were then rank-ordered based on overall attractiveness and split into three categories. The 20 individuals with the highest rating (from 3.7 to 6.5) were grouped into the high attractiveness (HA) category and the 20 individuals with the lowest rating (1.8 to 2.7) were grouped into the low attractiveness (LA) category. The remaining 20 individuals were grouped into the medium attractiveness (MA) category. The happiness ratings across images (see Figure 2D) were not used in the procedure to determine attractiveness categories.

Demographic information on the images was not included with the image set nor did the participants indicate what age and ethnicity they perceived the individuals in the images to be. However, possible differences in frequencies of different ages and ethnicities between the groups could be relevant for interpreting any observed differences between the different attractiveness categories. For each individual, the median age and most frequent ethnicity were determined across ratings made by the student authors. For the HA category, the mean age was 25, and age ratings ranged between 21 and 35. Only one individual sorted in the HA category was perceived as not being of European descent. For the MA category, the mean age was 40.5, and age ratings ranged between 21 and 62. Four individuals in the MA category were perceived as not being of European descent. Finally, for the LA category, the mean age was 49.2, and age ratings ranged between 27 and 63. As in the MA category, four individuals in the LA category were perceived as not being of European descent, so across the image set, nine individuals were rated as being of non-European descent.

Consistency in Ratings Across Images

Participants demonstrated a large degree of individual differences in how attractive they perceived particular photographed individuals to be (see Figure 2A). However, a Wilcoxon signed-rank test across all 1740 ratings revealed a bias with individuals being rated as more attractive when smiling ($z = 8.9$, $p < .001$). There was much greater consistency for the happiness ratings (see Figure 2B) with a clear majority of the images being rated as happier when exhibiting a smiling expression compared to a neutral expression ($z = 33.6$, $p < .001$).

In order to examine the relative effects of emotional expression on the attractiveness and happiness ratings amongst photographed individuals, averages were taken across all participants for each individual separately. On the one hand, emotional expression tended to have a consistent effect on attractiveness rating across attractiveness levels (see Figure 2C), evidenced by the best fit line through the data having a slope of 1.0. The effect was relatively small with a vertical shift of 0.29 on a 9-point scale. On the other hand, emotional expression tended to have a large effect on happiness rating (see Figure 2D), evidenced by the best fit line through the data having a vertical shift of 5.53 on a 9-point scale and a slope of 0.33. The average attractiveness rating, combined across emotional expressions, was not significantly correlated with the average happiness rating, $r(56) = 0.22$, $p = .088$, indicating that ratings were likely based on distinct aspects of the images.

Effects of Attractiveness Category

As noted above, there was a large variation regarding which pictured individuals participants found attractive. Even within the set of images determined to be in the upper third of attractiveness (i.e., the HA category) and the lower third of attractiveness (i.e., the LA category) across all participants, across emotional expressions there was substantial variability between participants in the attractiveness ratings assigned to these individuals (see Figure 3A). For the HA category, there was a median rating of 5, and for the LA category, there was a median of 2. There was also large variability in the happiness ratings for both attractiveness categories (see Figure 3B) though there was much more overlap in the distribution of ratings collapsing across emotional expression as the median for both HA and LA categories was 6.

Though individual differences in perceived attractiveness are inevitable, our goal was to have sufficiently consistent attractiveness ratings across the entire set of images within a category so that a different set of participants would experience the same differences between categories. The reliability of the effects of attractiveness category and emotional expression was therefore assessed across participants (see Figure 3C&D). In order to determine the effects of attractiveness group and emotional expression separately and in combination on both the attractiveness and happiness rating scales, a 2x2 RM MANOVA was run. Regarding attractiveness ratings (Figure 3C), there was a main effect of attractiveness, $F(1,28) = 119.6$, $p < .001$, $\eta^2 = .780$, as would be expected based on the data categorization procedure. There was also a main effect of emotion, $F(1,28) = 12.2$, $p = .002$,

η^2 = .008, though this was qualified by an interaction effect, $F(1,28) = 12.4$, $p = .002$, $\eta^2 = .003$. Further exploration using paired-sample *t*-tests demonstrated that smiling HA faces were consistently rated higher on attractiveness than neutral HA faces, $t(28) = 4.1$, $p < .001$, but smiling LA faces were not significantly different than neutral LA faces, $t(28) = 1.3$, $p = .196$. Regarding happiness ratings (Figure 3D), there was a main effect of emotion, $F(1,28) = 1032.5$, $p < .001$, $\eta^2 = .924$), as would be expected based on the nature of the stimuli. There was additionally a main effect of attractiveness, $F(1,28) = 114.8$, $p < .001$, $\eta^2 = .027$, which was also qualified by an interaction effect, $F(1,28) = 19.6$, $p < .001$, $\eta^2 = .007$. Though smiling individuals were rated higher in happiness than individuals with a neutral expression for both the HA and LA categories (*t*'s>26.6, *p*'s < .001), the increase in happiness ratings was larger for the HA category than the LA category, $t(28) = 4.40$, $p < .001$.

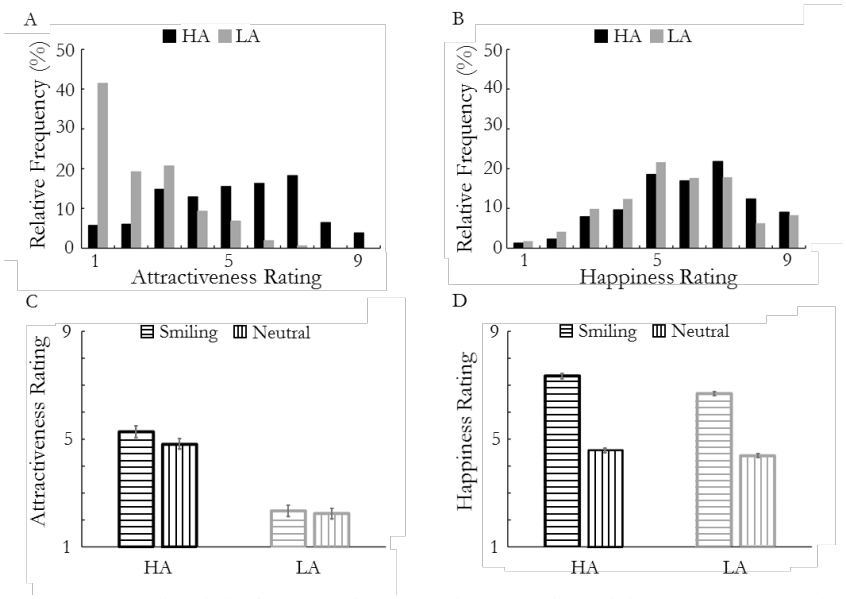

Figure 3. Effects of attractiveness category on the relative frequency of ratings used (A&B) and cumulative averages across participants (C&D). A) The relative frequency of attractiveness ratings across all images within the HA and LA categories. B) The relative frequency of happiness ratings across all images within the HA and LA categories. C) Attractiveness ratings averaged across all 20 images within a category as a function of the attractiveness level and emotional expression. D) Happiness ratings averaged across the same images as in C. Error bars show 1 standard error of the mean.

DISCUSSION

This study established image sets comprised of individuals of differing levels of attractiveness – HA and LA – displaying two different expressions – neutral and smiling. It was important that the image sets varied in attractiveness and perceived emotionality for the analysis of Study 1 data and so that the images could be used in Study 2 to examine brain responses during passive viewing.

Results of Study 1 replicated findings of prior research demonstrating a main effect of smiling on attractiveness, in that smiling faces were rated as more attractive than faces displaying a neutral expression.[6,7] These previous studies did not, however, address the influence of smiling as a function of attractiveness level, whereas that interaction was tested for and examined in Study 1.

A clear interaction pattern was established wherein the HA individuals were rated as higher in both attractiveness and happiness when smiling but the LA individuals were rated as higher only in happiness when smiling. Note that the lack of a difference in attractiveness ratings for the LA individuals as a function of emotional expression is in large part due to the high frequency of low attractiveness ratings (see Figure 2A) across participants. Thus, though individuals were included in the HA category if they were rated high in attractiveness for either emotional expression, the LA category only included individuals rated low in both emotional expressions. This restricted range effect is advantageous for Study 2, however, because as a result attractiveness varied based on emotionality in only the HA category. This difference in the response pattern, i.e., an effect of smiling for the LA category on happiness ratings but not on attractiveness ratings, is useful in drawing conclusions about the relative timing of the combination of attractiveness and emotionality in the brain. Namely, it helps answer the question of whether conveyed facial emotion needs to be processed before the physical emotional expression influences the brain response to attractiveness.

STUDY TWO

Investigating the timing of the neural processing of attractiveness, emotionality, and their interaction can offer insights into how these factors are processed in the brain. In order to do so, it is crucial to know the relative timing of each of these three processes. Initial brain processing that differentiates faces from other visual stimuli has been shown to occur within 200 ms of stimulus presentation[18] in parts of the left and right fusiform and inferior temporal gyri. With the use of scalp recordings, additional processing of the nature of the face, such as its relative attractiveness and emotional expression, has been found to occur after this and in different locations as detailed below.

Studies on the timing of the perception of attractiveness indicate that it is processed especially early. Werheid *et al.* discovered that an early posterior negativity (EPN; ~250 ms) occurs in addition to a late parietal positivity (400-600 ms) for attractive faces in relation to unattractive faces.[9] Task specific effects, such as those resulting from explicitly judging relative attractiveness, can occur even earlier, with differential effects found for the P1 (100-140 ms) and N170 (140-190 ms) ERP components with larger magnitudes for attractive faces.[9,20] Zhang and Deng also found an early N300 (230-330 ms) with greater negativity for attractive faces,[20] which we expect to be the early stage attractiveness response for passive viewing. We thus expected to observe an effect of attractiveness in the 230-300 ms time range during passive viewing.

Based on the literature, the timing of emotionality processing at first appears to be a bit more variable than that of attractiveness, but this is likely in part due to the various forms of emotion examined and the tasks required of participants. According to a recent review by Calvo and Nummenmaa, emotional expression is processed differently than neutral expression between 150 and 350 ms, the range of N170 to EPN.[21] The earliest differential responses are possibly based on negative emotions[22] and/or task demands of making an explicit choice.[17] Balconi and Pozzoli observed negative peaks around 230 ms elicited by faces displaying fear, anger, surprise, happiness, or sadness.[22] Schupp *et al.* differentiated responses to threatening, friendly, and neutral facial expressions and found a greater and earlier overall effect for threatening stimuli.[16] Furthermore, Calvo *et al.* found that for angry and fearful faces in relation to neutral ones, a differential response occurred between 175-250 ms and was present in the frontal and central scalp regions in addition to some posterior regions.[10] They found that differential responses for happy faces in relation to neutral ones occurred later – around 330-430 ms – and were predominantly located in the posterior regions.[10] We thus expect to observe an effect of happy versus neutral expressions starting in the 300-400 ms time range.

Our review of the literature suggests that passive viewing of facial stimuli would result in an earlier brain response to attractiveness than emotional expression. To examine post-stimulus effects across categories, we used a component-independent experimental design which looks into the timing of effects specifically rather than parsing out the ERP components involved in advance. In *An Introduction to the Event-related Potential Technique*, Luck states this is the best strategy to use in order to avoid the obscurity typically enmeshed with determining which components compose an ERP waveform.[23] Three possibilities exist for the timing of an interaction effect: (1) before the processing of emotion, (2) concurrent with the processing of emotion, or (3) after the processing of emotion. Respectively, these effects are estimated to occur approximately (1) less than 300 ms after the stimulus, (2) in the 300-400 ms time range, and (3) more than 400 ms after the stimulus. Early interaction effects occurring prior to explicit emotional processing (possibility 1) would suggest independent processing of attractiveness and emotional expression with emotional expression having a direct modulation on attractiveness processing. Late interaction effects occurring during or after both attractiveness and emotionality processing (possibilities 2 and 3) would suggest parallel or sequential dependent processing. As we are the first to our knowledge to look at these responses for the same set of stimuli in the same study, a specific hypothesis about the time at which an interaction would take place was not made.

METHODS

Participants

Data Set 1: For this study, 27 female students were recruited from the general student body. None were in Study 1. Students in psychology classes received course credit in exchange for their participation. No information about demographics or handedness was recorded. All participants had normal or corrected-to-normal vision. Anyone susceptible to seizures induced by flashing stimuli was not eligible to participate. The study was conducted in accordance with the guidelines of the Roanoke College Institutional Review Board, and all participants provided informed consent.

Data Set 2: For an unrelated study,[24] 112 students were recruited from the general student body. The data of 45 female students with a sufficient number of valid ERP trials (see below for more details) was also included in the current study. All other information given about participants in Data Set 1 is the same.

Equipment

The equipment setup was the same as for other projects in the lab.[25] A PowerLab 26T from AD Instruments was used to record the electroencephalography (EEG) signals. Five lead shielded electrodes transmitted signals from the participants' scalps to the PowerLab unit. The EEG signal was sampled at 400 Hz and underwent bandpass filtering between 0.5 Hz and 50 Hz with a Mains filter applied prior to recording. In order to reduce the possibility of artifacts, all trials were conducted in a dark room with minimal possible distractions. All software was run on a Dell XPS 15z laptop. Participants viewed stimuli on the laptop's internal 15" widescreen monitor. Stimuli were presented to participants using SuperLab 4.5 from Cedrus Corporation. An external 17" Dell monitor viewable only to experimenters presented the output of the EEG signals through LabChart 7 software from AD Instruments. The temporal presentation order of the stimuli was recorded using a StimTracker device from Cedrus Corporation.

Stimuli

Participants were presented with images of faces and of fingerprints. Fingerprints were included as a way to give participants the opportunity to blink during runs. The 80 face images were from the HA and LA categories of Study 1, i.e., they were of the 20 most attractive and the 20 least attractive individuals, each with two facial expressions. These images were 384 pixels high and 256 pixels wide, embedded in a gray rectangle (RGB = 127). The 20 pictures of fingerprints were 383 pixels high and 254 pixels wide with no gray border. The fingerprints appeared as black lines (RGB = 0) on a white background (RGB = 255). All images were presented using SuperLab 4.5 on a gray background (RGB = 130). At the beginning of each run, a white cross appeared in the center of the screen in order to guide fixation. The faces were centered on the same location as the middle of the cross, with the center of the screen being around the nose area of the face.

Procedures

Upon entering the experiment room, potential participants were prescreened and given an informed consent sheet. If they were susceptible to seizures, had metal in their heads, or did not have normal or corrected-to-normal vision, they were not eligible to participate (n = 0). Participants were additionally warned that the procedures of the study could possibly induce migraines and were asked to remove any jewelry on or around their face. After cleaning and abrading their forehead, the electrodes were put in place (see below for specific locations for each data set). Channel 1 focused on the anterior-posterior axis and included an electrode over the occipital lobe (Oz) in order to detect general visual processing. Channel 2 included an electrode over the right temporal-parietal region (TP10) in order to detect initial face processing and capture the right-lateralized EPN. This effect has been localized to the right posterior region in prior studies.[9,10,20] Though a small number of electrodes limits spatial resolution, the temporal resolution was sufficient to address our question of the relative timing of brain responses.

Participants were then asked to rest their chins on a pillow placed on top of a stack of books, and the lights were turned off. Stimuli were presented on a laptop screen approximately 30 cm from the participant. The LabChart 7 software was opened, and the experimenter evaluated the incoming signals to help ensure that the equipment was set up and functioning properly. Once the experimenter had tested the electrode connections for voltage consistently within the +/- 60 μV range, five runs were conducted which lasted about three minutes each. Sometimes additional runs were conducted if equipment malfunctioned or participants experienced some sort of issue. For example, the SuperLab 4.5 software crashed on approximately 5% of the runs. In between runs, participants were always given the opportunity to rest their eyes if necessary. For each run, LabChart recorded the continuous EEG output from the PowerLab unit and the discrete timing of the stimulus presentations from the StimTracker unit. The stimuli were presented to participants in a random order using SuperLab 4.5.

Each run consisted of the presentation of 100 trials, i.e., a single presentation of each of 80 face images and 20 fingerprint images. The task was simply to look at the face images while refraining from blinking and to blink when the fingerprint images appeared. At the beginning of each run, a white cross appeared in the center of the screen for 500 ms. There was a 600 ms pause between when the cross was removed and when the first stimulus was presented. A comment line was inserted into the EEG data trace 100 ms prior to the presentation of each stimulus, with different comments specifying the different stimulus categories. Each stimulus image was presented for 300 ms. Since judgements made based on viewing photographs for only 100 ms and without time constraints are highly correlated,[26] a 300 ms presentation was sufficient for people to form an impression based on appearance. The interstimulus interval between the end of the presentation of one stimulus and the start of the next stimulus was 1300, 1400, or 1500 ms. Each run took approximately three minutes. A large number of trials is needed to average out noise in the EEG signal, so the total number of runs was chosen to maximize the number of trials per stimulus that could be collected within a 30 min experimental time slot. After completion of the runs, participants were debriefed. The entire process took approximately 30 minutes per participant.

Data Set 1: The ground electrode was placed on the left side of the forehead approximately halfway between the eyebrows and hairline (FP1). The negative Channel 1 electrode was placed on the right side of the forehead halfway between the eyebrows and

hairline (FP2), and the positive Channel 1 electrode was placed approximately an inch above the inion (OZ). The negative Channel 2 electrode was placed on the left earlobe (A1), and the positive Channel 2 electrode was placed about an inch behind and above the right ear (TP10). The electrode placed on A1 was disposable and was affixed to the earlobe with its adhesive backing. The other four electrodes were attached to the scalp using electrode paste and held in place with an elastic headband. Participants viewed the stimuli five times each. As there were 20 images per stimulus category, there were 100 trials per stimulus category across the five runs.

Data Set 2: The procedures were nearly identical to those of Data Set 1, though participants completed four runs rather than five. Additionally, for two out of the four runs, fingerprint images were not displayed, and participants were asked not to blink for the duration of the run (approximately 2.5 minutes). Also, there were slight differences in the placement of the electrodes. For Channel 1, the electrodes were all in the same anatomical regions but the sides of the forehead were reversed for the positive electrode and the ground electrode. Therefore, the data for Channel 1 is believed to be comparable to the other data set. For Channel 2, the negative electrode was moved from the earlobe to the left temple (F7) and the positive electrode remained in the same location (TP10). All five electrodes were attached to the scalp using electrode paste and held in place with elastic headbands. As there were 20 images per stimulus category, there were 80 trials per stimulus category across the four runs.

Analysis

Following data collection, the data were exported from LabChart 7 into a format that could be analyzed in a custom-designed MATLAB program for ERP analysis. The data were exported as a matrix containing the entire time course of the EEG recordings with a separate matrix indicating the corresponding temporal location and nature of the comments that specified which stimulus was presented when. As the shortest duration between successive stimuli was 1300 ms, the ERP for each trial was determined for a time span of 1200 ms, starting 300 ms prior to the stimulus onset and ending 900 ms after the stimulus onset. The voltage values were scaled such that the baseline average, defined as the voltage amplitude from 300 ms to 100 ms prior to stimulus onset, was always 0.

Artifacts were determined by whether or not the voltage within a trial ever reached an amplitude greater than 60 μV above or below the baseline voltage, and a trial was discarded from further analysis if an artifact was present. Eye stability was not directly checked, though large eye movements or blinks would lead to artifacts in Channel 1 due to the location of an electrode on the frontalis muscle. Since artifacts in Channel 1 were associated with the loss of visual input for a given trial, corresponding trials in Channel 2 were also discarded when an artifact was present in Channel 1. Additional trials were discarded for Channel 2 alone if an artifact was present in only Channel 2. For Data Set 1, artifacts were present in 41% of trials across all participants, though the frequency of artifacts varied by participant from 1% to 91%. Participants who averaged less than 40 valid trials per condition in Channel 1 were eliminated from further analysis. The data of four participants was discarded, leaving 23 remaining participants. For the remaining participants, an average of 35% of trials had artifacts, i.e., 65% were considered valid trials. For Data Set 2, participants were only included from the beginning if they averaged at least 40 valid trials per condition in Channel 1 and individuals varied from 58% to 97%. An average of 26% of trials had artifacts, i.e., 74% were considered valid trials in Channel 1.

ERPs were determined separately for the four conditions (HA smiling, HA neutral, LA smiling, LA neutral) by averaging across all valid trials. The magnitude of the response for a particular category was the average voltage across time within a particular time window that could be used to examine main effects of attractiveness and emotionality as well as their interaction. Consistency of the effects across participants for each time window was determined by a 2x2 RM ANOVA with attractiveness category and emotional expression as the factors. RM ANOVAs were performed within 100 ms time windows using a sliding window procedure that shifted the analysis window by 10 ms for each subsequent analysis. Significance within a time window is reported based on the midpoint of the time window, i.e., a midpoint of 190 ms refers to a window of 140-240 ms. The p-values for the main effects and interaction effect were recorded within each time window so that subsequent analysis could be performed on when in time each effect was present. A nominal significance level of $p < .05$ was used. Caution needs to be used, however, in interpreting significance if the range of time windows that are significant is small, i.e., less than three consecutive windows, due to the high number of statistical tests being run on each time course and the substantial overlap of adjacent time windows.

RESULTS

Valid Trials

Not all participants ended up with the same number of valid trials because individual trials with artifacts were discarded from the data. It was possible that the number of valid trials varied with the conditions and, furthermore, behavioral artifacts could have been differentially caused by the conditions. In order to test this, a 3-way RM ANOVA was run for each data set with the attractiveness category (HA vs. LA), emotional expression (smile vs. neutral expression), and channel (Channel 1, including the frontalis muscle, vs. Channel 2, including the temporalis muscle) as factors.

Data Set 1: There was a main effect of channel, $F(1,22) = 5.2$, $p = .032$, $\eta^2 = .047$, with more valid trials in Channel 1 (66.7%) than Channel 2 (63.5%) and also of emotion, $F(1,22) = 10.6$, $p = .004$, $\eta^2 = .122$, with more valid trials for neutral (67.7%) than smiling (62.5%) expressions (Figure 4A). The interaction effect between channel and emotionality was also significant, $F(1,22) = 5.2$, $p = .032$, $\eta^2 = .047$, as the frequency of valid trials for the smiling expression dropped from 65.7% to 59.2% when artifacts specific to Channel 2 were included but no artifacts were found specific to Channel 2 for the neutral expression. Neither the main effect of attractiveness category (F<<1) nor any of the interaction effects associated with attractiveness category (F's<<1) were significant. Therefore, subsequent analyses looking at the timing of interaction effects in the ERPs within Channel 1 involving attractiveness can be considered unbiased with regards to the number of valid trials.

Data Set 2: Four participants were excluded from this analysis because they did not have any valid trials in the smiling conditions for Channel 2, though all results were qualitatively similar when they were included. There was a main effect of channel, $F(1,40) = 15.1$, $p < .001$, $\eta^2 = .077$, with more valid trials in Channel 1 (74.4%) than Channel 2 (69.2%) and also of emotion, $F(1,40) = 6.4$, $p = .016$, $\eta^2 = .044$, with more valid trials for neutral (73.8%) than smiling (69.8%) expressions (Figure 4B). The interaction effect between channel and emotionality was also significant, $F(1,40) = 15.1$, $p < .001$, $\eta^2 = .077$, as the frequency of valid trials for the smiling expression dropped from 75.0% to 64.6% when artifacts specific to Channel 2 were included but no artifacts were found specific to Channel 2 for the neutral expression. Neither the main effect of attractiveness category (F<<1) nor any of the interaction effects associated with attractiveness category (F's<<1) were significant.

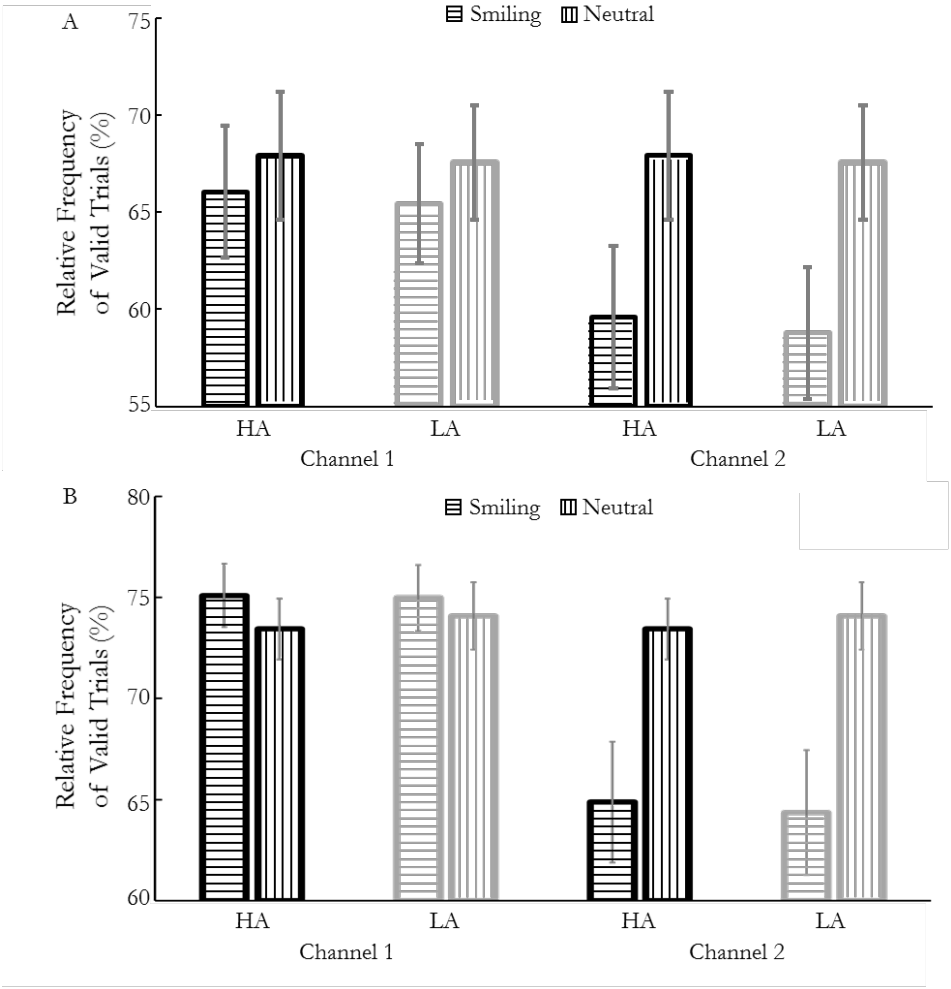

Figure 4. The percentage of trials for each of the Study 2 conditions averaged across participants, graphed separately by channel. Note that all trials excluded in Channel 1 were also excluded in Channel 2 but trials could be excluded in Channel 2 without being excluded in Channel 1. Error bars show 1 standard error of the mean. A) Results from Data Set 1. B) Results from Data Set 2.

Main Effects and Interaction Effects in ERP Data

The time course for the ERPs as a function of attractiveness category and emotional expression within Channel 1 is shown in Figure 5A for Data Set 1 and Figure 5B for Data Set 2. Of particular interest is the differential timing of when brain responses appear for the main effects of attractiveness and emotionality as well as the interaction effect between them. The nature of the observed interaction effect, i.e., whether the brain responses more closely resemble the observed interaction patterns in Study 1 based on attractiveness or on happiness ratings, is also of interest.

Data Set 1: A significant main effect of attractiveness with a more positive voltage for HA faces was found for time windows of midpoints 190 ms to 310 ms. A later main effect of attractiveness with a more negative voltage for HA faces was found from midpoints 730 ms to 780 ms. No post-stimulus time windows exhibited a significant main effect at the level of $p < .05$ for emotional expression, though a marginally significant main effect of emotional expression at the level of $p < .075$ was found for time windows of midpoints 300 ms to 340 ms. A significant interaction effect was seen for just a time window midpoint of 210 ms. Within that window, there was a difference between HA smiling faces and HA neutral faces, $t(22) = 2.7$, $p = .012$, but not between LA smiling faces and LA neutral faces, $t(22) = 0.2$, $p = .869$. This is the same pattern of interaction as was found for the attractiveness ratings in Study 1. There were no significant main effects or interaction effects in Channel 2 (all p's > .12).

Data Set 2: A significant main effect of attractiveness with a more positive voltage for HA faces was found for time windows of midpoints 200 ms to 330 ms. A significant main effect of emotionality with a more positive voltage for neutral faces was found for time windows of midpoints 360 ms to 490 ms. No post-stimulus time windows exhibited a significant interaction effect at the level of $p < .05$. Even for the specific time point where an interaction effect was observed in Data Set 1, the interaction effect was not close to significant ($p = .40$). For Channel 2, the data for individual participants was too unreliable to analyze.

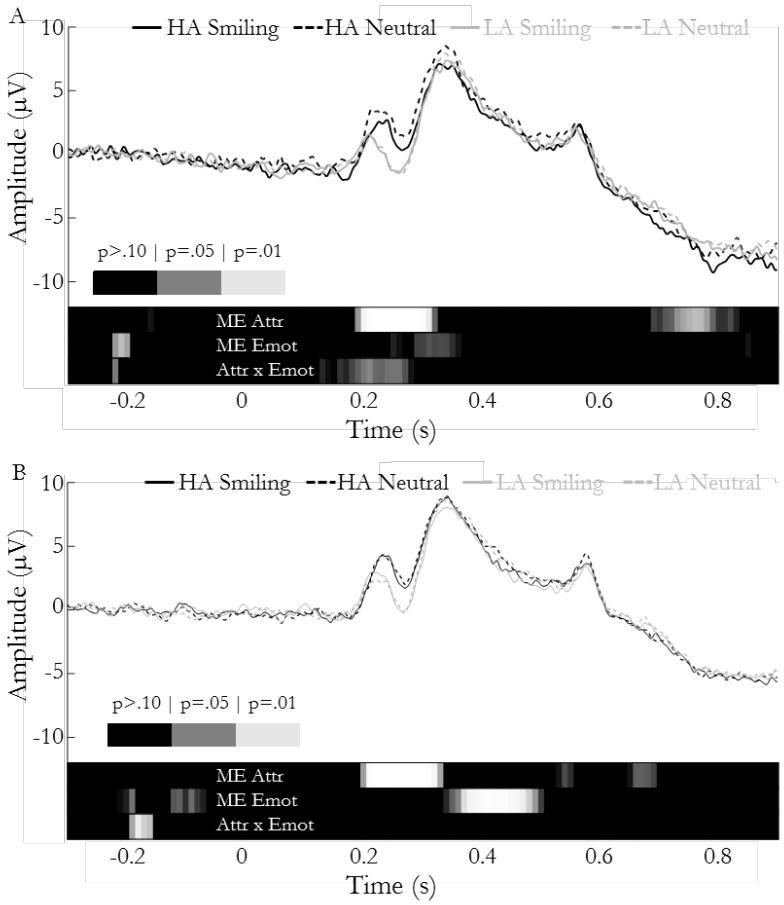

Figure 5. The ERP signals for each of the Study 2 conditions is shown across time. The output of RM ANOVAs run within various 100 ms analysis windows is also shown with the *p*-values reflected in the shading of the rectangles. The middle of the shaded rectangle is the midpoint of the 100 ms analysis window. Results are shown for Channel 1 only as Channel 2 data were not statistically significant. The stimulus onset was at time = 0 s. A) Results from Data Set 1. B) Results from Data Set 2.

DISCUSSION

The relative timing and magnitude of ERPs specific to facial attractiveness and emotional expression were observed across two sets of data using independent sets of participants but similar data methods, including identical stimuli. Consistent with prior research,[9,20] a main effect of attractiveness level was found roughly 200-300 ms after the stimulus. Also consistent with prior research,[10] a main effect of smiling versus neutral expression was observed subsequent to this, roughly 300-500 ms after the stimulus. Data Set 1 suggests that the established pattern of attractiveness ratings of HA and LA face stimuli in Study 1 may correspond to observable patterns of brain responses. The same interaction effect was observed in both Study 1 attractiveness ratings and Data Set 1 of Study 2, even though the studies had separate sets of participants and Study 1 involved explicit ratings of the faces while in Study 2 participants were not required to make explicit ratings of the stimuli but rather viewed them passively. The observed interaction pattern of early brain responses was consistent with those observed for attractiveness ratings but not for happiness ratings. Caution is required in drawing conclusions about the neural processing of attractiveness and emotionality based on this interaction effect, however, because it occurred at only a single analysis time point and failed to replicate in Data Set 2 with a larger number of participants.

It is important to note that while our observed timing of the attractiveness and emotionality ERPs was consistent with previous research, the direction of the magnitudes of the effects was inconsistent. That is, previous findings on the early attractiveness response, e.g., EPN, have tended to find that attractive faces have a more negative amplitude than unattractive faces[9, 15, 27] whereas we found a more positive response for attractive faces. Similarly, happy expressions consistently show a more negative amplitude than neutral expressions[17,18,20] whereas we found a more positive response for smiling versus neutral faces. This consistent inversion of the direction of the observed effects can be explained by the fact that we used a bipolar EEG setup and our EEG channel 1 was defined with the posterior lead, Oz, as the positive input versus the anterior lead, FP1/2, as the negative input. Most other EEG setups assess each location roughly independently in relation to a general or grand mean measure. Therefore, we still believe that our results show patterns of brain response for main effects of attractiveness and emotionality consistent with the published literature.

A novel finding in addition to that of the potential timing of the interaction effect was that the smiling conditions were associated with participants having fewer valid trials. Based on the observations made by researchers during data collection, it is suspected this resulted from participant facial responses such as smiling, as well as from laughter. This is supported by the fact that Channel 2 – which would have been more likely to pick up on mouth movements through the temporalis muscle – had artifacts beyond those in Channel 1 only for the smiling conditions. Previous research has found differential facial expressions in viewers wherein more negative expressions are made when viewing unattractive faces.[28] Our results are better explained by the chameleon effect – whereby people tend to unconsciously mimic the expressions of others[29] – as we observed more smiling behaviors in participants viewing happy expressions regardless of attractiveness level. The different levels of artifacts do not account for the observed interaction effect in brain responses though because the interaction effect found in ratings and brain responses was specific to different levels of attractiveness and the chameleon effect is only related to emotional expression.

GENERAL DISCUSSION

The current studies investigated the effects of attractiveness and smiling through behavioral and electrophysiological methods. In Study 1, self-report ratings were used to establish sets of high attractive (HA) and low attractive (LA) faces based on average participant responses and to determine a differential interaction effect of attractiveness and emotional expression based on attractiveness and happiness ratings. A difference in attractiveness ratings was found between smiling and neutral HA faces but not between smiling and neutral LA faces, whereas a difference in happiness ratings as a function of emotional expression was present for both the HA and LA categories but was of different magnitudes. Study 2 found that attractiveness categories were differentiated in the ERP prior to the differentiation based on emotional expression. The number of artifacts in Channel 2 varied with emotional expression, irrespective of attractiveness level. This was likely due to participants' mimicry of the smiling expressions as a result of the chameleon effect[29] and thus seemed to be reflective of a behavioral response lacking an interaction effect. The interaction effect of attractiveness and emotional expression, when observed, overlapped with the early attractiveness response and occurred before the emotionality-based response. Furthermore, a difference was found between the smiling and neutral images for HA individuals but not LA individuals, indicating that the observed ERP interaction effect was consistent with the self-report ratings of attractiveness but not of happiness. While such a set of findings supports an overall conclusion that smiles have a direct and initial impact on the perception of individuals' relative attractiveness prior to emotionality being explicitly processed, suggesting independent brain processing, the failure to replicate the neural interaction effect leaves open the possibility for parallel or sequential dependent processing with an interaction effect that takes place concurrent with or following the brain processing of emotional expression.

The establishment of four subsets of faces (HA smiling, LA smiling, HA neutral, and LA neutral) may be applicable in many studies of psychology with particular benefits if the same set of stimuli are used repeatedly across studies. For example, the subset may be used to test the chances of being hired based upon a mock employee resume and an included photo from the given subset in an industrial-organizational study. Further studies examining attractiveness with a similar approach could focus on establishing a subset of female faces for each condition while controlling for age-related effects. The FERET database[19] may be ideal for this purpose based on the large number of faces included in it and the fact that it is provided for free use in scientific research. An alternative could be the establishment of a stimulus repository so that various studies are able to use the same set of face images in order to increase consistency amongst research studies and to make replication easier.

In assessing the findings of these studies, it is important to note their limitations. First, female ratings of male faces were used, and therefore findings may not generalize to male ratings of female faces, especially considering the differences in how men and women perceive facial attractiveness and smiling. While happiness is the most attractive emotional expression in females, happiness is the least attractive male expression, and pride, the least attractive female expression, is the most attractive male expression.[30] An evolutionary hypothesis frames sex differences in the attractiveness of emotional expressions as due to a difference in mating strategies wherein women seeking long-term partners find smiling faces more attractive but women seeking short-term partners prefer masculine faces over smiling faces.[13] Therefore, we could surmise that part of our observed variance in ratings may have resulted from differences in participants' internalized mating strategies. Another consideration is that there may be less difference in the facial expressiveness of high and low attractive men compared to high and low attractive women,[14] further indicating that results may be different if males rate images of females.

Second, in Study 1, some participants rated some individuals as happier when displaying a neutral expression as compared to a happy expression. One individual was rated in such a manner overall. We created the HA and LA categories based solely on average attractiveness ratings. Happiness ratings were examined in a global sense to verify that smiling images were rated as happier than images with a neutral expression, but we did not look at relative happiness ratings for individual images until after running Study 2. It would likely be better for future studies using these same categories to exclude images in the LA category rated as less happy when smiling and possibly even any rated as similarly happy when smiling. More generally, future studies would benefit from checking that ratings of stimuli are consistent with what they should be on an individual basis in addition to on a global basis and from subsequently eliminating stimuli from analysis that do not pass this methodological check.

In addition, a substantial portion of the images were of individuals who were 20-40 years older than the participants. Images of older individuals were included in the image set to allow for broader conclusions about patterns in the perception of attractiveness and emotionality. This was beneficial in examining differences in the overall ratings of depicted individuals in Study 1 and thus in the discovery of the interaction effects in the ratings. However, it proved to be a confound in later analyses when the individuals were divided into groups based on their attractiveness levels, as older individuals tended to be rated as less attractive than younger individuals. Therefore, attractiveness was not the only characteristic that varied between the LA, MA, and HA groups, restricting our ability to draw conclusions. Since the raters were college-aged females, it was not surprising that images of older individuals were generally judged as less attractive. If the raters had been older, they may have found the older individuals more attractive than the participants used in the current studies, though there is evidence for a clear youth bias in attractiveness regardless of the age of the rater.[31] Furthermore, there was less ethnic diversity in the HA category than the LA category after sorting. This could have come about due to either a relatively homogenous set of participants or a discrepancy in the range of ethnicity in the images as a function of age. In future studies on the effects of attractiveness and emotional expression, age-related confounds could be avoided by using a wider range of images of individuals who are all of a similar age to that of the participants.

Overall, participants seemed to smile or laugh more while viewing smiling images than while viewing neutral images based on an increase in artifacts. Anecdotally, certain images within the smiling and neutral categories seemed to be associated with a high number of artifacts, but only the category of stimuli, not the individual image being displayed at a particular point, was recorded. Therefore, we are unable to determine whether participants reacted differently to individual stimuli, i.e., if particular images made participants laugh or smile, leading to an abundance of artifacts. Future studies may wish to record the presentation order of all individual stimuli so that differences in reactions to individual images can be analyzed.

Processing of attractiveness was expected to occur earlier than processing of emotion, but it was not previously clear whether emotional expression needed to be processed before it could influence perceived attractiveness. Our results partially imply modular neuroarchitecture for the neural processing of face qualities,[11] wherein explicit processing of emotionality is not required to occur before an interaction of attractiveness and emotional expression. Study 1 also suggested the existence of an interaction between attractiveness and perceived happiness levels, but we are not able to draw conclusions about the existence or timing of that interaction on a neural processing level due to the placement of our electrodes. A location such as Pz may have spoken to a

happiness-based interaction, as suggested by ERPs resulting from emotional stimuli centered on a more parieto-central location (see Figure 2 of Calvo *et al.*[18]; Figure 1 of Schupp *et al.*[16]). Further research with electrodes at a variety of locations could help to fully clarify the nature of the neurological combination of facial attractiveness and emotional expression.

REFERENCES

1. Sunnafrank, M., and Ramirez Jr., A. (2004). At first sight: Persistent relational effects of get-acquainted conversations. *Journal of Social and Personal Relationships, 21,* 361-379. doi:10.1177/0265407504042837
2. Lennon, S. (1988). Physical attractiveness, age, and body type. *Home Economics Research Journal, 16,* 195-203. doi: 10.1177/1077727X8801600304
3. Dion, K., Berscheid, E., and Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*, 285-290. http://dx.doi.org/10.1037/h0033731
4. Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., and Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*, 390-423. doi:10.1037/0033-2909.126.3.390
5. Dimberg, U. and Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology, 39,* 39-45. doi: 10.1111/1467-9450.00054
6. Lau, S. (1982). The effect of smiling on person perception. *The Journal of Social Psychology, 117*, 63-67. doi:10.1080/00224545.1982.9713408
7. Abel, M.H., and Watters, H. (2005). Attributions of guilt and punishment as functions of physical attractiveness and smiling. *The Journal of Social Psychology, 145*, 687-702. doi:10.3200/socp.145.6.687-703
8. O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., and Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41,* 147-155. doi: 10.1016/s0028-3932(02)00145-8
9. Werheid, K., Schacht, A., and Sommer, W. (2007). Facial attractiveness modulates early and late event-related brain potentials. *Biological Psychology, 76*, 100–108. doi:10.1016/j.biopsycho.2007.06.008
10. Calvo, M. G., Marrero, H., and Beltrán, D. (2013). When does the brain distinguish between genuine and ambiguous smiles? An ERP study. *Brain and Cognition, 81*, 237-246. doi:10.1016/j.bandc.2012.10.009
11. Bzdok, D, Langner, R., Hoffstaedter, F., Turetsky, B. I., Zilles, K., and Eickhoff, S. B. (2011). The modular neuroarchitecture of social judgements on faces. *Oxford University Press*, *22*, 951-961. doi:10.1093/cercor/bhr166.
12. Rolls, E.T. (2000). Precis of the brain and emotion. *Behavioral and Brain Sciences, 23,* 177-234. doi: 10.1017/s0140525x00002429
13. Okubo, M., Ishikawa, K., Kobayashi, A., Laeng, B., and Tommasi, L. (2015). Cool guys and warm husbands: The effect of smiling on male facial attractiveness for short- and long-term relationships. *Evolutionary Psychology*, 1-8. doi:10.1177/1474704915600567
14. Rennels, J. L., and Kayl, A. J. (2015). Differences in expressivity based on attractiveness: Target and perceiver effects? *Journal of Experimental Social Psychology, 60*, 163-172. doi:10.1016/j.jesp.2015.05.012
15. Chen, J., Zhong, J., Zhang, Y., Li, P., Zhang, A., Tan, Q., and Li, H. (2012). Electrophysiological correlates of processing facial attractiveness and its influence on cooperative behavior. *Neuroscience Letters, 517*, 65-70. doi:10.1016/j.neulet.2012.02.082
16. Schupp, H. T., Öhman, A., Junghöfer, M., Weike, A. I., Stockburger, J., and Hamm, A. O. (2004). The facilitated processing of threatening faces: An ERP analysis. *Emotion, 4*, 189-200. doi:10.1037/1528-3542.4.2.189
17. Rellecke, J., Sommer, W., and Schacht, A. (2012). Does processing of emotional facial expressions depend on intention? Time-resolved evidence from event-related brain potentials. *Biological Psychology, 90*, 23-32. doi:10.1016/j.biopsycho.2012.02.002
18. Calvo, M. G., and Beltrán, D. (2013). Recognition advantage of happy faces: Tracing the neurocognitive processes. *Neuropsychologia, 51*, 2051-2060. doi:10.1016/j.neuropsychologia.2013.07.010
19. Phillips, P. J., Raus, P. J., and Der, S. Z. (1996). *FERET (face recognition technology) recognition algorithm development and test results.* Army Research Laboratory.
20. Zhang, Z., and Deng, Z. (2012). Gender, facial attractiveness, and early and late event-related potential components. *Journal of Integrative Neuroscience, 11*, 477-487. doi:10.1142/s0219635212500306
21. Calvo, M. G., and Nummenmaa, L. (2015). Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion*, 1-26. doi:10.1080/02699931.2015.1049124
22. Balconi, M. and Pozzoli, U. (2003). Face-selective processing and the effect of pleasant and unpleasant emotional expressions on ERP correlates. *International Journal of Psychophysiology, 49,* 67-74.
23. Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: MIT Press.
24. Shields, S. M., Morse, C. E., and Nichols, D. F. (2015, March). *An exploration on the reduction of artifacts in EEG studies.* Poster presented at the annual SYNAPSE undergraduate neuroscience conference, Asheville, NC.

25. Hurless, N., Mekic, A., Pena, S., Humphries, E., Gentry, H., and Nichols, D. F. (2013). Music genre preference and tempo alter alpha and beta waves in human non-musicians. *Impulse*, 1-11.
26. Willis, J., and Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*, 592-598. doi:10.1111/j.1467-9280.2006.01750.x
27. Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., and Spencer, D. D. (1994). Face recognition in human extrastriate cortex. *Journal of Neuropsychology, 71*, 821-825.
28. Principe, C. P., and Langlois, J. H. (2011). Faces differing in attractiveness elicit corresponding affective responses. *Cognition and Emotion, 25*, 140-148. doi:10.1080/02699931003612098
29. Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893-910. doi:10.1037/0022-3514.76.6.893
30. Tracy, J. L., and Beall, A. T. (2011). Happy guys finish last: The impact of emotion expressions on sexual attraction. *Emotion, 11*, 1379-1387. doi:10.1037/a0022902
31. McLellan, B., and McKelvie, S. J. (1993). Effects of age and gender on perceived facial attractiveness. *Canadian Journal of Behavioural Science, 25*, 135-142. doi:10.1037/h0078790

ABOUT THE STUDENT AUTHORS

Paige Arrington graduated in May 2012, after the completion of Study 1. Stephanie Shields and Caitlin Morse graduated in May 2017.

PRESS SUMMARY

Prior research has shown that attractiveness and emotionality affect perception independently and interact to influence how a person perceives others, though the timing of the neural processing of this interaction is unknown. The present studies aimed to investigate this interaction and how it occurs through behavioral, specifically self-report, and physiological, specifically electrophysiological, methods using faces that varied in both attractiveness and emotional expression. Faces of high attractiveness were rated higher by participants on happiness in addition to attractiveness when smiling, but no effect of emotional expression on attractiveness ratings was found for faces of low attractiveness. A neurological interaction effect was observed in one data set coinciding with a main effect of attractiveness and prior to a main effect of emotional expression, suggesting independent processing of attractiveness and emotion. However, a neurological interaction effect was not seen in a second data set, though evidence was found consistent with modular processing of attractiveness and emotional expression. Further research is necessary to clarify the nature of the neurological combination of facial attractiveness and emotional expression.

How to Become a "Real Chicagoan" in No Time: The Promise and Pedagogy of Walking Tourism

*Jacob Henry**

Department of Anthropology, Memorial University of Newfoundland, St. John's, NL, CANADA

*Student: jlhenry@mun.ca**

ABSTRACT
This study takes seriously the tourist's desire to feel like a local and examines how walking tour guides work toward fulfilling that desire. The paper examines some of the techniques used by urban walking tour guides to convey local cultural cues. The tourist, armed with these cues, may feel able to fit into a new culture as a quasi-insider. Through qualitative methods, primarily participant observation, the researcher identifies three tactics that guides implement to make the tourist to feel like a local. These tactics are labeled agent alignment, urban alchemy, and material action. These tactics take place within a borderzone, the liminal time-space between insider and outsider status. A successful guide facilitates the border crossing, allowing the tourist to transition from tourist to perceived 'real Chicagoan.' However, the unsuccessful guide forces tourists to exit the borderzone unchanged, still as tourists. These findings highlight the uniqueness of walking tourism as a niche tourism and wade into the conceptual milieu of 'localism' and 'the local.'

KEYWORDS
Walking Tourism; Urban Tourism; Tour Guides; Localization; Interculturalism; Urban Alchemy; Agent Alignment; Chicago

INTRODUCTION

The tourists, a class of design students from a Mississippi university, huddled around their guide, Gerry. Backed by the grand marble staircase of the Chicago Cultural Center, Gerry (pseudonyms throughout) might have seemed somewhat muted in his earth tone, untucked button-down shirt, cargo pants, and well-worn baseball cap. However, as he began his well-rehearsed prose, using a thicker-than-normal Chicago accent, he commanded the tourists' attention. First, Gerry boasted about the quality of his start-up touring company: "we are the number one on TripAdvisor walking tours in Chicago," (they aren't). He then proselytized on Chicago's importance to the fields of architecture, music, and food and described how Chicago could only be understood as a mosaic of neighborhoods. Casually gesturing to the staircase behind him, Gerry told the tourists that the Cultural Center was the gateway to Chicago and that both the building and the city are places the tourists would not forget. Once he finished with the introduction to the city, Gerry paused for a moment, seeming to reflect on the task he and *these* tourists were about to undertake. After a beat, as if to size up the seriousness of this particular group, he offered the tourists the critical promise, "if you ask questions and pay attention, you'll become a real Chicagoan in no time."

This study explores how Gerry's guides try, with varying success, to fulfill the promise of turning tourists into 'real Chicagoans.' Tour guides in Chicago intuitively know to include the promise of 'realness' in their half-day neighborhood walking tours. They, and the many websites and apps competing with them for tourist attention,[1-3] understand that tourists, especially walking tourists, want to transcend 'being tourists' by 'becoming local.' The guides serve as gatekeepers to a knowledge about the city, a type of symbolic capital (in lay terms, insider knowledge)[4] that the tourists access through participation in ritual. Tourism, as Crick recognized long ago, is an act of transfer: economic capital turns to cultural capital.[5]

What does it mean to be a 'real Chicagoan?' The tourists certainly do not gain a deep understanding of a locality's cultural fabrics and economic anxieties after a morning tour in a neighborhood. There is no assertion that the tourists actually become local in the most literal sense. Instead, guides attempt to instill a local competence; the tourist should feel enlightened, intrigued, and better able to fit into a local situation than when the tour began. The tourist, armed with cultural cues, may feel like an insider for the remainder of their stay and on subsequent visits to the city. This process can be conceptualized as the transmission of a most basic, simplified *habitus* of a neighborhood or city. Traditionally, *habitus* is used to describe a "quasi-conscious" mental operation in which behavior and norms are predisposed based on social structures.[6] The caricatured *habitus* of walking tours skips the structural cause and jumps right to the predisposition effect; the tourists are taught to frame their behavior by asking how a real Chicagoan might approach a given situation. That is, they are empowered to make strategic calculations based on their understanding of 'Chicago *habitus*' --at least that is the guide's hope.

To date, scholarly writing on 'becoming local,' acquiring some level of *habitus*, has primarily focused on international tourism. Backpackers in India want to blend into the toured community[7] and international conservation volunteers in Thailand want to live the 'Thai way.'[8] However, Gerry's promise in the domestic tourism context remains largely unexplored. To gather data, the researcher conducted participant observation on multiple walking tours in Chicago neighborhoods. The subsequent analysis revisits classic concepts in tourism studies like the borderzone and authenticity, framed within the context of domestic walking tourism. The researcher will also show how at least three touring tactics need to be successfully implemented for the guide to facilitate the crossing from tourist to 'real Chicagoan.' *Agent alignment* describes how guides position themselves relative to the toured community and the tourists. *Urban alchemism* is adapted for the current study to describe how guide metanarratives shape the urban landscape and how the guides help tourists make sense of the city. *Material action* is the process in which guides utilize movement, physicality, and material consumption to supplement the touristic metaphor of border crossing from tourist to 'real Chicagoan.'

Two tours are presented as case studies. One tour, led by Alex, was by all accounts successful. The tourists smiled, took ample photographs, and asked questions that proved they were engaged with Alex's narrative. At the end of this tour, Alex employed a rhetorical device (see below) which suggested the tourists had successfully 'become local.' Another tour, led by Fred, yielded different results and is presented here as a cautionary tale. On this tour, the tourists seemed disengaged, distant, and occasionally antagonistic toward their guide. Fred was unable to transmit the cultural cues needed to enlighten the tourists or to see them invested in the locality. While each tour operated in a unique context with unique tourists, the (in)ability of the guide to facilitate the crossing to 'real Chicagoan' likely contributed to the varying end results.

METHODS

Tourism in Chicago is big business. In 2015, over 50 million tourists visited the city. This generated $935 million in tax revenue.[9] The same year, Chicago was ranked as the 9th most internationally-visited American city. It also consistently ranks in the top ten domestically-visited cities and is often the only Midwestern city on such lists.[10] Such an impressive tourist presence encourages entrepreneurship among locals who spin their own narratives and start their own tour companies. Gerry founded Out and About Chicago after retiring and now makes it his primary business to share Chicago with others.

Out and About employs fifteen guides, all of whom have come to tourism from other professions. Many are retired and some maintain other, more substantial, employment. Because most of the guides treat their work with Out and About as hobby-like rather than necessary employment, they tend to only work tours as their schedules allow. While Gerry works almost every tour, the other guides are interchangeable topical generalists. The exceptions to this are Out and About's gastro-tourism specialists and an architecture guide. The company offers about a dozen highly customizable tours, some are in Chicago's downtown Loop and others are in the city's neighborhoods. Personal tours can be booked for $25 to $40. Group discounts are offered to schools and corporate organizations. Both tours showcased in this study were school groups and presumably were offered these discounts.

At the beginning of each observed tour, Gerry delivered an opening monologue giving the tourists a brief overview of what was to come. After Gerry's introduction, the large tour groups were divided into subgroups of about ten people and matched with a guide. One tour showcased in this study took place in the downtown Loop and the other occurred in the Mexican-American enclave, Pilsen. Each tour lasted a few hours, the better part of a morning, and ended at the group's pre-determined lunch location. The guides did not reconvene after the tour was completed. Both tours covered about a mile of ground in mild Fall weather, which seemed comfortable to the researcher. However, during post-tour reflection, both guides indicated that the routes had been too long and planned to ask Gerry to shorten them for future tours.

The Loop tour was presented by Alex, a middle aged attorney who became a "certified" Chicago tour guide after he moved to the city from Los Angeles ten years ago. His resume includes work with the Cultural Center of Chicago and Gerry describes him as a, "virtual encyclopedia of contemporary and historic Chicago." Alex chose to work with Gerry and Out and About Chicago after his main career as a lawyer "took off" and demanded more of his time. Out and About's small business style allows Alex to keep guiding, which he enjoys, and practice law. The tourists on this tour were college students enrolled in a design program at a large Mississippi university; they were spending a week in Chicago studying skyscraper architecture and art. The professors leading the trip knew Gerry and had used Out and About for their bi-annual field trip for as long as they could remember. Both of the professors were complimentary of Gerry, Alex, and the entire experience with Out and About.

The Pilsen tour was led by Fred, a retired teacher who claims he, "didn't want to give up teaching, [he] just wanted to give up grading papers." Fred taught and still resides in the suburbs. While he did teach courses on Chicago, he has never lived within city limits. Fred seemed unfamiliar with the Pilsen tour; he carried a white plastic binder which contained a written version of the tour and referenced it liberally. He is proudly of Irish heritage, and while Pilsen was never predominantly inhabited by Irish immigrants, he frequently brought them into the narrative as exemplarily of the immigrant struggle. The tourists were high school

students from a northwestern suburb who were enrolled in a Spanish class. The school was affiliated with the Catholic Church and the students were all in uniform. Teachers and classroom aides also walked along with the group as chaperones. These adult tourists' behavior and level of engagement was similar to that of their students.

The researcher, in a sense, was also a tourist. Participant observation allows the body of the researcher to become the instrument of data collection.[11] Thus, in addition to recording the actions of the tourists, the researcher too felt the excitement, lulls, boredom, and shock on tour. While the researcher's own feelings are not foregrounded in this analysis, they undoubtedly inform it.

During ethnographic fieldwork, a researcher hopes the observed group will conduct business as usual, avoiding what is commonly referred to as the Hawthorne effect.[11] This normalcy is typically acquired through building rapport, but given the transitory nature of tourism, the best that could be achieved for the current project was innocuity. The researcher walked at the back of the group and did not initiate any interactions with the student-tourists and only briefly spoke with their instructors. Neither the researcher nor the tour guides noted any major changes in tourist behavior due to a researcher's gaze, though this possibility cannot be totally ruled out. After the initial introductions in which Gerry presented the other tour guides and the researcher (along with his academic affiliation and the erroneous fact that he was a graduate student) to the tourists, they more or less ignored the research process. In order to constantly remind the tourists that they were being studied, the researcher wore clothing with recognizable branding from his academic institution. The tourists could also see the researcher making jottings in a notebook.[12] After each tour, these jottings were converted and expanded into what Bernard calls analytic and descriptive field notes.[11] This project utilizes notes from about fifteen hours in the field observing tour groups and conversing with Gerry. The project design was pre-approved by Gerry as well as professors at Elmhurst College. While the institution lacks an official IRB process, this independent research was supported by the College and department.

Finally, taking cues from newer forms of research such as participant action research,[13] initial research findings were shared with Gerry, who was enthusiastic to learn about the process. The main theme of these conversations was representation in Pilsen. Gerry confessed that he has always been uncomfortable giving tours in Pilsen but could not understand the reasons for this affective reaction. After lengthy conversations, the researcher suggested that Gerry's discomfort may arise from controlling the narrative about a group of people, Mexican-Americans, to which he does not belong. Gerry readily agreed that such representation is problematic. He realized that in other neighborhoods he could use his own lived experience, whereas in Pilsen he had to rely on "their" experience. When asked why he did not hire Pilsen residents to give the tour, he was quick to say that he had done so in the past but it, "hadn't worked out." This conversation concluded with the researcher recommending Gerry approach a non-profit student housing complex in the neighborhood with a position offer. However, over a year after the fieldwork concluded, Out and About's website does not indicate this hire was made.

LITERATURE REVIEW

The question of who tourists are and what they want has intrigued scholars for the last fifty years. While some scholars believe that tourists search for the real lived experience of the toured[14] others propose that tourists want to invert their regular lives in order to temporarily play a queen or peasant.[15] Other scholars propose that tourists simply want to have fun, preferably in an individualized, hyper-realistic setting.[16] In a critical paper to tourism studies, Erik Cohen acknowledges that all these positions have some empirical support but 'tourists' as a homogenous group do not exist.[17] Cohen proposed a typology to categorize tourists by their level of integration into local life. This model attempted to account for all tourists, from those who just want a distraction from their jobs to those who search for profound experiences. Bruner also found differentiation between tourists, recognizing that sites with wide appeal attract people with varying expectations.[18] However, after conducting multi-sited ethnographies, Bruner noted that tourists with similar demographics and characteristics seem to self-delineate by type-of-tourism and venue.[19, 20]

Urban walking tourists seem to be somewhat delineated from other categories. Most notably, walking tourists blur the distinction between local and tourist. They are often domestic tourists, sometimes even from the same city or the nearby suburbs.[21, 22] To them, tourism is educational and cultural *work*. They want to understand the city on a deeper level and gain the cultural cues to move beyond the tourist bubble (in their own home). These tourists resemble what Butcher termed 'New Moral Tourists.' They are like amateur anthropologists who are self-aware enough to realize that they, the tourists, have an impact on the places they visit and thus they tread lightly in order to respectfully study the local culture they enter.[23] The grounded nature of urban walking tourism is seen as a perfect option for metro residents and tourists from further afield to access areas of a city that seem complex, unfamiliar, or intimidating[24] while also avoiding the massive touristic industries that are perceived to destroy the environments they extol.[23] To maintain the classic binary of the tourist at play and the host at work[5, 25] may not provide a complete picture of walking tourism. Instead, while the tourists are studying hard, their guides[26] and the random locals on the street who 'contribute' to tours[27] are sometimes less than serious and may poke fun at the contrived nature of the touristic situation.

New Moral Tourists are reflective and critical learners, when prompted.[20] They are the newest iteration of tourist-learners in long history of tourism-as-overt-pedagogy. From the Grand Tour[16] to today, sites have been designed to accommodate schoolchildren and thus serve an overt educational function.[18] Walking tourism is especially well-suited to create a pedagogy of the city.[27] Many guides believe they are providing a public service by spreading niche knowledge as public historians.[30] They frame the encounter between tourist and toured and often teach a celebration of difference and localism.[22] Thus, walking tour guides' urban alchemism (see below) is always an act of pedagogy, even without formal students, as it teaches the tourists a way to know the city. While some scholars question whether tourism is a just instrument of education,[28] it continues to be a complex form of cultural transmission. It is not unusual for walking tours to serve as school-sanctioned pedagogical field experiences and overlap with institutional education at all levels.[27, 29]

The discussion of 'real Chicagoans' requires a brief foray into one of the most enduring themes in tourism literature: *authenticity*. In a much citied work, Dean MacCannell suggests that it is every tourist's quest to find life as it is authentically lived in the tourist destination. However, the tourist will always fail to find the true 'backstage' lived experience of the toured community.[31] In this view, the world of tourism is divided into public and private spaces for the tourist to navigate. Authenticity is juxtaposed with superficiality and whatever is easily accessed by the tourist is inauthentic.[32] Cohen and Cohen take a different view, asserting that authentication is an ongoing process, not an innate quality. Sites can be authenticated based on emotions and beliefs of the tourists (hot authentication) and also based on official agencies' and academics' recommendations (cool authentication).[33] This means that authenticity is often contested and disputes between the 'from below' and 'from above' authentication sources are common. Bruner expands the discussion of contestation by suggesting there are at least four meanings of authenticity. Sites may be authentic if they are recognized by so-called locals; by the tourists who imagine the site; if they are originals, not reproductions; or if they are 'coolly authenticated' by authoritative agencies.[18] Bruner is generally more critical of authenticity than Cohen—he adopts a constructivist approach that rejects the notion that there is a hidden 'real' version of the toured culture.[34] Instead of asking what makes a 'real Chicagoan' he might ask who makes the 'real Chicagoan,' paying special attention to the struggle over the interpretations of emerging cultural discourse.[18] The *culture on tour* is real. The tourists, the toured, and the tourism in which they engage are real. Authenticity, to Bruner, is not a debate over front and backstage but rather it is about the socio-political processes which brought both cultures to this particular display.[20]

Becoming a 'real Chicagoan' certainly must include celebrating the small and seemingly insignificant characteristics that make the city unique.[22] Walking tourists are quite literally putting feet to pavement as they explore 'localized' urban areas. They breathe the air of the city and bump elbows with locals on crowded sidewalks. The tourists gaze and are gazed upon as they become part of the urban landscape. This contrasts with other forms of urban mass tourism, like bus tours, which are more confining and sterile experiences.[30] By getting the tourists 'out amongst the folks' (folks who sometimes interrupt tours), guides present a city that would be recognizable to locals, as it is imagined by the tourists, and in a non-simulated form. While guides may 'schmaltz' the stories or add small embellishments, the Chicago of the walking tours is usually *authentic*.[18, 27] Walking tourism with a political message may even take up the same questions about cultural constructivism as Bruner and act as a space of dissent against hegemonic narratives of place.[35] Even walking tour guides who partner with institutions of cultural production are often independent contractors who play at the boundaries of cool authentication and socially sanctioned cultural narratives.[21]

Recently, scholars have asked, with some alarm, why tourists continue leave home when cultural amenities and opportunities are often available in their home cities.[36, 16] This concern seems especially relevant for the urban tourism sector. The post-modern cityscapes of North America supposedly blend together; they are themed and simulated so that the world can be represented in compact spaces, even within a single mall.[16, 37] Why visit the original Chicago, the post-modern argument goes, when you can have a better, simulated Chicago in your hometown? Walking tour guides fight this de-differentiation by emphasizing the important and unique cultural features, beyond the simulacra, of their localities. Their stories, their livelihoods, and arguably portions of their identity are bound to the premise that real, not just hyper-real, places are worth teaching and exploring. These tours turn urban features that may seem ordinary or homogenized into moments of localized meaning.[35] It is not the showcasing of simulation that makes walking tourism interesting, it is the pedagogical nature of 'becoming local' which promises the tourist something they cannot get from theme parks: it is this promise and pedagogy that entices walking tourists to leave home, even in the post-modern era.

DISCUSSION

MacCannell's frontstage-backstage is not as helpful for understanding Gerry's promise as Bruner's concept of the borderzone. The walking tourists do not move continually toward a more authentic version of the toured, but rather enter and attempt to successfully cross a metaphorical borderzone, the liminal time-space between tourist and local *habitus*. According to Bruner, the borderzone exists both as a moment within reality but also as a theatrical imaginary with a defined beginning and end.[20] In Bruner's classic example, the Maasai transform themselves from villagers to colorful performers to entertain tourists. The tourists

are mobile and their presence initiates the 'point of conjuncture' while the Maasai are seen as permanent fixtures of the Kenyan landscape.[34]

In walking tourism, every tour occurs in the borderzone. The tourist begins the tour (enters the border) as an outsider, a non-member of the physical and social space they are touring. During the tour (in the borderzone), the guide uses at least three tactics to help the tourist 'become local.' If the guide is successful, the tourist will end the tour (cross the other side of the border) with a new knowledge and the cultural cues to avoid 'looking like a tourist.' Unlike Bruner's tourists on safari, the domestic walking tourists are likely to return to the place they are touring. Thus, in walking tourism, there is less emphasis on 'peak experiences' and more focus on how it "takes about three hours to become a New Yorker"[27] or how one can "become a real Chicagoan in no time."

Like Bruner's Maasai, urban tour guides are locals who come out to meet the tourists within the borderzone encounter. However, unlike the Maasai who conclude each performance having brought their tourists to the *end* of the borderzone, the urban walking tour guides deliver their tourists *across* it. While promotional advertisements in Kenya may suggest that the tourists will *see* the real Maasai, there is certainly no promise that they will become a 'real Maasai in no time.' This is a critical difference between the two types of tourism. For the Maasai (and Bruner), the borderzone is an in/out binary, whereas the walking tourists' experience changes drastically depending on the guide's facilitation and *how* they exit the borderzone—did they 'become local' or are they still a tourist?

The borderzone in the original and current contexts, remains a creative space for the invention of culture.[20] In a correctly facilitated borderzone crossing, cities and neighborhoods become constructed as comfortable and welcoming while notions of exclusion or unfamiliarity are dismantled.[22] The tourists must share common rituals, narratives, and practices with the toured community and distinguish themselves from *true* outsiders.[38] Upon crossing the borderzone—after 'becoming local'—the tourists are happy, thank the guide, and are amiable and optimistic about any future interaction in the city, even if they are only staying for dinner. When the promise to become a 'real Chicagoan' is unfulfilled, tourists react differently. They might be combative or disengaged, talking amongst themselves and disrespecting the guide. In order to successfully facilitate the transition, the guide must correctly employ the tactics of agent alignment, urban alchemism, and material action.

Agent Alignment

At the beginning of a typical walking tour, multiple culture-groups are represented: the tourist, the toured, and the guide who skillfully mediates the interactions. The guide is the critical force, facilitating the crossing between local and tourist status. Guides have previously been compared to parents who socialize their children, the tourists.[39] While this metaphor may be patronizing to the walking tourist, guides certainly do 'raise' tourists into another world. The personalities and training of each individual guide will change the nuance of her presentation,[40] however the overall alignment must not change. The successful guide needs to be an agent of the borderzone, carefully aligned as a third culture intermediary who can act with the *habitus* of the toured but who also is dedicated to transmitting knowledge.[21] Guides are those who 'go-between' cultures and borders,[40, 41] facilitating the tourists' transitions to 'inside dopesters.'[24] Guides must appear to be passionate advocates for both the toured and the tourists' causes.[32]

The agent must demonstrate competence in the "culture of the toured.' In most cases the toured do not get to speak for themselves, rather the guide funnels thousands of voices, dead and alive, to create one exciting, if not untrue story. Oftentimes, the guide is the singular adjudicator deciding which narratives get relayed and which are left out.[30] Tourists trust the guide's information to be accepted as cultural currency in the wider community. However, if the guide aligns too closely with the "culture of the toured," by only rattling off his knowledge of place, he may be perceived as simply showcasing his home and neglecting the promise of 'becoming local.' The walking tour guide is a curator as much as he is an entertainer and an educator.[24]
He directs the tourist gaze toward certain narratives and away from others, using linguistic devices like, "And I think it's so interesting,' "It's very interesting, isn't it," "Now this is the most fantastic story," and "Anyhow, it's a great story."[32] The guide must be friendly with the tourists and demonstrate that she understands their concerns and will support them as they enter an unfamiliar environment.[41] However, if the guide steers too close to the "culture of the tourists" she may be perceived as lacking insider knowledge about the toured community and thus fail to facilitate the borderzone crossing.

Urban Alchemism

The second tactic walking tour guides use to transition tourists across the borderzone is urban alchemy. This term originates with Jack Katz who describes it as, "the trick of selling versions of the public to the public."[42] For Katz, the cityscape is presented for all to observe but only for some to appropriate.[42] Wynn adapts the concept for the study of walking tourism, describing it as the utilization of space to create narrative and, especially, using leftover space to make a career.[30] Wynn's comprehensive study on New York walking tour guides aligns with the current study: guides are often not corporate or bureaucratic cogs, but rather active constructors in what counts as the "culture of the toured." Many walking tour guides are on a mission to tell a forgotten part of

their city's history. Wynn labels these guides, 'urban alchemists;' they re-enchant public space that city governments or bodies of authentication may have neglected.[30] To make leftover and less obvious spaces important to tourists, guides intuitively use MacCannell's method of sight sacralization (naming, framing, elevating, enshrinement, reproduction) which transforms background or seemly mundane elements of the urban landscape into noteworthy sites.[43, 32] After a site is sacralized--given a story and meaning-- it transforms (alchemy) from being leftover space to becoming important to the story of the city. Urban alchemy also challenges the privatization of public space by leading tourists into the otherwise non-accessible public and private territory.[16, 30, 41] Guides reclaim the street-level public stories and connect them to larger historical forces with an enthusiasm that aims to transform and enlighten the eager tourists. There is a certain irony that tourists, in this case, pay to walk in (usually) public space yet the guides see themselves as reenchanting the urban commons.[44]

Guides typically bind disparate sites together into a metanarrative in order to convey official and folk knowledge about the neighborhood, the city, and how to be local.[21] They frame the route of the tour within a larger context, some kind of point they want to make about Chicago and its residents. Guides take physical spaces and cultural imaginaries and turn them into consumable and compact narratives, a type of shorthand that the tourist can use to conceptualize and conduct herself within the larger local community. *Metanarratives* here are different from Bruner's metanarratives, which are ideas that travel unknowingly with the tourist, a type of taken for granted assumption.[20] Rather, here, they are conceptually closer to his 'tourist tales' which denotes how tourists recall their trip.[20] However, the walking tour guide's metanarratives of urban alchemy are more pedagogical than personal, and are conceptually absent in much of the literature.

Material Action

The third tactic guides use to facilitate the borderzone crossing is material action, which encompasses physicality and materialist consumption. Tourists must enter, replicate, and consume the "culture of the toured." This concept is related to Cohen's guide *animation* which also involves using physicality to enhance the tourism experience.[41] However, material action has a greater symbolic value than simply animating the tourist to increase their excitement. Material action must always be an embodied representation of the symbolic borderzone crossing taking place on tour.

Guides use a variety of physical experiences to prove that the promise of 'becoming local' is being fulfilled. One of the most common methods involves enshrined sites-- sites housed within larger sites.[14] By leading tourists to an enshrined site, the guide ensures that the group will cross numerous smaller, often real, borders (thresholds). As tourists comfortably cross these thresholds to observe what is within, they share experiences with locals entering the same spaces to see the enshrined treasure. Fine and Speer believe that such movement across thresholds creates a ritual transformation in which 'segregated tourists' become part the host community through the shared experiences.[32] This transfer of cultural capital and co-existent *habitus* is exactly what the tourists paid for and were promised.

Material action can also include food consumption (see below) or something as seemingly insignificant as removing the sticker or pin that was attached to one's clothing at the beginning of the tour.[27] Tourists may also replicate the materiality of their experience through photographs. While picture taking is a quintessential tourist activity, tourists who realize they are not getting across the borderzone seem to take fewer photographs than tourists who are in the process of 'becoming local.' One could speculate that the tourists who are 'becoming local' might understand that they will soon need to shed the tourist *habitus*, so they take an increased number of photographs. Alternatively, they might just be enjoying the tour more than those who are not on the path to 'real Chicagoan.' Regardless of the specific methods encouraged by the guide, by the end of a tour successful material action and physical engagement is required for the tourists to garner the experiences needed to cross the borderzone.

Case 1: A cautionary tale from Pilsen

Touring ethnic neighborhoods has been a middle class White American pastime since the 1920s.[45] While 'slumming' –or cautiously gazing at poverty while pretending not to—was a feature of the London upper class, urban and suburban White Americans openly and unabashedly ogled at the 'exotic' enclaves of newly arrived groups from around the world.[22, 46] These Americans were less eager to see class difference and more interested in essentialized ethnicities. Today, tourists can still wander an 'imagined Orient in miniature' in Chicago's neighborhoods, complete with Orientalist discourse.[22, 47] The stereotypes are persistent, especially in a city as racially segregated as Chicago[48] where dominant groups can make every effort to avoid seeing who they might consider 'Third World peoples' in the city, (even as they jet set off to Bangkok or Nairobi).[20]

Walking tour guides, the urban alchemists, control some amount of discourse about these racialized communities. It is undeniable that tourism reinforces dominant imaginaries and legitimizing rationales.[22] While it is undoubtedly problematic, often exploitative, for guides to overlay their own narratives of place,[49] the walking tour can open a small window for counter-hegemonic stories to be told. It is possible for guides to contest normative assertions or stereotypes and such re-framings can lead to the best tours and most engaged tourists.[30] It is possible to use the voices from the community to break dichotomies—not to advocate for

colorblindness—but to truly unpack what it means to live in a multicultural, hyper-segregated, metropolis and how social and historical factors have caused and continue to maintain an inequitable situation.[22, 50] Surely, to become a 'real Chicagoan,' one cannot ignore the neighborhoods, which even the official tourism website points to as the 'heart and soul' of the city.[51] The walking tourist should not, however, take essentialized divisions at face value,[52] even if such acritical acceptance is a quality of many real Chicagoans.

It is within this social context and complicated history that Fred, who is not a resident of Pilsen or a Latino person, enters the borderzone with his tourists. Things first went awry when Fred chose to align himself with what he perceived to be the "culture of the toured," rather than as an agent of the borderzone. Since Fred lived in the same area as the tourists, he tried to renegotiate himself as one of them. He knew that everyone on the tour was affiliated with the Catholic Church, so he made liberal reference to his Catholic faith. In order to align himself with the tourists, Fred would joke about rivalries within local Catholicism: "How do you feel about those people who go to [a rival church]? We hate 'em. Even we Catholics, we don't like people from the other side. Isn't that insane, it's not what our faith is based on." This rhetoric ended up creating a dichotomy between the Pilsen Catholic community (Latino/a) and the suburban Catholic community (White).

When discussing White Catholics, Fred easily uses "we." However, when describing Pilsen residents, he often used third person pronouns and "Mexicans," (even though Mexican-Americans would be more accurate). When presenting the National Museum of Mexican Art, he said, "they want to invite people to be invited to the culture," when talking about the pre-fire housing of Pilsen, he said, "Look at these houses! Would you see these in Mexico? I've never been to Mexico, you tell me." Even when Fred tried to speak as a learned authority on Pilsen culture, he produced awkward, stereotypical phrases constructed from an outsider's lens: "in the Mexican culture, the mother image is huge," or "the good Mexican male, you gotta be the tough guy." While this effort *may* have increased his connection with the tourists, it pushed him further from being able to deliver them across the borderzone. The plastic three-ring binder of answers to common Pilsen questions that Fred carried and referenced at almost every stop might have given the tourists a more nuanced presentation of the neighborhood and would have lacked Fred's intermittent, often incorrect, interpretation.

Fred's urban alchemy spun the metanarrative of preservation versus progress which, according to him, was one of the major issues facing communities like Pilsen (as if they are somehow pre-modern). Again, Fred constructed a dichotomous narrative that implied the gears of modernity (Whiteness) could not co-exist with the historical (Latino/a). Near the end of the tour, the group stopped outside of the home of internationally-renowned artist Hector Duarte. The house has a colorful two-story mural of a man casting off his chains painted onto an exterior wall. For Fred, this site was the epitome of his metanarrative. He turned to the tourists and asked if such a mural would be appropriate in their neighborhood. When they shook their heads, he told them that it probably was also *technically* illegal in Pilsen; however, "maybe it's against some city regulation, but they let it slide because it's important to the neighborhood." Probably unwittingly, Fred neutralized White culture as the monotonous output of modernity while positioning the Latino artist as someone with a culture so strong that it can bypass city ordinances. While this is problematic enough on its own accord, the continued dichotomy also meant that the tourists were not collecting a cultural currency that would be accepted in wider-Pilsen.

For material action, Fred showed the tourists the shrine to Our Lady of Guadalupe which is housed in St. Pius Church. Fred offered little commentary about the church or its role in the development of the neighborhood, but did emphasize the ritual of taking holy water before entering. While this signaled his comfort crossing thresholds in *a* church, he was clearly out of place here. In that moment, the church-goers were all older, quiet Latino men, not loud Irish-Catholic tour guides. Thus, even Fred's movement into semi-public space only served to highlight his inability to help his tourists become any closer to 'real Chicagoans' than they were at the start of the tour.

The tourists seemed to recognize this failure to facilitate the crossing. They were polite but inattentive at the beginning of the tour and only became more disconnected throughout the day. This eventually culminated into two major challenges to Fred. In the first, he argued with a tourist over her multi-racial identity and in the second another tourist boldly corrected Fred's (incorrect) spelling of "*Czech*." The tour ended without any borderzone exit ritual or rhetoric. Certainly, the tour took place in a challenging social milieu and the debate can be had over the appropriateness of such a tour even existing. However, it is also apparent that Fred's poor utilization of agent alignment, urban alchemy, and material action contributed to the exceptional problems he faced.

Case 2: A borderzone crossing

From the beginning of the downtown Loop tour, Alex positioned himself as an insider: he told the tourists that he used to work in the Cultural Center, was dressed in slacks which are commonplace in the Loop, and carried a Chicago Architecture Foundation bag around his shoulder. First, Alex told the tourists how much he loved walking in the city and how his job with Out and About allowed him to spend time simply being outdoors. This language probably reassured the tourists that Alex, like them, enjoys

gazing and consuming; he is not so local that he forgets to appreciate the sites. The tourists seemed to trust him as a cultural liaison: they believed in his abilities to move them from Mississippians to 'real Chicagoans.'

Not ten seconds after stepping out of the Cultural Center onto the sidewalk, Alex unexpectedly stopped the group. He beckoned for the tourists to huddle around a lamppost and pointed to a metal stamp design only a few inches in diameter. Alex told the group that it was Chicago's municipal device and after explaining the symbolism of the design, revealed that it could be found hidden in plain sight on most city buildings and structures. According to Alex, many Chicagoans never notice the stamp on mundane items, like lampposts. In this pedagogical moment, Alex aligned himself both as an expert with a keen eye and as an instructor with a desire to transmit cultural capital.

Alex's metanarrative was that Chicago is a city in a garden. The familiar slogan dates back to the city's first planners and Alex reinforced it as a local and proud narrative of Chicago. A significant portion of the Loop tour did occur within the boundaries of public parks and Alex noted, "as much green space as you can fit into a city, Chicago tries to do that." He often presented sites as special opportunities to see more of the city and the gardens. For example, Alex told the group to, "swing through the garden, because I want you to see an art piece." Not only does this reinforce the metanarrative, it also highlights his ability to utilize public space within the narrative. The implications of his statement are clear, without a guide, the tourists would miss a part of Chicago that could help them 'become local.' As Alex stood with his tourists in front of a popular outdoor art installation, the boundaries between tourists photographing the art and real Chicagoans photographing the art began to blur.

Alex's material action featured the enshrined Tiffany domes of the Cultural Center. Alex noted that the Cultural Center is, "well publicized now but used to be a hidden gem" and that it holds, "colossal, impressive, expensive treasures." The glass domes in the ceiling are easily missed if one does not look up at the right moment. When Alex shouts, "come with me I want to show you something special," before showing the tourists where to stand to see the one small piece of untarnished Tiffany glass within the dome, he is actively helping the tourists cross the borderzone.

However, his most effective utilization of material action involved food. The Loop tourists passed a Garrett's Popcorn shop, a well-known institution in Chicago. One of the tourists shouted, "someone told me no matter how long the line is, wait for it." Upon hearing this, Alex stopped the tour and used his allotted funds to buy a bag of popcorn for the group to share. Before giving it to them, he joked, "Now, nobody has an addictive personality, right?" Because the tourists had heard about Garrett's Popcorn from others who had gained 'local' status in Chicago, this moment was especially impactful in their transition across the borderzone. This physicality, eating expensive popcorn, symbolized the greater consumption of the Chicago imaginaries taking place on tour.

Alex concluded his Loop tour with all the tourists circled up in a seemingly random aisle of Macy's. After thanking them for visiting Chicago, he paused. Seeming to remember one last thing, he quietly told the group, "look up." Above their heads was yet another Tiffany glass dome, one of the largest in the city. The tourists stood in awe, they thanked Alex, knowing they never would have found this treasure on their own. Alex responded, "sometimes you just have to be a dumb tourist and just look up, it's the key to enjoying Chicago." This was more than a 'big reveal' at the end of the tour, Alex's dialogue gave the tourists permission to occasionally cross back into the borderzone to be tourists. In effect, this legitimized their acquired cultural knowledge and verified that they must have become 'real Chicagoans' as promised.

CONCLUSIONS

While these two tours were quite different, they are good examples of agent alignment, urban alchemy, and material action as conceptualized in this study implemented well and poorly. While Fred's troubles did not emanate solely from his difficulties with the borderzone tactics, failing to utilize them well contributed to the ongoing problems. Like the safari tourists Bruner observed, Fred's tourists exited the borderzone without 'becoming local.' This is not standard form for walking tours which attempt to be more grounded and pedagogical than other types of cultural tourism, such as cultural theme parks or bus tours. In contrast, Alex skillfully guided his tourists through the Loop and completed the tour with a rhetorical flourish validating their newfound local-ness and (perceived) Chicago *habitus*.

This study was limited (there can always be more participant observation), but proposes a grounded theory of borderzone tactics, reaffirms the unique qualities of walking tourism, and advances the conversation on the pedagogical nature of tourism more broadly. There may be other borderzone tactics for future research to identify. Additionally, more in-depth research into counter-hegemonic guiding narratives within racialized communities could have wide appeal both in the academic and guiding communities. Future scholars may also follow up with tourists to understand how they value and implement their new 'local *habitus*' after the walking tour concludes.

ACKNOWLEDGMENTS
The author thanks Dr. Connie Mixon and Dr. Mike Lindberg at Elmhurst College, Keji Kujjo at Bowling Green, and the people of the pseudonymous tour company in Chicago for their critical contributions to this project. Portions of this project were presented at the Elmhurst College Research Showcase (2016) and the Aldrich Conference in St. John's, Newfoundland (2017).

REFERENCES

1. Hewitt, E. (n.d.) Tourist No More: Three Secrets for Traveling like a Local, IndependentTraveler.com, http://www.independenttraveler.com/travel-tips/travelers-ed/tourist-no-more-three-secrets-for-traveling-like-a-local (accessed Oct 2016)

2. Hoeller, S. C. (2015) 10 Travel Apps That Will Make You Feel Like a Local, Time Magazine, http://time.com/4023144/travel-apps-feel-local/ (accessed Oct 2016)

3. LikeALocalGuide.Com (2016) Chicago City Guide, https://www.likealocalguide.com/chicago, (accessed Oct 2016)

4. Bourdieu, P. (1986) The Forms of Capital, in *Handbook of Theory of Research for the Sociology of Education* (Richardson, J. E., Ed.) 241–258. Greenwood Press, New York.

5. Crick, M. (1989) Representations of International Tourism in the Social Sciences: Sun, Sex, Sights, Savings, and Servility, *Annual Review of Anthropology* 18: 307–344.

6. Bourdieu, P. (1995 [1977]) *Outline of Theory and Practice* (trans. Nice, R.) Cambridge University Press, Cambridge.

7. Gillespie, A. (2006) Tourist photography and the reverse gaze, *Ethos* 34(3): 343–366. doi:10.1525/eth.2006.34.3.343

8. Broad, S. (2003) Living the Thai Life—A Case Study of Volunteer Tourism at the Gibbon Rehabilitation Project, Thailand, *Tourism Recreation Research* 28(3): 63–72. doi:10.1080/02508281.2003.11081418

9. Office of the Mayor of Chicago, Mayor Emanuel and Choose Chicago Announce Record Tourism in 2015, https://www.cityofchicago.org/content/dam/city/depts/mayor/Press%20Room/Press%20Releases/2016/April/4.26.16MayorChooseChicagoAnnounceRecordTourism.pdf (accessed April 2017)

10. National Travel and Tourism Office, Overseas Visitation Estimates for U.S. States, Cities, and Census Regions, 2015, http://travel.trade.gov/outreachpages/download_data_table/2015_States_and_Cities.pdf (accessed April 2017)

11. Bernard, H. R. (2006) *Research Methods in Anthropology* 4th ed., AltaMira, Lanham, MD.

12. Berreman, G. D. (2007) Behind Many Masks: Ethnography and Impression Management, in *Ethnographic Fieldwork* (Robben, A. C. G. M., Sluka, J. A., Eds.) Blackwell, Malden, MA.

13. Staddon. S. (2014) So What Kind of Student Are You? The Ethics of 'Giving Back' to Research Participants, in *Fieldwork in the Global South: Ethical Challenges and Dilemmas* (Lunn, J., Ed.) 249–261, Routledge, New York.

14. MacCannell, D. (2013 [1976]) *The Tourist: A New Theory of the Leisure Class*. University of California Press, Berkeley, CA.

15. Gottlieb, A. (1982) American's Vacations, *Annals of Tourism Research* 9: 165–187. doi:10.1177/004728758302200166

16. Urry, J. (1990) *The Tourist Gaze* Sage, London.

17. Cohen, E. (1979) A Phenomenology of Tourist Experiences, *Sociology* 13(2): 179–201. doi:10.1177/003803857901300203

18. Bruner, E. M. (1994) Abraham Lincoln as Authentic Reproduction: A Critique of Postmodernism, *American Anthropologist* 96(2): 397–415. doi:10.1525/aa.1994.96.2.02a00070

19. Bruner, E. M. (2001) The Maasai and the Lion King: Authenticity, Nationalism, and Globalization in African Tourism, *American Ethnologist* 28(4): 881–908. doi:10.1525/ae.2001.28.4.881

20. Bruner, E. M. (2005) *Culture on Tour* University of Chicago, Chicago.

21. Wynn, J. (2011) Elective Affiliations: Marginal Urban Characters Negotiating Legitimacy and Autonomy in Urban Culture, *International J of Social Inquiry* 4(1), 133–157.

22. Santos, C. A., Belhassen, Y. and Caton, K. (2008) Reimagining Chinatown: An analysis of tourism discourse, *Tourism Management* 29: 1002–1012. doi:10.1016/j.tourman.2008.01.002

23. Butcher, J. (2003) *The Moralisation of Tourism: Sun, sand . . . and saving the world?* Routledge, New York.

24. Schmidt, C. J. (1979) The Guided Tour: Insulated Adventure, *Urban Life,* 7(4), 441–467. doi:10.1177/089124167900700402

25. Nash, D. (1989) Tourism as a Form of Imperialism, in *Hosts and Guests: The Anthropology of Tourism*, (Smith, V., Ed.) 37–52, University of Pennsylvania Press, Philadelphia.

26. Wynn, J. R. (2007) Field Note: Guiding Ideas, *Contexts* 6: 56–57.

27. Wynn, J. R. (2005) Guiding Practices: Storytelling Tricks for Reproducing the Urban Landscape, *Qualitative Sociology* 28(4), 399–417. doi:10.1007/s11133-005-8365-2

28. Higgins-Desbiolles, F. and Powys Whyte, K. (2013) No High Hopes for Hopeful Tourism: A Critical Comment, *Annals of Tourism Research* 40: 428–433. doi: 10.1016/j.annals.2012.07.005

29. Lopez, E. M. (1980) The Effect of Leadership Style on Satisfaction Levels of Tour Quality, *J of Travel Research* 18(4): 20–23. doi:10.1177/004728758001800403

30. Wynn, J. R. (2010) City Tour Guides: Urban Alchemists at Work, *City & Community* 9(2), 145–164. doi:10.1111/j.1540-6040.2010.01322.x

31. MacCannell, D. (1973) Staged Authenticity: Arrangements of Social Space in Tourist Settings, *Am J of Sociology* 79(3): 589–603.

32. Fine, E., and Speer, J. H. (1985) Tour Guide Performances of Sight Sacralization, *Annals of Tourism Research* 12, 73–95. doi:10.1016/0160-7383(85)90040-4

33. Cohen, E, and Cohen, S. A. (2012) Authentication: Hot and Cold, *Annals of Tourism Research* 39(3): 1295–1314. doi:10.1016/j.annals.2012.03.004

34. Bruner, E. M., and Kirshenblatt-Gimblett, B. (1994) Maasai on the Lawn: Tourist Realism in East Africa, *Cultural Anthropology* 9(4): 435–470. doi:10.1525/can.1994.9.4.02a00010

35. Obrador, P. and Carter, S. (2010) Art, politics, memory: Tactical Tourism and the route of anarchism in Barcelona, *Cultural Geographies in Practice* 17(4): 525–531. doi:10.1177/1474474010368610

36. MacCannell, D. (2001) Remarks of the Commodification of Cultures, in *Hosts and Guests Revisited: Tourism Issues of the 21st Century* (Smith, V. L., and Brent, M., Eds.) 380–390 Cognizant Communication, New York.

37. Dann, G. (1996) *The Language of Tourism: A Sociolinguistic Perspective,* CABI, Wallingford.

38. McCallion, M. J. (2007) In-Groups and Out-Groups, in *Blackwell Encyclopedia of Sociology* (Ritzer, G., Ed.) Blackwell Reference Online. doi:10.1111/b.9781405124331.2007.x

39. Schuchat, M. G. (1983) Comfort of Group Tours, *Annals of Tourism Research* 10, 465–477. doi:10.1016/0160-7383(83)90003-8

40. Tsaur, S., and Teng, H. (2017) Exploring tour guiding styles: The perspective of tour leader roles, *Tourism Management* 59: 438–448. doi:10.1016/j.tourman.2016.09.005

41. Cohen, E. (1985) The Tourist Guide: The Origins, Structure and Dynamics of a Role, *Annals of Tourism Research* 12, 5–29. doi:10.1016/0160-7383(85)90037-4

42. Katz, J. (2010) Time for new urban ethnographies, *Ethnography* 11(1): 25–44. doi:10.1177/1466138109346999

43. MacCannell, D. (2004). Sightseeing and Social Structure: The Moral Integration of Modernity, in *Tourists and Tourism: A Reader* (Gmelch, S. B., Ed.) 55–70 Waveland Press, Long Grove, IL.

44. Harvey, D. (2012) *Rebel Cities*, Verso, London.

45. Graburn, N. H. H. (1983) The Anthropology of Tourism, *Annals of Tourism Research* 10: 9–33. doi:10.1016/0160-7383(83)90113-5

46. Stienbrink, M. (2012) 'We did the Slum!' – Urban Poverty Tourism in Historical Perspective, *Tourism Geographies* 14(2): 213–234. doi:10.1080/14616688.2012.633216

47. Said, E. W. (2003 [1978]) *Orientalism* 25th Anniv. ed., Vintage Books, New York

48. Bowean, L. (2016). Segregation declines in Chicago, but city still ranks high, census data show, Chicago Tribune (4 Jan), *http://www.chicagotribune.com/news/ct-segregation-declines-neighborhoods-change-met-20160103-story.html*

49. Conforti, J. M., (1996) Ghettos as Tourism Attractions, *Annals of Tourism Research* 23(4): 830–842. doi:10.1016/0160-7383(96)00010-2

50. Dyson, P. (2012) Slum Tourism: Representing and Interpreting 'Reality' in Dharavi, Mumbai, *Tourism Geographies*, 14(2), 254–274. doi:10.1080/14616688.2011.609900

51. Choose Chicago, Discover Chicago Neighborhoods, *http://www.choosechicago.com/neighborhoods-and-communities/* (accessed April 2017)

52. Wolf, E. (1982) *Europe and the People Without History* University of California Press, Berkeley, CA.

ABOUT THE AUTHOR

Jacob Henry received his B.A. in Sociology and Intercultural Studies from Elmhurst College in 2016. This work was completed as part of his independent intercultural research. He is currently pursuing an M.A. in anthropology at Memorial University of Newfoundland where he studies tourists who teach and teachers who tour, especially in southern Africa.

PRESS SUMMARY

Has a friend ever come back from vacation and said, "I really felt like a local?" It wouldn't be surprising, given that there are plenty of resources devoted to making travel more local-like. It seems everybody wants to avoid looking like a tourist. This paper explains the tactics walking tour guides in Chicago utilize to make you feel like a local. After reading this paper, you will know what to look for when guides are employing the rhetorical and symbolic methods that help you feel connected and invested in a new place.

Characterization of Rectifying and Sphere Curves in $\mathbb{R}^3$

*Julie Logan and Yun Myung Oh**

Department of Mathematics, Andrews University, Berrien Springs, MI

Student: julie@andrews.edu
Mentor: ohy@andrews.edu

ABSTRACT
Studies of curves in 3D-space have been developed by many geometers and it is known that any regular curve in 3D space is completely determined by its curvature and torsion, up to position. Many results have been found to characterize various types of space curves in terms of conditions on the ratio of torsion to curvature. Under an extra condition on the constant curvature, Y. L. Seo and Y. M. Oh found the series solution when the ratio of torsion to curvature is a linear function. Furthermore, this solution is known to be a rectifying curve by B. Y. Chen's work. This project, uses a different approach to characterize these rectifying curves.
This paper investigates two problems. The first problem relates to figuring out what we can say about a unit speed curve with nonzero curvature if every rectifying plane of the curve passes through a fixed point x_0 in $\mathbb{R}^3$. Secondly, some formulas of curvature and torsion for sphere curves are identified.
KEYWORDS
Space Curve; Rectifying Curve; Curvature; Torsion; Rectifying Plane; Tangent Vector; Normal Vector; Binormal Vector

INTRODUCTION

Consider a unit speed curve $\alpha: I \to \mathbb{R}^3$, where $I = (a, b)$ is an interval on the real number line. Since it is a unit speed curve (i.e. $\alpha'(s)$ has magnitude 1), the unit tangent vector is $T(s) = \alpha'(s)$.
Definition 1. The curvature, $\kappa(s)$, of a unit speed curve, α, is defined as

$$\kappa = \kappa(s) = |T'(s)|. \qquad \textbf{Equation 1.}$$

The principal normal vector, $N(s)$, is defined by dividing $T'(s)$ by its magnitude:

$$N(s) = \frac{T'(s)}{|T'(s)|} = \frac{T'(s)}{\kappa(s)}. \qquad \textbf{Equation 2.}$$

We then have $T'(s) = \kappa(s)N(s)$. The binormal vector which is perpendicular to both $T(s)$ and $N(s)$ is defined as $B(s) = T(s) \times N(s)$.

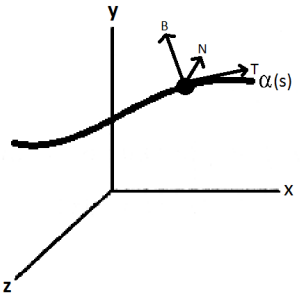

Figure 1. The Frenet-Serre frame.

These three vectors $\{T(s), N(s), B(s)\}$ form an orthonormal basis in $\mathbb{R}^3$ along the curve which is called the Frenet-Serre frame (**Figure 1**). So,

$$B'(s) = a(s)T(s) + b(s)N(s) + c(s)B(s) \qquad \textbf{Equation 3.}$$

for functions a, b, and c. It is easy to see that c is zero from the fact that B is a unit vector field. The fact that a is zero follows from the perpendicularity of T and B together with the fact that $T' \cdot B = 0$, which follows from $T'(s) = \kappa(s)N(s)$. We are ready to define the torsion of the curve.
Definition 2. The torsion, $\tau(s)$, of the curve, $\alpha(s)$, is defined by the equation $B'(s) = -\tau(s)N(s)$.
The curvature measures the deviation of a curve from being a line and torsion measures the deviation of a curve from being contained in a single plane.
Since we know that $T(s)$, $N(s)$, and $B(s)$ are all mutually perpendicular to each other, we have $N(s) = B(s) \times T(s)$. Using the facts that $T'(s) = \kappa(s)N(s)$ and $B'(s) = -\tau(s)N(s)$ we derive the following:

$$N'(s) = B'(s)\times T(s) + B(s)T'(s)$$
$$= -\tau(s)N(s)\times T(s) + B(s)\times\kappa(s)N(s)$$
$$= -\kappa(s)T(s) + \tau(s)B(s). \quad \textbf{Equation 4.}$$

Thus, we have the following Frenet-Serret formula:

$$T' = \kappa(s)N(s)$$
$$N' = -\kappa(s)T(s) + \tau(s)B(s)$$
$$B' = -\tau(s)N(s) \quad \textbf{Equation 5.}$$

According to the Fundamental Theorem of Curves, any regular curve in 3D space is completely determined by its curvature and torsion, up to position. Several characterization facts have been found over the years. We know that any curve with constant curvature and zero torsion is a circle and the curve with a constant ratio of torsion to curvature is known to be a general helix.

Now, we need to introduce three types of planes along the curve. The osculating plane to a unit speed curve $\alpha(s)$ is the plane perpendicular to $B(s)$, the normal plane of $\alpha(s)$ is the plane perpendicular to $T(s)$ and the rectifying plane of $\alpha(s)$ is the plane perpendicular to $N(s)$ (**Figure 2**).

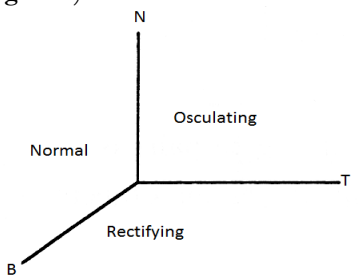

Figure 2. The normal, osculating and rectifying planes.

It has been shown that if every normal plane to the curve $\alpha(s)$ passes through a fixed point x_0 in $\mathbb{R}^3$, then the curve lies on a sphere, and if every osculating plane to the curve $\alpha(s)$ passes through a fixed point x_0 in $\mathbb{R}^3$, then the curve lies on a plane.[1] Thus, it is natural to investigate the case where every rectifying plane goes through a given point x_0.

Definition 3. A rectifying curve is a space curve whose position vector lies in its rectifying plane.[3]

The idea of rectifying curves was introduced by B. Y. Chen and he provided many fundamental properties of the curves together with classification results.[3]

For the second problem investigated in this paper, we will need the following fact about the curvature and torsion of "sphere curves", i.e. curves in $\mathbb{R}^3$ lying on a sphere.

Proposition: If $\alpha(s)$ is a unit speed curve with $\kappa \neq 0, \tau \neq 0$, then $\alpha(s)$ lies on a sphere if and only if

$$\frac{\tau}{\kappa} = \left(\frac{\kappa'}{\tau\kappa^2}\right)'. \quad \textbf{Equation 6.}[1]$$

RESULTS

The first problem we worked on was finding a necessary condition for a unit speed curve if every rectifying plane contains the point x_0 in $\mathbb{R}^3$. Since the rectifying plane is orthogonal to N, we have $(\alpha(s) - x_0)\cdot N = 0$. Thus, taking the derivative of both sides we find

$$0 = \alpha'(s)\cdot N + (\alpha(s) - x_0)\cdot N'. \quad \textbf{Equation 7.}$$

Then by substituting from the Frenet-Serret formula we have

$$0 = T\cdot N + (\alpha(s) - x_0)\cdot\left(-\kappa(s)T(s) + \tau(s)B(s)\right)$$
$$= (\alpha(s) - x_0)\cdot\left(-\kappa(s)T(s) + \tau(s)B(s)\right) \quad \textbf{Equation 8.}$$

since $T\cdot N = 0$. We know

$$0 = -\kappa(s)(\alpha(s) - x_0)\cdot T(s) + \tau(s)(\alpha(s) - x_0)\cdot B(s). \quad \textbf{Equation 9.}$$

Thus,

$$\kappa(s)(\alpha(s) - x_0)\cdot T(s) = \tau(s)(\alpha(s) - x_0)\cdot B(s). \quad \textbf{Equation 10.}$$

Then, taking another derivative of **Equation 8**, we obtain

$$0 = \left[(\alpha(s) - x_0)\cdot\left(-\kappa(s)T(s) + \tau(s)B(s)\right)\right]'$$
$$= T(s)\cdot\left(-\kappa(s)T(s) + \tau(s)B(s)\right) + (\alpha(s) - x_0)\cdot\left(-\kappa'(s)T(s) - \kappa(s)T'(s) + \tau'(s)B(s) + \tau(s)B'(s)\right)$$
$$= -\kappa(s) + (\alpha(s) - x_0)\cdot\left(-\kappa'(s)T(s) - \kappa^2(s)N(s) + \tau'(s)B(s) - \tau^2(s)N(s)\right)$$
$$= -\kappa(s) - \kappa'(s)(\alpha(s) - x_0)\cdot T(s) - \kappa^2(s)(\alpha(s) - x_0)\cdot N(s) + \tau'(s)(\alpha(s) - x_0)\cdot B(s) - \tau^2(s)(\alpha(s) - x_0)\cdot N(s)$$
$$= -\kappa(s) - \kappa'(s)(\alpha(s) - x_0)\cdot T(s) + \tau'(s)(\alpha(s) - x_0)\cdot B(s). \quad \textbf{Equation 11.}$$

So by **Equation 10**, we have $\kappa(s) = -\kappa'(s)(\alpha(s) - x_0) \cdot T(s) + \tau'(s)\left(\frac{\kappa(s)}{\tau(s)}\right)(\alpha(s) - x_0) \cdot T(s)$ and by simplification we obtain

$$\kappa(s) = (\alpha(s) - x_0) \cdot T(s)\left[-\kappa'(s) + \tau'(s)\left(\frac{\kappa(s)}{\tau(s)}\right)\right]. \quad \textbf{Equation 12.}$$

Therefore,

$$(\alpha(s) - x_0) \cdot T(s) = \frac{\kappa(s)}{-\kappa'(s)+\tau'(s)\left(\frac{\kappa(s)}{\tau(s)}\right)} \quad \textbf{Equation 13.}$$

$$(\alpha(s) - x_0) \cdot B(s) = \frac{\kappa(s)}{\tau(s)}\frac{\kappa(s)}{-\kappa'(s)+\tau'(s)\left(\frac{\kappa(s)}{\tau(s)}\right)}. \quad \textbf{Equation 14.}$$

Then working with the denominator of **Equation 13,**

$$-\kappa'(s) + \tau'(s)\left(\frac{\kappa(s)}{\tau(s)}\right) = \frac{\tau'(s)\kappa(s)-\kappa'(s)\tau(s)}{\tau(s)} = \frac{\left(\frac{\tau(s)}{\kappa(s)}\right)'\kappa^2(s)}{\tau(s)}. \quad \textbf{Equation 15.}$$

Equation 13 becomes

$$(\alpha(s) - x_0) \cdot T(s) = \frac{\kappa(s)}{-\kappa'(s)+\tau'(s)\left(\frac{\kappa(s)}{\tau(s)}\right)} = \frac{\tau(s)}{\left(\frac{\tau(s)}{\kappa(s)}\right)'\kappa(s)} = \frac{\frac{\tau(s)}{\kappa(s)}}{\left(\frac{\tau(s)}{\kappa(s)}\right)'}. \quad \textbf{Equation 16.}$$

In a similar manner, since we know that $(\alpha(s) - x_0) \cdot B(s) = \frac{\kappa(s)}{\tau(s)}(\alpha(s) - x_0) \cdot T(s)$,

$$(\alpha(s) - x_0) \cdot B(s) = \frac{1}{\left(\frac{\tau(s)}{\kappa(s)}\right)'}. \quad \textbf{Equation 17.}$$

Therefore

$$\alpha(s) - x_0 = \frac{\frac{\tau(s)}{\kappa(s)}}{\left(\frac{\tau(s)}{\kappa(s)}\right)'}T(s) + \frac{1}{\left(\frac{\tau(s)}{\kappa(s)}\right)'}B(s). \quad \textbf{Equation 18.}$$

We then let $t(s) = \frac{\tau(s)}{\kappa(s)}$. **Equation 18** becomes

$$\alpha(s) - x_0 = \frac{t(s)}{t'(s)}T(s) + \frac{1}{t'(s)}B(s). \quad \textbf{Equation 19.}$$

We take another derivative of **Equation 18**, to find

$$(\alpha(s) - x_0)' = \left(\frac{t(s)}{t'(s)}\right)'T(s) + \left(\frac{t(s)}{t'(s)}\right)T'(s) + \left(\frac{1}{t'(s)}\right)'B(s) + \left(\frac{1}{t'(s)}\right)B'(s) \quad \textbf{Equation 20.}$$

and since $(\alpha(s) - x_0)' = T(s)$,

$$T(s) = \left(\frac{t(s)}{t'(s)}\right)'T(s) + \left(\frac{t(s)}{t'(s)}\right)\kappa(s)N(s) + \left(\frac{1}{t'(s)}\right)'B(s) - \left(\frac{1}{t'(s)}\right)\tau(s)N(s). \quad \textbf{Equation 21.}$$

We then obtain

$$0 = \left(-1 + \left(\frac{t(s)}{t'(s)}\right)'\right)T(s) + \left(\frac{\kappa(s)t(s)-\tau(s)}{t'(s)}\right)N(s) + \left(\frac{1}{t'(s)}\right)'B(s) \quad \textbf{Equation 22.}$$

and we now note that T, N and B are linearly independent. Thus,

$$0 = -1 + \left(\frac{t(s)}{t'(s)}\right)' \quad \textbf{Equation 23.}$$

$$0 = \left(\frac{1}{t'(s)}\right)'. \quad \textbf{Equation 24.}$$

Then, working with **Equation 24**, we have $t''(s) = 0$, $t'(s) = c$ and $t(s) = cs + d$ for some constants c, d and arc length s. This solution satisfies **Equation 23** as well.

Here is the summary of the first result:

Theorem A: Suppose α is a unit speed curve with nonzero curvature.

If every rectifying plane contains the point x_0 in $\mathbb{R}^3$, i.e., if α is a rectifying curve, then its ratio τ/κ is a linear function.

Note: This result was obtained by B. Y. Chen in Theorem 2,[3] but it was derived here by a different method.

Our next task is to get the formula for the curvature and torsion for a sphere curve. From the Proposition,

α is a sphere curve in $\mathbb{R}^3$ if and only if $\frac{\tau}{\kappa} = \left(\frac{\kappa'}{\tau\kappa^2}\right)'$. Let $f(s) = \frac{\kappa'}{\tau\kappa^2}$. Then $\tau = f'(s)\kappa$. Thus,

$$\frac{\kappa'}{\kappa} = \tau f(s). \quad \textbf{Equation 25.}$$

Substituting $\tau = f'(s)\kappa$,

$$\frac{\kappa'}{\kappa^2} = f'(s)\kappa f(s) \quad \textbf{Equation 26.}$$

which yields

$$\kappa' = \frac{d\kappa}{ds} = f'(s)\kappa^3 f(s). \quad \textbf{Equation 27.}$$

By rearranging the terms

$$\frac{d\kappa}{\kappa^3} = f'(s)f(s)\, ds. \quad \textbf{Equation 28.}$$

Taking the integral of both sides to solve the separable differential equation we obtain $\frac{-1}{2\kappa^2} = \frac{1}{2}f(s)^2 + c_0$ for a constant c_0 and then $\frac{1}{\kappa^2} = -f(s)^2 + c$ for a constant $c = -2c_0$.

By solving for κ, we find

$$\kappa = \frac{1}{\sqrt{c - f(s)^2}}, \quad \textbf{Equation 29.}$$

and since we know that $\tau = f'(s)\kappa$, we use this formula for κ to get

$$\tau = \frac{f'(s)}{\sqrt{c - f(s)^2}}, \quad \textbf{Equation 30.}$$

for a differentiable function $f(s)$ and a constant $c = -2c_0 > 0$.

The above result can be summarized as follows:

Theorem B: For a sphere curve in $\mathbb{R}^3$, its curvature κ and τ are given by

$$\kappa = \frac{1}{\sqrt{c - f(s)^2}} \text{ and } \tau = \frac{f'(s)}{\sqrt{c - f(s)^2}} \quad \textbf{Equation 31.}$$

for a constant $c > 0$ and a differentiable function $f(s)$.

DISCUSSION

Centrodes are a particular kind of rectifying curve and they are useful in mechanics and joint kinematics.[4] Another example of application appeared in the Journal of Hand Surgery.[5]

Using the formulas for curvature and torsion of sphere curves we have found, future work would include finding a formula for the original sphere curve satisfying given curvature and torsion with specific c and differentiable function $f(s)$. This is a continuation of work by Ye Lim Seo and Yun Myung Oh published in American Journal of Undergraduate Research in Jan, 2015.[6]

REFERENCES

1. Millman, R. and Parker, G. (1977) Elements of differential geometry, Prentice Hall, Englewood Cliffs, N. J. , 33–37
2. Curvature, Torsion and the Frenet Frame. *blogimages.bloggen.be/gnomon/attach/203774.pdf*.
3. Chen, B. Y. (2003) When Does the Position Vector of a Space Curve Always Lie in Its Rectifying Plane?, *Amer. Math. Monthly* 110**,** 147–152.
4. Chen, B. Y. and Dillen, F. (2005) Rectifying Curves as Centrodes and Extremal Curves, *Bulletin of the Institute of Mathematics Academia Sinica* 33 No. 2, 77–90.
5. Weiler, P.J. and Bogoch, R. E. (1995) Kinematics of the distal radioulnar joint in rheumatoid-arthritis-an in-vivo study using centrode analysis, *J. Hand Surgery* 20A, 937–943.
6. Oh, Y. M. and Seo, Y. L. (2015) A Curve Satisfying $\tau/\kappa = s$ with constant $\kappa > 0$, *American Journal of Undergraduate Research* 12**,** 57–62.

ABOUT THE STUDENT AUTHOR

Julie Logan is an undergraduate student at Andrews University and this paper is a requirement of the honors program at Andrews University. She has interest in mathematics and has earned several awards for excellence in mathematics classes. She is planning to apply for graduate work in mathematics.

PRESS SUMMARY

It is known that for every rectifying curve, the ratio τ/κ is a linear function. In this paper, we give a different proof of the result. Besides, we also derive formulas for the curvature and the torsion of sphere curves as well.

ACKNOWLEDGEMENTS

The authors thank the Andrews University Research Office and J. N. Andrews Honors program for assistant through the 2015–2016 Undergraduate Research Assistant Scholarship. The authors are very grateful to the referees for several valuable suggestions.

Previvors' Perceptions of Hereditary Breast and Ovarian Cancer Health-related Information

Rachel Koruo[a], Marleah Dean[a], Courtney L. Scherr[b], Meredith Clements[a], Amy A. Ross[b]*

[a]Department of Communication, University of South Florida, Tampa, FL
[b]Department of Communication Studies, Northwestern University, Chicago, IL

*Student: rachelkoruo@mail.usf.edu**
Mentor: marleahdeank@usf.edu
Co-Author(s): courtney.scherr@northwestern.edu, mclements@mail.usf.edu, amyross2019@u.northwestern.edu

ABSTRACT
The purpose of this study is to identify female previvors' perceptions of hereditary breast and ovarian cancer (HBOC) health-related information. Previvors are individuals who tested positive for a harmful *BRCA* genetic mutation, which increases their lifetime risk for HBOC, but who have never been diagnosed with cancer. As a part of a larger research project where 25 qualitative interviews were conducted, this manuscript reports on the analysis of ten interviews which are most relevant to the research focus. Using the constant comparative method, themes were created and developed from the interview data. Results indicate previvors view information as a source of power. These women reported feeling personally responsible for seeking and sharing information, while also relying on medical professionals to provide credible sources of information. Furthermore, previvors emphasized a desire for medical professionals to be more informed about *BRCA* in order to assist them in making personal health decisions. This study presents the perceptions regarding HBOC information as reported by this population of previvors. The findings indicate that information is not provided in an organized way relative to their specific needs. Therefore, the authors recommend an educational intervention tool for previvors and their medical professionals.

KEYWORDS
BRCA; Communication; Qualitative; Hereditary Cancer; Health Experiences; Previvors; Medicine; Patient Perspectives; Health Information

INTRODUCTION

What if I told you that there is up to an 87% chance you will develop cancer in your lifetime, and each of your children had a 50% chance to inherit this risk? These are statistics previvors face. A previvor is an individual who is highly predisposed to hereditary breast and ovarian cancer (HBOC) due to a genetic mutation in the *BRCA1* or *BRCA2* (*BRCA*) gene, but who has not had a personal diagnosis of cancer.[1] While genetic test results provide information regarding an individual's lifetime risk for developing HBOC, such results are not always sufficient or helpful in making health decisions to prevent HBOC.[2] This mutation occurs in both men and women, however, the current study examines women's health experiences.

Within the context of HBOC, receiving information—or lacking information—regarding a genetic predisposition to such serious health conditions can produce negative effects such as emotional distress, anxiety, and uncertainty.[3] At the same time, information can also provide patients with a sense of empowerment and comfort.[4] To understand these nuances, this study investigates previvors' perceptions regarding HBOC health-related information. To understand previvors' perceptions, first, we provide an overview of previous literature examining HBOC, and then we describe the current study's research approach and methods. Ultimately, the results are analyzed and discussed to convey the perceptions of the interviewed previvors.

Receiving genetic test results

Women can undergo genetic testing to determine if they have a high lifetime risk for developing HBOC. Testing positive for a *BRCA* genetic mutation means that one of the individual's parents has the mutation, and the individual's siblings and biological children have a 50% chance of carrying the same mutation.[1] The knowledge of being at high risk for developing HBOC can drastically change a previvor's emotional well-being, as she not only worries about her own lifetime risk, but also her family's possible genetic risk.[1] Moreover, while testing positive for the *BRCA* genetic mutation reveals information about one's risk for HBOC, genetic testing results do not determine *when* and *where* a previvor will develop cancer, nor do such test results guarantee a woman will develop cancer.[1]

Because of this uncertain future, previvors often experience negative emotions.[3] A qualitative study conducted with thirteen Canadian women who tested positive for *BRCA* and received a breast cancer diagnosis experienced four negative emotions—

anger, frustration, grief, and regret.[5] More specifically, these women reported the lack of information regarding their genetic testing, rigorous screening, and preventative options produced sadness and regret.[5] It is important to note that Canada has universal insurance coverage, which is not the case in the US. It is possible that this can change the experiences of previvors, as insurance coverage changes access to care. However, the women in the study still reported lack of information and negative emotions, regardless of their access to health insurance. Results confirm that genetic testing results create a need for further information regarding HBOC, in order to make decisions. Providing patients with organized information in ways that will aid in managing such negative emotions is essential for previvors' emotional well-being.

To reduce such negative emotions and cultivate positive ones, patients can seek health information. Although each patient is sure to have unique information-seeking needs, behaviors, and experiences, analyzing the needs of a specific population—previvors—will allow for a better understanding of how information may assist in producing better health outcomes for HBOC patients. Information regarding HBOC risk can be overwhelming; however, it is necessary for previvors to be aware of their personal risk and their options in order to manage their health and make decisions. Information can be provided in an appropriate and personalized way to increase its effectiveness and helpfulness. For instance, while previvors may take comfort in having an abundance of information resources, they may find unorganized or unreliable sources as overwhelming and unhelpful.[4] An analysis of previvors' perceptions will provide a basis for creating effective information resources to manage the anxiety and uncertainty associated with testing positive for *BRCA*.

Making health decisions

After receiving genetic test results, previvors are confronted with making health-related decisions. First, previvors must decide who to disclose their genetic test results to and in what way. In a recent qualitative study conducted in Canada, the authors found women feel great concern for future generations and family members who may be at high risk, as well.[5] More specifically, some women who test positive for *BRCA* report feeling guilty for handing down their 'bad genes' to their children, and worry about the risk their children may have inherited.[1] While some women perceive their positive test results as a dead end, others see it as an opportunity to take preventative action.[1]

Second, previvors can decide to monitor their bodies through cancer screening appointments to detect cancer early or undergo preventative surgeries to prevent the diagnosis of cancer.[4] Yet, the decision-making process is far from simple, and can sometimes consume patients' lives for extended periods of time.[1] Thus, seeing specialists who can provide information and work with a patient to make the best decision in terms of their personal needs is essential. Armed with information, patients may feel empowered and motivated to fight the odds.[1]

However, even after decisions are made, previvors' anxiety and uncertainty may still persist.[4] For instance, electing to undergo increased surveillance means constant cancer screenings. Four common screening methods are as follows: breast self-exam, clinical breast exam, mammogram, and magnetic resonance imaging (MRI).[4] This opens up an entirely new window of uncertainty and fear as increased surveillance may lead to more false-positive screening results.[6] False-positive results are defined as identifying normal breast tissue as suspicious and incentivizing unnecessary biopsies.[4] Receiving false-positive results may increase thoughts regarding a cancer diagnosis and create distress, anxiety, and worry.[7] False-positive results also tend to provoke a greater perceived risk among women.[7] Being reminded of their cancer risk right before each cancer screening is likely to have an impact on a patient's well-being. In fact, some patients mention missing their mammograms to gather more information regarding *BRCA* and their options.[1] Important here is that while increased surveillance is an appropriate option for *BRCA*-positive women it does not prevent HBOC.

Previvors interested in prevention options for HBOC may consider chemoprevention and preventative surgery. Chemoprevention is defined as using medicine/drugs to prevent cancer from developing.[8] Yet, there are several ways chemoprevention may also create additional negative emotions and uncertainty. First, although chemoprevention reduces the likelihood of HBOC, it does *not* ensure that cancer will not develop.[4] Second, even though chemoprevention is noninvasive, it has potential side effects such as menopause-like symptoms, increase of blood clots, and increase risk for uterine cancer.[4] Third, many healthcare providers question whether or not this form of prevention treatment is effective (i.e., Tamoxifen and Raloxifene seem to reduce risk by only 50%).[4] In short, it is important for providers to offer previvors clear information about chemoprevention, so they can weigh their options.

The third health option for previvors is preventative surgery. Preventative surgery is the most effective risk reduction option for HBOC.[4] One example is a prophylactic bilateral mastectomy (PBM), which involves surgically removing one's natural breast tissue to reduce breast cancer risk.[1] A PBM has been found to decrease breast cancer risk by about 90%.[4] A second option is a prophylactic bilateral salpingo-oophorectomy (BSO) —removing one's ovaries and fallopian tubes in order to prevent ovarian cancer—which reduces one's risk for ovarian cancer by 80% and risk for breast cancer by 50%.[4, 9] It is important to note that

while these procedures do not guarantee that cancer will not develop, they do reduce a previvor's risk to below that of a women with no mutation.[1] As such, previvors need to understand their risk and not perceive it to be greater than it actually is.[4]

With the previous literature in mind, this study asked the following research question: What are previvors' perceptions regarding HBOC-related health information?

METHODS AND PROCEDURES

Recruitment and participants

After receiving IRB approval from the University of South Florida (IRB#: Pro00022422), participants were recruited through Twitter and Facebook using an IRB-approved flyer. Eligible participants were: 1) age 18-years or older, 2) female, 3) *BRCA*-positive, and 4) interested in research on HBOC issues.

Demographics	Total Number *N*		Selected Transcripts n	
Age Range				
20–30	5	(20%)	2	(20%)
31–40	10	(40%)	4	(40%)
41–50	3	(12%)	1	(10%)
51–60	6	(24%)	3	(30%)
61 and older	1	(4%)	0	(0%)
Race				
Asian	0	(0%)	0	(0%)
White	24	(96%)	10	(100%)
Other	1	(4%)	0	(0%)
Ethnicity				
Hispanic	2	(8%)	1	(10%)
Non-Hispanic	23	(92%)	9	(90%)
Ashkenazi Jewish Heritage				
Yes	8	(32%)	4	(40%)
No	17	(68%)	6	(60%)
Marital Status				
Married	17	(68%)	7	(70%)
Single	3	(12%)	1	(10%)
Other	5	(20%)	2	(20%)
Type of Mutation				
BRCA1	14	(56%)	5	(50%)
BRCA2	11	(44%)	5	(50%)
Type of Health Decisions				
Increased Surveillance	6	(24%)	4	(40%)
Chemoprevention	0	(0%)	0	(0%)
Preventative Surgeries	19	(76%)	6	(60%)
Health Insurance Coverage				
Workplace	19	(76%)	7	(70%)
Purchased	4	(16%)	2	(20%)
Government	2	(8%)	1	(10%)

Table 1. Participant Demographics. (*N* = 25) (n = 10). Note: The demographics for Ashkenazi Jewish Heritage were collected, because there is a significantly high frequency of *BRCA* mutations among this specific ethnic group, with 1 in 40 testing positive, when compared to the general population, where 1 in 350 will test positive.[4]

The final sample included 25 positive female *BRCA* carriers. Twenty-four (96%) of these women identified as white, and one (4%) as other. Two of these women identified as Hispanic (8%), and twenty-three as non-Hispanic (92%). Five of these women were age 20–30 (20%), ten were 31–40 (40%), three were 41–50 (12%), six were 51–60 (24%), and one was 61+ (4%). Seventeen of these women were married (68%), three were single (12%), and five identified as other (20%). Fourteen of these women carried a *BRCA1* gene mutation (56%), and eleven carried a *BRCA2* gene mutation (44%).

The ten selected transcripts represent two women age 20–30 (20%), four age 31–40 (40%), one age 41–50(10%), and three age 51–60 (30%). All ten women identified as White (100%). Nine of the women identified as non-Hispanic (90%) and one as Hispanic (10%). Seven of these women were married (70%), one was single (10%), and two identified as other (20%). Five women carried a *BRCA1* gene mutation (50%) and five women carried a *BRCA2* gene mutation (50%). See **Table 1** for participant demographics.

Data collection

After receiving informed consent from participants, the second and fourth author—trained in qualitative interviewing techniques—conducted phone interviews with participants. The interviews were recorded and transcribed by a professional transcription service. The transcribed interviews were then audio checked by research team members including the first author. Interviews lasted between 30–90 minutes. Example interview questions include: "Immediately after testing positive for the *BRCA* genetic mutation, did you feel like you needed additional information?" "Did you seek information for particular reasons?" "Do you feel like you have all of the information that you need to manage your health?" "Since testing positive for *BRCA*, have you encountered any challenges related to managing your health that you wish you could have known about ahead of time?"

Data analysis

Framed by the constant comparison approach and the selective approach to qualitative analysis, this particular manuscript analyzes 10 of the original 25 interviews, which were collected as part of a larger research study, based on the relevance to the research question.[10, 11] To begin, the first author played the interview recordings and took initial notes, followed by reading the interview transcripts on paper and taking additional notes. These activities were part of the data immersion phase.[12] Coding was conducted by the first author, with guidance from the second author. The transcripts were then read a second time, and a coding scheme was developed using the perspectives that emerged. This stage of coding is known as the primary cycle coding.[12] A list of potential codes and short descriptions were noted during the primary cycle coding. During secondary cycle coding the codes were organized and second-level codes were created.[12] Five transcripts were used to develop a codebook, and revisions were made to the codebook using the last five interview transcripts. The themes in the codebook were identified based on three criteria—recurrence, repetition, and forcefulness.[13] The interview transcripts were coded using direct quotations from the transcripts to represent the themes and subthemes of the codebook.

RESULTS

Analysis of previvors' perceptions of HBOC health information revealed four main themes: 1) information as power, 2) responsibility for seeking and sharing information, 3) trust and comfort in information sources, and 4) stage of life. Themes one and four are directly related to a previvor's state of mind and life, while themes two and three are related to the sources of information. Below, we present the four themes with exemplar quotes to support the identified themes.

Information as power

The first theme to emerge from the data provides an overall framework for how previvors view information. Broadly, previvors perceive information as a source of power, which encompasses two subthemes of control and empowerment. Control and empowerment are described below, with exemplar quotes included.

Control

The first subtheme of information as power is control. Previvors explained that information enables them to know more about their situation, which creates a sense of control over their health and fate. Control was important to previvors, because receiving positive *BRCA* genetic test results made them feel out of control. In other words, information was viewed as a tool in the decision-making process to gain control. For example, Savannah emphasized: "I feel like there's something hiding under the bed and I can't see it but I can sense that it is there. And there is nothing I can do about it, and all I want to be able to do is turn on the light. If I can turn on the light to see it's there then we can deal with it. And this was a piece of data" (p. 4).

Empowerment

The second subtheme is empowerment. Previvors described that simply having information about *BRCA* and HBOC alleviated their feelings of powerlessness, making them feel empowered. While previvors did not always act on the information they sought and found, the very idea of having the information made the previvors feel able to cope with their *BRCA* status. For instance, Iris

explained it this way: "it puts out this sense of empowerment. You can do something about it; whatever you choose to do, whether it's surgery or surveillance. You're getting ahead of something and then information is power, knowledge is -- it's good to know about this and it's good to know if you're at risk and can make decisions based on that" (p. 7).

Responsibility for seeking and sharing information

The second theme to emerge from the data notes that previvors view all parties involved in the communication of HBOC information as responsible for both seeking and sharing this information. Previvors explained that seeking and sharing HBOC information was a collective effort, and no one person can be held accountable for all their information needs. More specifically, the following subthemes emerged based on the key individuals previvors believe should be involved in seeking and sharing HBOC health information: previvor with self, previvor with medical professionals, previvor with previvor, and previvor with family/friends. These subthemes are further described below, with exemplar quotes included.

Previvor with self

First, previvors believe it is their personal responsibility to seek and share HBOC health information. Previvors noted that there is a lot of information available on HBOC, and they just had to seek out the information by doing research, asking questions, and sorting through the information. This mentality was best described by Chloe: "I'm an information seeker I guess. So I think it's out there that you got to look for it" (p. 11).

Previvor with medical professionals

Secondly, previvors believe medical professionals play an important role in seeking and sharing HBOC health information. The need for medical professionals to communicate HBOC information to previvors was evident in previvors' complaints about medical professionals' lack of *BRCA* knowledge as well as their appreciation for medical professionals who sought and shared HBOC information during their consultations. Many previvors reported their medical professionals had no knowledge or limited knowledge about *BRCA*. Some previvors even described situations where they had to explain to their medical professionals what being *BRCA*-positive meant. Previvors also emphasized that medical professionals should not only share knowledge about HBOC, but help them seek information to make informed decisions. As Iris articulated: "I think that's the most important thing, just to have access to professionals that have the data and can share that with you in a way that makes sense to you" (p. 10). Jasmine also stated: "the most important way for a person to get their information is through their doctors and their genetic counselor" (p. 20).

Previvor with previvor

Third, previvors believe it is important to seek and share HBOC information with other previvors. Many previvors talked about wanting to share the information they acquired, through their information seeking journey, in order to help other previvors. They interpreted sharing information with others as a way to raise awareness, offer support, or discuss their own personal stories. This was true for both the women who received support from others, and women who struggled to find useful information sources. To share information, previvors volunteered for HBOC organizations, interviewed with news outlets, and blogged about their experiences. In short, an implied responsibility for being a part of the previvor community was seeking and sharing information with other previvors. Mia's comment articulates this viewpoint: "the type of information which is more like how can I help? What can I do? I'm now looking for less of how can somebody else help me? If that makes any sense. That's sort of the place I'm at right now" (p. 15). Chloe shares: "I have been on local TV and I have been in the newspaper with an article because my genetic counselor is very intentional about getting the word out about prevention. So I think that more people talk about mammograms, the more in that kind of prevention, the more people hear about the previvor and that this is even an option, is huge" (p. 4).

Previvor with family/friends

Finally, previvors believe seeking and sharing information with their family members and friends is also important. Previvors reported a sense of responsibility to share the information with those around them, yet they often struggled with how to disclose their genetic test results and subsequent health decisions. Jessica states her familial responsibility: "I felt like it's important for me to get involved and find out everything about this because I need to be able to help Hannah [my niece] make some tough decisions" (p. 8). Some previvors noted receiving great support after sharing. Meanwhile, others recounted that sharing strained their family relations, and in some cases, relationships were ended altogether. One particular concern for previvors was sharing health information with their children. As such, many previvors conducted research and talked to other previvors to learn the best approach for disclosing their genetic test results to their children. For example, Chloe communicated this concern when she said: "The biggest anxiety was I have kids and I got to tell them I'm doing these surgeries. How do I explain it?" (p. 7).

Trust and comfort in information sources

The third theme to emerge from the data presents the ways in which previvors perceive information sources. Previvors reported the need to trust the information sources that they utilized. This perception derived from the belief that many information

sources are unreliable or biased, which causes frustration and confusion. In other words, trust was a prerequisite for feeling comfortable in using the information source for decision making. Previvors' key information sources included the following: medical professionals, Internet/media, and social support. The following sections describe the trust/comfort, or lack, that previvors experience with each of the key information sources, with exemplar quotes included.

Medical professionals

The first source was medical professionals. Medical professionals were a key source of information to previvors. Many previvors expressed strong relationships with doctors; some even had their physician's cell phone numbers. Previvors emphasized the importance of having a personal connection to their medical professionals, which enhanced their trust for the professionals and gave them confidence in the information the professionals provided. In other words, if the previvor trusted the medical professional, she was more likely to act on the provided information. For instance, Sophie noted: "I'm just kind of follow the provider's lead, that there's kind of—assuming that they're covering everything that's important, which is usually true, but it isn't necessarily always true" (p. 11).

In contrast, and as discussed earlier, many providers were not always knowledgeable about HBOC, which made previvors critically reflect on the information. Maggie said it this way: "I questioned the ability and the agenda of a lot of doctors. And in my own experience in my area, that is saturated with medical knowledge in cutting edge research. I still find doctors that aren't knowledgeable. So it's definitely being choosy about source" (p. 14). Although previvors can mistrust certain medical professionals, they also acknowledge their medical professionals' clinical expertise in healthcare.

Internet and the media

The second source was the Internet/media. Previvors found that much of the available online sources were "cloudy", with a few reputable websites that they were comfortable using. For example, Charlotte stated: "I think I didn't want to go online and do an exhaustive search of all the information because I felt that that was—It's just an unfiltered barrage of information that I wasn't sure if I could handle" (p. 11). One important issue that produced discomfort and distrust with online sources was how the media sexualized breast cancer. Mia mentions this when she said: "I also feel like I've been living with blinders on because breast cancer is super I think sexualized and a lot about the breast. We don't say that on anything else. We don't say, 'Save the liver.' Or like we say, 'Save the person.' Why is it with breast cancer -- even ovarian cancer, we don't say, 'Save the ovaries' we say, 'Save the person.' In breast cancer we say, 'Save the tatas,' or 'Save the boobies'" (p. 17). Some previvors noted their appreciation of Angelina Jolie's disclosure of her positive *BRCA* test result, because they felt it produced more information about *BRCA* and HBOC. Yet at the same time, others felt that she should say more because many of the online personal stories, such as blogs, could be "confusing and contradictory." Overall, previvors wanted to use the Internet for information, but also understood that issues of reliability and credibility exist.

Social support

The third source was social support. Previvors described using support groups as supplemental information sources, after consulting their medical professionals. Previvors believed support groups were a more relatable source of information, as opposed to online clinical data. However, sometimes the information shared during social support meetings made previvors feel uncomfortable because they feared similar issues might happen with them. Previvors directly stated that they felt they could trust the information from the support groups because it was "practical information" from someone who has experienced it firsthand. For example, Charlotte mentioned: "I felt like getting the information from the genetic counselor and my doctor was beneficial but I think more the emotional support and knowing like the reality of stuff and what is my life going to look like and that came more from the FORCE group" (p. 16).

Stage of life

The fourth theme to emerge from the data was previvors' perceptions of information needs related to their stage of life. Previvors navigated and interpreted HBOC information based on whether they were young adults, mid-life, or later life. Young adults include previvors in their 20's and 30's, but who have not yet had children and are not married. The following paragraph describes their specific needs and struggle to find information specific to their young age. Previvors in mid-life have established relationships and sometimes children, but are pre-menopausal. The Mid-life category below describes how their information needs changed, sometimes in relation to their careers/relationships and children. The later life category includes previvors who have well established lives and are post-menopausal. This section explains how their information needs are no longer relative to child-bearing or their careers. Exemplar quotes are included for each stage of life.

Young adult

The first category was young adulthood. Young adulthood refers to previvors who were not married and/or did not have children. Previvors in their early 20's and 30's who do not have children struggled to identify relevant information about *BRCA*.

Some previvors reported their medical professionals did not know how to advise younger *BRCA*-positive patients. For instance, Sophie explained: "When I first – met with just like my regular OB/GYN last summer and told her that we were thinking of starting a family but we were going to look into IVF and PDG first and she was like, she thought it was really great that we were doing it but she was so surprised that we were in that position because she was like, I have never met anyone who found out that they had *BRCA* mutation before they had kids" (p. 18). Previvors during this stage of life discuss the difficulty of receiving insurance coverage for increased cancer surveillance, due to their young age. In addition, many young previvors expressed the loneliness associated with testing positive for *BRCA* because they did not know many other young women going through similar experiences. Young previvors reported often avoiding social support group meetings because they are not "tailored" to their age. Finally, younger previvors who had established romantic relationships found it easier to seek out HBOC information, because they felt like they were in a more stable situation.

Mid-life

The second category was mid-life. Mid-life previvors had established romantic relationships, oftentimes with children, but were pre-menopausal. In this stage of life, previvors' information needs related to coping with the decisions they had made or were planning to make as it related to preventative surgery. On one hand, mid-life previvors expressed that sometimes they found it easier to make preventative health decisions because they already completed their families and had partners to support them in the recovery process. For example, Chloe stated: "I was glad I didn't know about my positive *BRCA* until I had already had my kids because that would be a hard time and maybe that's why it was easier for me" (p. 18). On the other hand, mid-life previvors explained that their careers and family demands also made it difficult to cope with their genetic predisposition and undergoing preventative surgeries.

This was illustrated by Kenzie when she stated: "When I was teaching, I hated it and I would always be like, 'Wow! If I got cancer now, at least I'd get some time off.' But now it's like, 'Okay. If I get cancer now, I could probably still handle my classes. I'll try to get to my internships.' It's literally always in my mind because I know it's such a huge possibility. And really, I do want to get breast surgery but I just can't right now with my schedule" (p. 5). Other information needs for mid-life previvors included coping with early on-set menopause after undergoing a preventative oophorectomy and disclosing their *BRCA* genetic mutation to their children.

Later life

The last category was later life. Previvors in this stage of life were in well-established relationships, were often retired, and were post-menopausal. While receiving *BRCA* positive genetic test results did not go away with time, later life previvors did explain that their information needs changed. Previvors in later life articulated that it was easier to make decisions about undergoing preventative surgeries to reduce their HBOC risk, because they no long were concerned about careers and/or child-bearing. For instance, Iris stated: "I had the oophorectomy right away so that really wasn't -- that to me that was a no brainer. I wanted to do that because that was my biggest fear and I really didn't need my ovaries anymore at that point because of my age" (p. 12).

In sum, previvors viewed health information as power, and also acquired a sense of responsibility to share identified health information with others including previvors, family members, and friends. The types of information sources previvors turned to were determined by their perception of trust and comfort with the information source, as well as their stage of life. The main valued source of information for previvors was their medical professionals.

DISCUSSION

Addressing information as power

The most important finding of this study was that previvors viewed information as a source of power. Although participants admitted being overwhelmed with information seeking and decision making at times, overall previvors believed information was essential to be in control and feel empowered. This finding extends previous research examining breast cancer patients, which found women diagnosed with breast cancer were primarily concerned with receiving comprehensive information in a timely fashion in order to feel empowerment.[14] It is important to note that this previous research was conducted in Iran, although the need is reflected among previvors interviewed in this study as well.

Addressing the responsibility of medical professionals

The primary source of information —regardless of the quality or quantity—for patients immediately after receiving a diagnosis are the medical professionals involved in the diagnosis. This interaction can be over the phone, through a letter, and in person. There is not only a perceived responsibility by previvors for medical professionals to provide adequate information, but also a responsibility for medical professionals to provide, at the very least, reliable sources of information to patients. Specifically, the findings of this study indicate that medical professionals need to have more information regarding HBOC to provide previvors with. Having to explain what your diagnosis means to your medical provider creates a sense of distrust/discomfort among previvors, since it is the provider's responsibility to educate and empower patients. A recent study which used an interactive risk

assessment tool to effectively increase knowledge about HBOC, collect family history, and spark patient-provider discussions about HBOC screening, shows that interactive tools are useful for communicating HBOC information but are not by any means a replacement for the patient-provider discussions that are essential to successful health outcomes.[15] It is important to note that this population of previvors turns to the Internet/media as sources of information when the information provided by medical professionals is not adequate, but that the internet/media should not be a replacement for patient-provider discussions. For instance, articles on Angelina Jolie's decision to undergo PBM could influence *BRCA*-positive women who are seeking information online.[16] However, these articles emphasize Jolie's gender identity as a sexual icon, partner, mother, and humanitarian.[16] *BRCA*-positive previvors need reliable health information that is unbiased from their medical professionals or reputable health sources, unlike the information provided through unfiltered Internet/media articles.

Addressing the function of support groups

The current study also demonstrates the function of support groups and previvors' perceptions of the relationships/information obtained through support groups. Previvors turn to support groups for emotional support, personal experience, and other relevant information. Support groups are very central to the previvors' experiences and information seeking journey. Studies suggest the support group dynamic as one that alleviates frustrations with medical professionals, while also offering emotional and experiential information in a trustworthy and safe environment.[17] This study also discusses women seeking information regarding early menopause, because physicians provide limited knowledge and internet/media was contradictory.[17] These are similar to the issues revealed in the data that indicate young previvors are not provided enough information regarding the consequences of early menopause.

Addressing the role of the previvor

Finally, this study addresses the role of the previvor and the perceptions of previvors' lived experiences. Although every previvor in this population used social support, there are nevertheless drawbacks and flaws with support groups. These flaws are particular to the stage of life that the participant is experiencing in relation to *BRCA*. Previvors felt out of place in their support groups due to the fact that they were the youngest members. A study conducted on young previvors, ages 18–24, notes that these women felt vulnerable and pressured to make decisions, although there were no well-defined guidelines for screening and prevention options for this population of young women.[18]

Disclosure varies based on stage of life, specifically depending on whether the previvor has kids, a stable relationship, or a career. They describe disclosure as something that is important, yet difficult. This is noted in a study conducted on family communication of *BRCA* results where age may affect a patient's decision to communicate results and information to family, because parents also try to protect children from anxiety producing information.[19] Another study highlights the struggle of non-married carriers and the anxiety and fear that is present before disclosing their results to their partners, although most receive positive support after disclosure.[20] It is essential that previvors are provided with the means to communicate their HBOC information to those around them, specifically those that the diagnosis could have immediate importance to such as children and siblings, or other relatives.

The current study reveals the following about previvors' perceptions about HBOC health information: 1) previvors desire information to enhance control and create empowerment; 2) previvors desire information sources that are reliable and relevant to their present situations; and 3) previvors' information needs vary based on their stage of life.

CONCLUSION

Limitations & Future Directions

In summary, this study delves into previvors' perceptions of HBOC information by analyzing qualitative interview transcripts. While previous research indicates information assists in managing uncertainty, this research study does not investigate the well-being of the interviewed previvors. Instead, the aim of this study is to present previvors' perceptions of information and highlight the gaps in information or information delivery as identified by previvors. The study emphasizes the sense of power that reliable information can provide previvors. It also reveals the need for medical professional education in regards to *BRCA* and HBOC. The main limitation of this study is the sample. The interview participants are all educated with either some or extensive college completed. Additionally, the sample is not very ethnically diverse, with primarily White participants and a few Hispanic participants.

At the same time, this study contributes to the field of health communication and hereditary cancer in three ways. First, the study demonstrates the importance of patient perspectives in understanding health information. Not only does this study show that medical professionals are helpful information sources for previvors, but so too are support groups and the Internet/media. Second, previvors' emphasis on information as a source of power suggests that medical professionals should assist previvors in the information seeking process especially as it relates to their stage of life. Finally, given this study's findings, it is clear an educational intervention tool would be helpful for previvors and their medical professionals.

ACKNOWLEDGEMENTS

The authors thank the participants for sharing information on HBOC, as well as FORCE for allowing recruitment of participants. Support for this research was provided by an internal grant from the University of South Florida.

REFERENCES

1. Roth Port, D. (2010) *Previvors: Facing the breast cancer gene and making life-changing decisions*, Avery, New York.
2. Joseph, G., and Guerra, C. (2014) To worry or not to worry: Breast cancer genetic counseling communication with low-income Latina immigrants, *J Community Genet, 6*(1), 63–76. DOI: 10.1007/s12687-014-0202-4
3. Dean, M. (2016) "It's not if I get cancer, it's when I get cancer": BRCA-positive patients' (un)certain health experiences regarding hereditary breast and ovarian cancer risk, *Soc Sci Med*, *163,* 21–27. DOI: 10.1016/j.socscimed.2016.06.039
4. Friedman, S., Sutphen, R., Steligo, K. (2012) *Confronting hereditary breast and ovarian cancer: Identify your risk, understand your options, change your destiny*, Johns Hopkins University Press, Baltimore.
5. Joseph, M., Rab, F., Panabaker, K., and Nisker, J. (2015) Feelings of women with strong family histories who subsequent to their breast cancer diagnosis tested BRCA positive, *Int J Gynecol Cancer, 25*(4), 584–592. DOI: 10.1097/IGC.0000000000000403
6. Portnoy, D. B., Loud, J. T., Han, P. K. J., Mai, P. L., and Greene, M. H. (2015) Effects of false-positive cancer screenings and cancer worry on risk-reducing surgery among BRCA1/2 carriers, *Health Psychol, 34*(7), 709–717. DOI: 10.1037/hea0000156
7. Brewer, N. T., Salz, T., and Lillie, S. E. (2007) Systematic review: The long-term effects of false-positive mammograms, *Ann Intern Med, 146*(7), 502–W133. DOI: 10.7326/0003-4819-146-7-200704030-00006
8. Koch M. (2015) Chemoprevention, Salem Press Encyclopedia of Health, Release January 2015, Available from: Research Starters, Ipswich, MA. (accessed Nov 16)
9. Metcalfe, K., Lynch, H. T., Ghadirian, P., Tung, N., Olivotta, I., Warner, E., Olopade, Ol., Eisen, A., Weber, B., McLennan, J., Sun, P., Foulkes, WD., Narod, SA. (2004) Contralateral breast cancer in BRCA1 and BRCA2 mutation carriers. *J Clin Oncol*, *22*(12), 2328–2335. DOI: 10.1200/JCO.2004.04.033
10. Lindlof, T. R., and Taylor, B. C. (2011) *Qualitative communication research methods*. London: SAGE.
11. van Manen, M. (1990) *Research lived experience: Human science for an action sensitive pedagogy*. State University of New York Press, Albany.
12. Tracy, S. J. (2013) *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. Wiley-Blackwell, Chichester, West Sussex, UK.
13. Owen, W. F. (1984) Interpretive themes in relational communication, *Q. J. Speech*, 70(3), 274–287.
14. Taleghani, F., Bahrami, M., Loripoor, M., and Yousefi, A. (2014) Empowerment needs of women with breast cancer: a qualitative study, *Iran Red Crescent Med J,* 16(11). DOI: 10.5812/ircmj.16379
15. Rupert, D. J., Squiers, L. B., Renaud, J. M., Whitehead, N. S., Osborn, R. J., Furberg, R. D., et al. (2013) Communicating risk of hereditary breast and ovarian cancer with an interactive decision support tool, *Patient Educ Couns, 92*(2), 188–196.
16. Dean, M. (2016). Celebrity health announcements and online health-information seeking: An analysis of Angelina Jolie's preventative health decision, *Health Commun*. DOI: 10.1080/10410236.2014.995866
17. Kenen, R. H., Shapiro, P. J., Friedman, S., and Coyne, J. C. (2007) Peer-support in coping with medical uncertainty: Discussion of oophorectomy and hormone replacement therapy on a web-based message board, *Psychooncology, 16*(8), 763–771. DOI: 10.1002/pon.1152
18. Werner-Lin, A., Hoskins, L. M., Doyle, M. H., and Greene, M. H. (2012) 'Cancer doesn't have an age': Genetic testing and cancer risk management in BRCA1/2 mutation-positive women aged 18–24, *Health (London), 16*(6), 636–654. DOI:10.1177/1363459312442420
19. Peshkin, B. N., DeMarco, T. A., and McKinnon, W. C. (2007) Life after BRCA1/2 testing: Family communication and support issues, *Breast Dis, 27*(1), 127–136. DOI: 10.3233/BD-2007-27108
20. Hoskins, L. M., Roy, K., Peters, J. A., Loud, J. T., and Greene, M. H. (2008) Disclosure of positive BRCA1/2-mutation status in young couples: The journey from uncertainty to bonding through partner support, *Fam Syst Health, 26*(3), 296–316. DOI: 10.1037/a0012914

ABOUT THE STUDENT AUTHOR

Rachel Koruo graduated in May 2017 with a dual degree in Biomedical Sciences and Communication. She plans to work in the communication field, before beginning dental school during the 2018–2019 academic year.

PRESS SUMMARY

Women who test positive for a *BRCA* mutation are at increased risk for developing hereditary breast and ovarian cancer (HBOC) during their lifetime, and are frequently referred to as previvors. This study identifies previvors' perceptions regarding HBOC health information. Analysis revealed previvors view health information as power and sense a responsibility to share identified health information with others including other previvors, family members, and friends. The types of information sources previvors turn to are determined by their perception of trust and comfort with the information source as well as their stage of life. The main valued source of information for previvors was their medical providers. Therefore, it is essential for medical professionals to be knowledgeable about *BRCA* and provide HBOC information to these women so they can make informed health decisions.

Determination of Fitted Size Distribution for Atmospheric Aerosols

Kaitlin M. DuPaul, Adam T. Whitten*

Physics Department, College of Saint Benedict/Saint John's University, Collegeville, MN

*Student: kmdupaul@csbsju.edu**
Mentor: awhitten@csbsju.edu

ABSTRACT
A synthetic set of aerosol optical depths (AODs) generated from a standard set of aerosol size distributions was analyzed by a parameter based particle swarm optimization (PBPSO) routine in order to test the reproducibility of the results. Junge and lognormal size distributions were consistently reproduced. Gamma and bimodal distributions showed large variability in solutions. $\bar{\chi}^2$ values were used to determine the best subset of possible solutions allowing quantification of parameters with uncertainties when using PBPSO. AODs measured by a sun photometer on a clear day (20160413) and a foggy day (20160508) were then processed by the PBPSO program for both bimodal and lognormal distributions. Results showed that in general the foggy day has smaller $\bar{\chi}^2$ values indicating that the PBPSO algorithm is better able to match AODs when there is a larger aerosol load in the atmosphere. The bimodal distribution from the clear day best describes the aerosol size distribution since the $\bar{\chi}^2$ values are lower. The lognormal distribution best describes the aerosol size distribution on the foggy day (20160508).

KEYWORDS
Atmospheric Aerosols; Size Distributions; Junge; Bimodal; Gamma; Lognormal; Particle Swarm Optimization; Inverse Problem; Aerosol Optical Depth

INTRODUCTION

Atmospheric aerosols are of particular interest due to their environmental and public health impact.[1, 2] For these reasons it is important to monitor atmospheric aerosol concentrations and size distributions. Atmospheric aerosols are tiny particles ranging in diameter from 0.001-10 μm suspended in the atmosphere. Natural aerosols are commonly made up of soil, minerals, salts, and various other chemicals mixed with water. Anthropogenic aerosol sources include industrial pollution and vehicle emissions. Aerosols are non-uniformly distributed throughout the troposphere and stratosphere with highly variable size distributions. Atmospheric aerosols can provide a surface on which chemical reactions occur, serve as condensation and ice nuclei, scatter and absorb light, and can influence the electrical properties of the atmosphere.[3]

Aerosol optical depth (AOD) is a common method for measurement of air quality and aerosol research. AOD can be inferred by comparing solar irradiance values to the standard E-490 solar spectrum created by the American Society for Testing and Materials[4] and subtracting off Rayleigh scattering terms. Higher AODs correspond to more scattering and/or absorption of sunlight, indicating that one or more of the aerosol size distribution properties has increased (i.e., particle number, particle size, particle shape, and/or index of refraction). Given a set of AODs, the aerosol size distribution may be inferred by inverting a set of Fredholm equations,

$$\tau(\lambda) = \int \pi r^2 Q_{ext}(r, \lambda, \eta) n(r) dr \qquad \textbf{Equation 1.}$$

where Q_{ext} is the Mie extinction coefficient and *n(r)* is the aerosol size distribution. Matrix inversion,[5–7] iterative,[8–10] and maximum entropy spectral[11] methods have been used to retrieve aerosol size distributions. However, inversion of **Equation 1** is an ill-posed problem classified as a NP (nondeterministic polynomial) problem meaning that the inversion can be performed using the two-step process for solving any NP problem: (1) guess a solution in a non-deterministic way and (2) use a deterministic algorithm to verify or reject the guess.[12] It is important to note that accepting a guess does not guarantee that the best possible solution has been generated; it merely indicates that the solution is one of many possible valid solutions that fit the deterministic criteria.

Recently, particle swarm optimization (PSO) algorithms that implement the two-step NP problem solving method have been used to retrieve size distributions from **Equation 1**. PSO is an iterative algorithm that uses a swarm of particles randomly generated in the solution space to identify size distribution parameters that yield calculated AODs close to those inferred from solar irradiance measurements. Closeness is quantified by a fitness function and the iterations continue until a stop criteria that depends on the fitness function is reached or until a predefined maximum number of iterations have been performed. The goal of PSO is to

minimize the fitness function by moving each particle in the swarm simultaneously toward its own and the swarm's previous best positions.[13, 14] Each particle's position is used to calculate an AOD and fitness value; the particle with the smallest fitness value is presumed to represent the best inversion of the AODs inferred from irradiance measurements and corresponds to the best position. However, given the nature of NP problems, multiple runs of PSO algorithms will in general converge to different solutions corresponding to local minima or the global minimum in the essentially infinite solution space.[15, 16]

Examples of PSO algorithms in use include Yuan *et al.*'s (2010, 2011) Stochastic Particle Swarm Optimization (SPSO)[17, 18] and Mao & Li's (2015) Improved Particle Swarm Optimization (IPSO).[19] Both use a fitness function based on the difference between measured and inferred AODs: SPSO uses the square root of the average variance; IPSO uses the inverse of half the sum of the variances. SPSO's stop criteria are (1) when changes in the best fitness function are less than 10^{-10} and (2) when 1000 iterations have been performed. IPSO stops after 30 iterations without placing constraints on fitness function convergence. The remaining difference between these two algorithms is that SPSO generates a random particle with each iteration in order to explore more of the solution space while IPSO uses a variable velocity to speed up convergence. Neither method specifies what type of boundary conditions are employed. Neither paper discusses the reproducibility of results when the algorithms are run on the same set of data more than once. There is no measure of how often their algorithms may focus in on a local minimum of the fitness function instead of the global minimum. Furthermore, Yuan *et al.*'s (2011) stop criterion of the fitness function changing by less than 10^{-10} restricts the uncertainty in AODs inferred from spectral measurements (typically 0.1-2.5%[20, 21]) by many orders of magnitude.

To address these deficiencies, a parameter-based particle swarm optimization (PBPSO) algorithm was written. Boundary conditions are such that particles moving outside the solution space are destroyed and a new random particle is generated. This boundary condition allows more exploration of the solution space. During each iteration, ten additional particles are generated to test for better solutions and then are destroyed. PBPSO's fitness function is the reduced chi-squared of calculated and inferred AODs. Like previous algorithms, the fitness function for PBPSO is used to guide evolution of the swarm towards a best solution that is, ideally, the global minimum. Unlike previous algorithms, the stop criterion is based on the relative changes in size distribution parameters, not on the fitness function or some fixed number of iterations. Using a size parameter convergence criterion ensures that a swarm has converged to a solution. The PBPSO stop criteria are (1) the size distribution parameters have a relative change of less than 10^{-30}% for 100 consecutive iterations or (2) the algorithm reaches 5000 iterations. After multiple runs of PBPSO on the same data set, the quality of each solution can be assessed using the corresponding fitness function values.

METHODS AND PROCEDURES

This research tests the reproducibility of the PBPSO algorithm for four different size distributions. This algorithm implements the following steps to solve the inverse radiation problem.

Step 1: Input aerosol optical depths at 412, 441, 463, 479, 500, 520, 556, 610, 675, 750, 778, 870, and 1020 nm. These include the World Meteorological Organization (WMO) standard wavelengths (412, 500, 610, 675, 778 nm), WMO standard wavelengths shifted out of absorption bands (463, 870, 1020 nm), and five additional wavelengths (441, 479, 520, 556, 750 nm).[22] Input the Mie extinction coefficient, Q_{ext}, for each of these wavelengths spaced at radius intervals $\delta r = 0.001$ μm. Choose bounds for the parameters depending on distribution type. Initialize a particle swarm and calculate AOD, $\tau(\lambda)$, using **Equation 2** at each wavelength, λ, for each particle. Q_{ext} was calculated by using the Bohren-Huffman Mie scattering subroutine[23] from $r_{min} = 0.04$ μm to $r_{max} = 10$ μm with a complex index of refraction $\eta = 1.5+0i$ giving a 2×9961 array. r_{min} was chosen as 0.04 μm since for $r \leq 0.04$ μm, $Q_{ext} < 0.0012$ and the contribution to the AOD is negligible. The real part of η was chosen to be between water (1.33) and silica (1.54).

$$\tau(\lambda) = \int_{r_{min}}^{r_{max}} \pi r^2 Q_{ext}(r,\lambda,\eta) n(r) dr \approx \pi \delta r \sum_{i=1}^{9961} r_i^2 Q_{ext}(r_i,\lambda,\eta) n(r_i) \quad \text{where} \quad n(r) = \frac{dN(r)}{dr} \qquad \textbf{Equation 2.}$$

Step 2: Calculate reduced chi-squared ($\bar{\chi}^2$) value (**Equation 3**) for each particle where $d = 13-\#$ of independent parameters.

$$\bar{\chi}^2 = \frac{1}{d}\sum_{i=1}^{13}\left(\frac{\tau_{i,obs} - \tau_{i,calc}}{\tau_{i,calc}}\right)^2 \qquad \textbf{Equation 3.}$$

Step 3: Compare current $\bar{\chi}^2$ with the previous $\bar{\chi}^2$ for each particle to determine if the new position is lower; if so it becomes a new local minimum, P_i, otherwise retain the previous local minimum.

Step 4: From the list of all P_i choose the one with the lowest $\bar{\chi}^2$ and set it as the global minimum, P_g.

Step 5: Introduce ten randomly generated particles and see if their $\bar{\chi}^2$s are less than P_g; if so replace P_g with the lower value.

Step 6: Check the stop criteria. (1) If the size distribution parameters have changed by less than 10^{-30}% for 100 consecutive iterations then the program is terminated (**Equation 4**). (2) If the program has gone through 5000 iterations without reaching the stop criterion (1) terminate the program. Otherwise evolve the swarm and loop back to step 2.[13]

$$\frac{1}{j}\sum_{i=1}^{j}\frac{\left|P_g[i](t+1)-P_g[i](t)\right|}{P_g[i](t)} \leq 10^{-5} \qquad \textbf{Equation 4.}$$

Step 7: Repeat steps 1-6 ten times. Examine the final $\bar{\chi}^2$s for each run's P_g: the lowest $\bar{\chi}^2$ is corresponds to the solution that is most likely to be the global minimum while the others correspond to solutions that are local minima. Since there is still no guarantee that the global minimum has been found, use a subset of the ten solutions defined by $\bar{\chi}^2 \leq 100\bar{\chi}^2_{min}$ to find size distribution parameter averages and standard deviations of the mean (SDOM). This process eliminates solutions that are deemed valid by step 2 of the NP selection process, but which represent outliers corresponding to local minima.

A synthetic set of AODs was generated from a standard set of size distribution parameters (N, α, and β) for each distribution (**Table 1**) in order to test the PBPSO algorithm. Aerosol optical depths at the 13 wavelengths were calculated for the four distributions listed in **Table 1** for an assumed set of values for the parameters (N_0, α, and β for the single mode distributions and α_0, β_0, N_1, α_1, and β_1 for the bimodal distribution) using the Bohren-Huffman Mie scattering subroutine.[23] Parameter values were chosen to give AODs that might typically be obtained from a Kipp-Zonen PGS-100 sun photometer. These AODs were then used as input to the PBPSO program to evaluate how well the algorithm could reproduce the distribution parameters. The resulting $\bar{\chi}^2$ values are a measure of the difference between the algorithm's calculated AODs and the inputted AODs. Since multiple runs might produce different parameter values, it was hypothesized that choosing a subset of the results with the smallest $\bar{\chi}^2$s and then finding the mean and standard deviation of the mean (SDOM) of this subset of parameters would give a reasonable representation of the actual size distribution. The PBPSO algorithm was run ten times for each distribution type and the results were plotted against the known distribution. These results were evaluated as a standard for identifying the best way to analyze atmospheric data.

Distribution	Equation	Parameters	Solution Space
Junge	$n_J(r) = N_0\beta r^{-(\alpha+1)}$	N_0=2.5×10⁶, α=3, β=0.3	$1.0\times10^3 \leq N_0 \leq 1.0\times10^9$ $2.0 \leq \alpha \leq 5.0$ $0.1 \leq \beta \leq 0.5$
Gamma	$n_G(r) = N_0\, r^{\beta} e^{-\alpha r}$	N_0=2.5×10⁹, α=0.1, β=0.35	$1.0\times10^6 \leq N_0 \leq 1.0\times10^{12}$ $0.05 \leq \alpha \leq 0.50$ $0.1 \leq \beta \leq 0.5$
Lognormal	$n_L(r) = \frac{N_0}{\sqrt{2\pi}}\frac{1}{\beta r} e^{-.5\left[\frac{\ln(r/\alpha)}{\beta}\right]^2}$	N_0=1.0×10⁷, α=0.5, β=0.307	$1.0\times10^4 \leq N_0 \leq 1.0\times10^{10}$ $0.05 \leq \alpha \leq 1.0$ $0.1 \leq \beta \leq 0.5$
Bimodal	$n_B(r) = \frac{N_0}{\sqrt{2\pi}}\frac{1}{\beta_0 r} e^{-.5\left[\frac{\ln(r/\alpha_1)}{\beta_1}\right]^2} + \frac{N_1}{\sqrt{2\pi}}\frac{1}{\beta_1 r} e^{-.5\left[\frac{\ln(r/\alpha_1)}{\beta_1}\right]^2}$	N_0=1.0×10⁷, α_0=0.1, β_0=0.307 N_1=2.5×10⁵, α_1=2, β_1=0.307	$1.0\times10^5 \leq N_0 \leq 1.0\times10^{11}$ $0.05 \leq \alpha_0 \leq 0.50$ $0.1 \leq \beta_0 \leq 0.5$ $1.0\times10^3 \leq N_1 \leq 1.0\times10^9$ $0.50 \leq \alpha_1 \leq 5.0$ $0.1 \leq \beta_1 \leq 0.5$

Table 1. Typical distributions and solution space. To access synthetic data sets used and results, see http://www.physics.csbsju.edu/~awhitten/pso.html.

RESULTS AND DISCUSSION

Synthetic Data

For Junge and lognormal distributions there was little to no variation in the results, but for the gamma and bimodal distributions the results seemed to vary so additional runs of the algorithm were made (1) holding α constant while allowing β to vary and (2) holding β constant while allowing α to vary.

Junge Distribution

After analyzing the results for the Junge distribution it was noticed that the plots of the resulting calculated number density distributions were identical. For all ten runs of the PBPSO program the $\bar{\chi}^2$ values are equal (2.05×10^{-13}) and although there is variance in N_0 and β, α is consistently 3.0. The variation in N_0 and β is explained by the fact that they are not independent

parameters and their variability is constrained only by the choice of solution space given in Table 1. Considered separately the ten values of N_0 and β have distributions that are narrower than a normal distribution (all values fall within the 95th percentile), but because of their dependent nature it is their product, not the product of their individual averages, that is important. The PBPSO program gave $N_0 \times \beta = 7.5 \times 10^5$ for all ten runs. Furthermore, the PBPSO program was able to exactly reproduce the AOD's for each wavelength every time.

Lognormal Distribution

The lognormal distribution of the sample data shows a generally uniform distribution with two small outliers (**Figure 2**). These two outliers prove to be runs two and eight which correspond to the results with the highest $\bar{\chi}^2$ values (~10^{-7}) and shouldn't be taken into consideration as the lowest $\bar{\chi}^2$ value for this data set is 3.3×10^{-13}. The distribution parameter values were $N_0 = 2.5 \times 10^6$ cm^{-3}, $\alpha = 0.5$ μm, and $\beta = 0.307$. As shown in **Table 2**, the output means for each variable were in agreement with the distribution parameter values when taking the standard deviation of the mean (SDOM) into account.

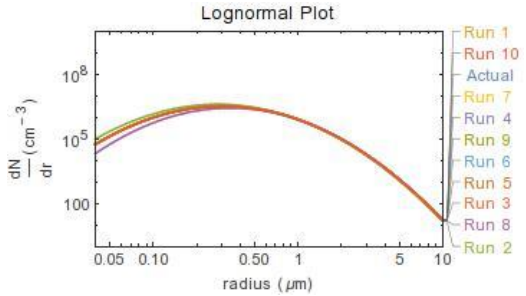

Figure 2. Size distribution functions output from the PBPSO algorithm reveal a tight grouping. The actual distribution function is hidden behind the red line.

	N_0 (cm^{-3})	α (μm)	β
Mean	2.500×10^6	0.5004	0.30698
SDOM	3×10^3	0.0003	0.00001

Table 2. Lognormal results with neither α nor β held constant.

Gamma Distribution

The gamma distributions showed some small, but significant variations. Therefore these distributions have been plotted on a smaller scale for the $\frac{dN}{dr}$ (cm^{-3}) axis in order to highlight the difference between runs. Despite these apparent differences, the PBPSO routine was able to reproduce the target AODs with $\bar{\chi}^2$ values ranging from 10^{-11} to 10^{-9}.

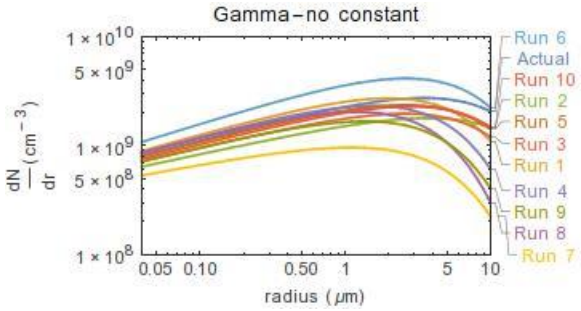

Figure 3.1

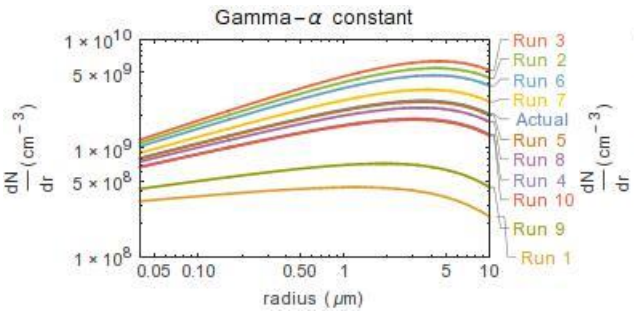

Figure 3.2

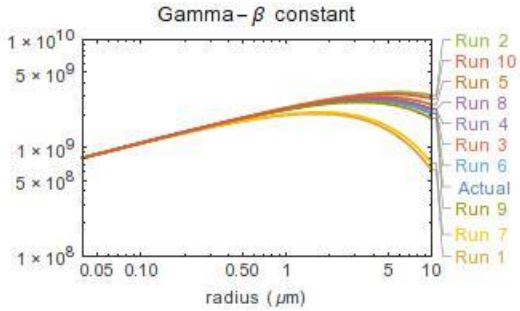

Figure 3.3

Figure 3. The actual distribution function is represented as a dark blue line in each of the graphs. The PBPSO program has more trouble reproducing the gamma distribution as shown in Figure 3.1. When α is held constant the PBPSO program can zero in on a wide range of solutions with smaller N_0 and β (a wider distribution) or higher N_0 and β (a narrower distribution) as shown in Figure 3.2. When β is held constant the PBPSO program can produce a more consistent set of solutions as shown in Figure 3.3. This suggests that falling in to a local minimum for β can greatly influence the results (see text).

When no parameters are held constant, a wide range of solutions are seen (**Figure 3.1**). To understand why this might be the case, additional optimizations were performed holding α or β constant. When α is held constant, β and N_0 vary significantly resulting in little to no overlap of the values (**Figure 3.2**). With β held constant, α varies from [0.06, 0.22], but N_0 stays relatively constant (**Figure 3.3**). This behavior is the result of the analytic forms of the Mie extinction coefficient and the gamma distribution. For visible and near infrared wavelengths Q_{ext} peaks at about 4.4 for $r < 1$ μm and then exhibits decreasing oscillations as shown in **Figure 4**.

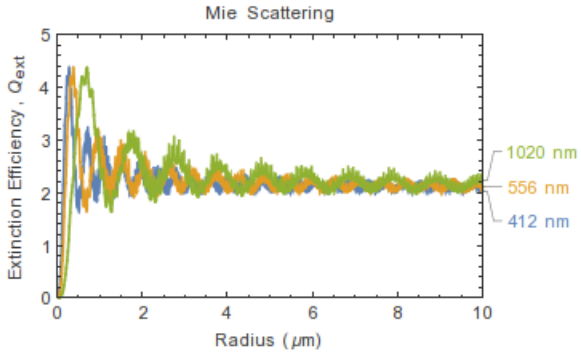

Figure 4. Mie extinction coefficient, Q_{ext}, calculated using the Bohren-Huffman subroutine for three wavelengths of light and complex index of refraction $\eta = 1.5+0i$. The values at large radii decrease to approximately one half of the peak values that occur at radii less than 1 μm.

This variation in Q_{ext} means that small changes in the contribution to AODs from small radii particles can be compensated by larger changes from large radii particles to yield the same AOD value. As β increases in the gamma distribution the r^{β} term reduces contributions to the AOD for $r < 1$ μm and increases contributions to the AOD for $r > 1$ μm. The exponential term reduces contributions to the AOD for $r > 1/\alpha$, but when combined with the r^{β} term it results in a peak in the size distribution at $r_0 = \beta/\alpha$. When α is held constant at 0.1, then $r_0 = 10\beta$ constraining the peak of the distribution to the range 1.0 μm ≤ r ≤ 5.0 μm

which is greater than the radii of the peak values of Q_{ext}. Larger values of β give larger values of N_0 because the decrease from r_0 down to $r < 1$ μm (where Q_{ext} is highest) is more rapid. When β is held constant at 0.35, $r_0 = 0.35/\alpha$ constraining the peak of the distribution to 0.7 μm ≤ r ≤ 7.0 μm. This range for r_0 is also greater than the radii of the peak values of Q_{ext}, but since β is fixed, the contribution to the AOD from small radii particles is well defined. Therefore, the variability in α and N_0 with β fixed is smaller than the variability in β and N_0 with α fixed.

For the case when neither α nor β are held constant the combined effect of the dependence on radius of both Q_{ext} and the gamma distribution makes it difficult for PBPSO to zero in on the correct solution because multiple solutions meet the fitness criterion. Therefore, to use this PBPSO routine to retrieve gamma size distribution parameters it is recommended to run the optimization multiple times and extract the best retrievals based on $\bar{\chi}^2$ values. Using solutions with $\bar{\chi}^2 \leq 100\bar{\chi}^2_{min}$ eliminates outliers corresponding to local minima and the remaining solutions are used to calculate the means and SDOMs of the size distribution parameters (**Tables 3.1-3.3**). The set parameter values were $N_0 = 2.5\times10^9$ cm^{-3}, $\alpha = 0.1\mu m^{-1}$, and $\beta = 0.35$. Not shown are the results when α and β are both held constant. This constraint consistently reproduces the correct N_0 in agreement with the results of Yuan *et al.* 2010.

	N_0 (cm^{-3})	α (μm^{-1})	β
Mean	2.50×10^9	0.19	0.348
SDOM	1.5×10^8	0.03	0.009

Table 3.1. Gamma results with neither α nor β held constant.

	N_0 (cm^{-3})	α (μm^{-1})	β
Mean	3.1×10^9	0.1	0.37
SDOM	5×10^8	constant	0.02

Table 3.2. Gamma results with α held constant.

	N_0 (cm^{-3})	α (μm^{-1})	β
Mean	2.5000×10^9	0.11	0.35
SDOM	3×10^5	0.02	constant

Table 3.3. Gamma results with β held constant.

Bimodal Distribution

The bimodal distributions shown in **Figure 5** indicate large variations in the retrieved size distribution parameters despite being able to reproduce the target AODs with small $\bar{\chi}^2$ values. **Figure 5.1** shows that when no parameters are held constant there is variability in solutions that give similar aerosol optical depths ($10^{-9} < \bar{\chi}^2 < 10^{-5}$). To examine this variability and understand which parameters PBPSO has trouble determining, α and/or β were held constant. When α is held constant (**Figure 5.2**), N_0 and β_0 stay relatively constant, but N_1 and β_1 do not ($10^{-11} < \bar{\chi}^2 < 10^{-5}$). When β is held constant (**Figure 5.3**) there is also more variability in N_1 and α_1 than in N_0 and α_0 ($10^{-9} < \bar{\chi}^2 < 10^{-7}$). PBPSO is less consistent when determining the distribution at larger radii. The source of this inconsistency is most likely the relatively flat Mie extinction profiles at larger radii for visible wavelengths of light making it difficult for the algorithm to pick out α_1. It should be noted that when α_0, α_1, β_0 and β_1 are all held constant, N_0 and N_1 values are exactly reproduced which is consistent with the findings in Yuan *et al.* 2011.

To quantify the inconsistencies, subsets of results based on $\bar{\chi}^2 \leq 100\bar{\chi}^2_{min}$ values were analyzed to find the means and SDOMs of the size distribution parameters. For no constants (**Figure 5.1**) run seven (light yellow) is the standard and solutions with a $\bar{\chi}^2 > 3.43\times10^{-7}$ were ruled out, resulting in five outliers for this set. In **Figure 5.2** run seven (light yellow) is the standard as it has the lowest $\bar{\chi}^2$ value for the data set (2.7×10^{-11}). After applying the $\bar{\chi}^2$ criterion three of the ten solutions are left. In **Figure 5.3** run six (light blue) is the standard with $\bar{\chi}^2 = 1.0\times10^{-9}$. It is much easier to identify the outliers in this plot as they all lie within the second lognormal distribution and after applying the $\bar{\chi}^2$ criterion five of the ten solutions are left. In order to have significant averages for the six size distribution parameters it is necessary to eliminate outliers.

The distribution parameters values were $N_0 = 1.0\times10^7$ cm^{-3}, $\alpha_0 = 0.1$ μm, $\beta_0 = 0.307$, $N_1 = 2.5\times10^5$ cm^{-3}, $\alpha_1 = 2.0$ μm, and $\beta_1 = 0.307$. Following the $\bar{\chi}^2$ rule stated above, the resulting size distribution parameters for no constants, both αs held constant, and both βs held constant are shown in **Tables 4.1-4.3**.

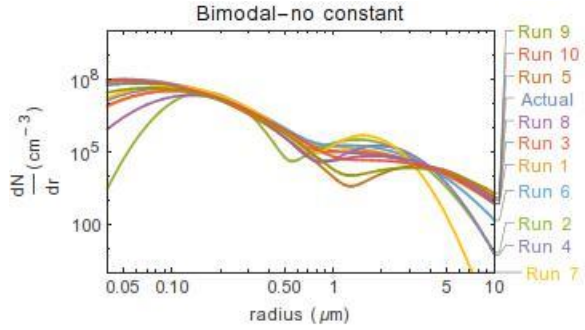

Figure 5.1

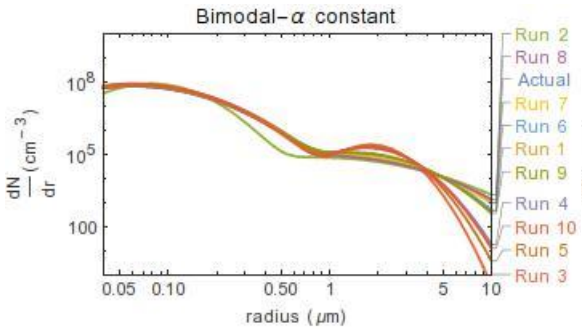

Figure 5.2

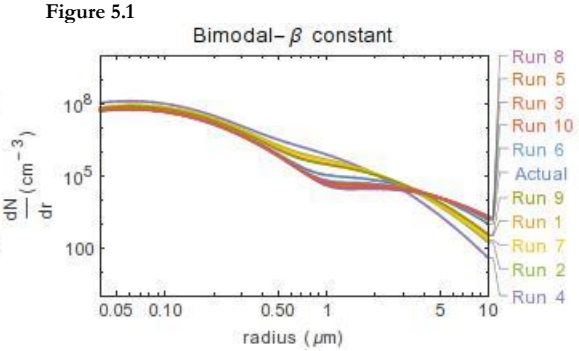

Figure 5.3

Figure 5. The actual distribution function is indicated in dark blue. Figure 5.1 shows that when all parameters are allowed to vary there are multiple solutions that reproduce the aerosol optical depths. Figure 5.2 shows that when α is held constant there is more variability in the width of larger aerosols. In Figure 5.2 the actual distribution is hidden behind the dark green line. Figure 5.3 shows that when β is held constant there is more variability in the radius of larger aerosols.

	N_0 (cm^{-3})	α_0 (μm)	β_0	N_1 (cm^{-3})	α_1 (μm)	β_1
Mean	9.20×10^{7}	0.11	0.29	2.6×10^{5}	2.5	0.24
SDOM	1.5×10^{6}	0.01	0.02	5×10^{4}	0.4	0.02

Table 4.1. Bimodal results with neither α nor β held constant.

	N_0 (cm^{-3})	α_0 (μm)	β_0	N_1 (cm^{-3})	α_1 (μm)	β_1
Mean	1.030×10^{7}	0.1	0.3056	2.8×10^{5}	2	0.26
SDOM	1.0×10^{5}	constant	0.0004	7×10^{4}	constant	0.02

Table 4.2. Bimodal results with α held constant.

	N_0 (cm^{-3})	α_0 (μm)	β_0	N_1 (cm^{-3})	α_1 (μm)	β_1
Mean	9.10×10^{6}	0.09990	0.307	1.62×10^{5}	2.8	0.307
SDOM	1.0×10^{5}	0.00004	constant	8×10^{3}	0.1	constant

Table 4.3. Bimodal results with β held constant.

Atmospheric Data

The PBPSO procedure is now applied to data taken by a Kipp & Zonen PGS-100 sun photometer on a clear day (2016-04-13) and a foggy day (2016-05-08). Aerosol optical depths inferred from irradiance measurements at the St. John's University observatory were processed ten times with the PBPSO algorithm. Solutions with $\bar{\chi}^2 > 100\bar{\chi}^2_{min}$ were eliminated and the remaining solutions were used to find the averages and SDOMs for size distribution parameters. Data from the two days was processed using a bimodal size distribution and a lognormal size distribution.

Bimodal Distribution

Aerosol optical depths at 13 wavelengths inferred from the PGS-100 and calculated for each of the tens runs of PBPSO for the clear day are shown in **Table 5**. Inferred AODs are reproduced by the PBPSO algorithm to within 3–7% for all wavelengths except for 610 nm, 750nm, and 778 nm where they are within 12–14%.

AOD from:	Wavelength (nm)													$\bar{\chi}^2$
	412	441	463	479	500	520	556	610	675	750	778	870	1020	
PGS-100	0.241	0.207	0.201	0.200	0.191	0.195	0.187	0.181	0.131	0.103	0.095	0.091	0.070	–
Run 1	0.227	0.219	0.211	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.067	1.4E-3
Run 2	0.230	0.220	0.211	0.205	0.197	0.189	0.176	0.156	0.136	0.115	0.108	0.089	0.065	1.5E-3
Run 3	0.229	0.220	0.212	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3
Run 4	0.230	0.219	0.211	0.205	0.197	0.190	0.176	0.157	0.136	0.115	0.108	0.089	0.065	1.4E-3
Run 5	0.229	0.220	0.212	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.067	1.4E-3
Run 6	0.230	0.220	0.211	0.205	0.197	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3
Run 7	0.229	0.220	0.211	0.205	0.197	0.190	0.176	0.156	0.135	0.114	0.108	0.088	0.066	1.4E-3
Run 8	0.229	0.220	0.211	0.205	0.197	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3
Run 9	0.230	0.220	0.211	0.205	0.197	0.189	0.176	0.156	0.135	0.115	0.108	0.089	0.066	1.4E-3
Run 10	0.228	0.220	0.212	0.206	0.198	0.190	0.176	0.156	0.135	0.114	0.107	0.088	0.066	1.4E-3

Table 5. Aerosol optical depths for a clear day (20160413) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a bimodal distribution.

Table 6 shows the AODs inferred from the PGS-100 and calculated for each of the tens runs of PBPSO for the foggy day. Inferred AODs are reproduced to within 0.4–7% for all wavelengths.

AOD from:	Wavelength (nm)													$\bar{\chi}^2$
	412	441	463	479	500	520	556	610	675	750	778	870	1020	
PGS-100	0.704	0.633	0.611	0.590	0.564	0.540	0.503	0.452	0.347	0.274	0.250	0.205	0.146	–
Run 1	0.691	0.652	0.620	0.597	0.566	0.539	0.490	0.422	0.352	0.285	0.263	0.205	0.141	7.0E-4
Run 2	0.691	0.652	0.619	0.596	0.566	0.539	0.490	0.422	0.352	0.285	0.264	0.206	0.140	7.1E-4
Run 3	0.691	0.652	0.619	0.597	0.566	0.539	0.490	0.422	0.352	0.285	0.263	0.206	0.141	7.1E-4
Run 4	0.693	0.652	0.619	0.596	0.565	0.537	0.488	0.420	0.351	0.285	0.264	0.207	0.143	7.2E-4
Run 5	0.691	0.652	0.620	0.597	0.566	0.539	0.490	0.421	0.352	0.285	0.263	0.205	0.142	7.0E-4
Run 6	0.689	0.651	0.618	0.596	0.566	0.538	0.490	0.422	0.352	0.286	0.264	0.206	0.140	7.2E-4
Run 7	0.691	0.652	0.620	0.597	0.566	0.539	0.490	0.421	0.352	0.284	0.263	0.205	0.142	6.8E-4
Run 8	0.687	0.650	0.618	0.596	0.567	0.539	0.491	0.424	0.354	0.287	0.265	0.207	0.140	7.2E-4
Run 9	0.690	0.651	0.619	0.597	0.566	0.539	0.490	0.422	0.352	0.285	0.264	0.205	0.140	7.0E-4
Run 10	0.691	0.651	0.619	0.596	0.566	0.538	0.490	0.422	0.353	0.286	0.264	0.206	0.138	7.4E-4

Table 6. Aerosol optical depths for a foggy day (201600508) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a bimodal distribution.

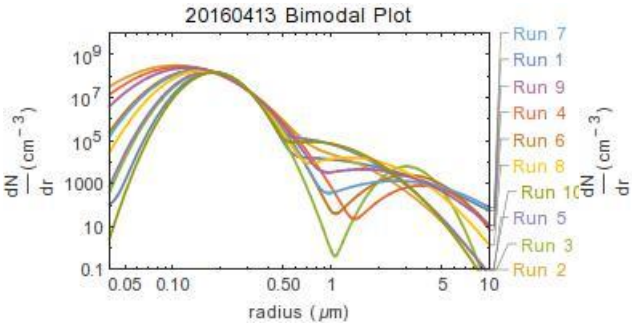

Figure 6.1

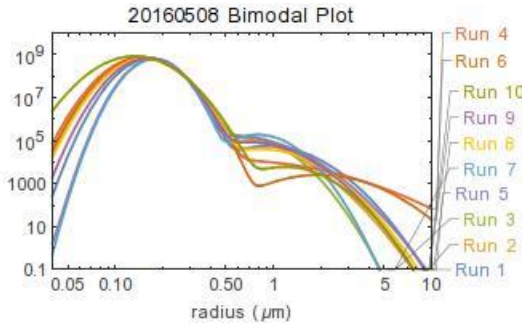

Figure 6.2

Figure 6. For the given days 20160413 in Figure 6.1 and 20160508 in Figure 6.2, all parameters are allowed to vary and there are multiple solutions that reproduce the aerosol optical depths. The reduced chi-squared values in the bimodal distribution are closer with little to no variance.

The bimodal results (**Figures 6.1-6.2**) show that the $\bar{\chi}^2$ values are all within 10^{-4} of each other with little to no variance so following the earlier stated rule all values must be considered. For the 20160413 results there are two groups of solutions roughly 40% of which are narrower with higher α_1 values and 60% that are wider with lower α_1 values. Results show a 4% uncertainty in the median radius α_0 of the small aerosol mode for the clear day and a 2% uncertainty in determining the median radius α_0 of the small aerosol mode for the foggy day. For the large mode aerosols the uncertainties in the median radius α_1 are 18% and 16% for the clear and foggy days, which confirms the assessment for synthetic data that the median radius of the small aerosol mode is easier to pick out than the large aerosol mode (**Tables 7.1-7.2**).

	N_0 (cm^{-3})	α_0 (μm)	β_0	N_1 (cm^{-3})	α_1 (μm)	β_1
Mean	2.78×10^{7}	0.167	0.147	3.5×10^{4}	2.2	0.212
SDOM	2.7×10^{6}	0.007	0.009	1.2×10^{4}	0.4	0.023

Table 7.1. Bimodal results 20160413 (clear day).

	N_0 (cm^{-3})	α_0 (μm)	β_0	N_1 (cm^{-3})	α_1 (μm)	β_1
Mean	8.6×10^{7}	0.171	0.124	7.3×10^{4}	1.12	0.214
SDOM	4×10^{6}	0.003	0.005	1.6×10^{4}	0.18	0.022

Table 7.2. Bimodal results 20160508 (foggy day).

Lognormal Distribution

Aerosol optical depths for the 13 wavelengths inferred from PGS-100 measurements and calculated for 10 runs of the PBPSO algorithm are shown in **Table 8** for the clear day and **Table 9** for the foggy day. AODs are reproduced to within 4–19% for the clear day and to within 0.2–7% for the foggy day.

AOD	**Wavelength (nm)**													
from:	**412**	**441**	**463**	**479**	**500**	**520**	**556**	**610**	**675**	**750**	**778**	**870**	**1020**	$\bar{\chi}^2$
PGS-100	0.230	0.164	0.162	0.159	0.154	0.165	0.156	0.156	0.100	0.086	0.083	0.099	0.084	–
Run 1	0.206	0.189	0.178	0.170	0.162	0.154	0.142	0.127	0.113	0.102	0.098	0.089	0.081	3.5E-3
Run 2	0.206	0.189	0.178	0.170	0.162	0.154	0.142	0.127	0.113	0.102	0.098	0.089	0.081	3.5E-3
Run 3	0.203	0.189	0.178	0.171	0.163	0.155	0.143	0.127	0.113	0.101	0.097	0.089	0.081	3.5E-3
Run 4	0.205	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 5	0.205	0.189	0.178	0.171	0.162	0.155	0.143	0.127	0.112	0.099	0.095	0.088	0.087	3.3E-3
Run 6	0.204	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 7	0.206	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 8	0.205	0.189	0.178	0.171	0.162	0.154	0.142	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 9	0.203	0.189	0.178	0.171	0.163	0.155	0.143	0.127	0.113	0.101	0.098	0.089	0.081	3.5E-3
Run 10	0.204	0.189	0.178	0.171	0.163	0.156	0.143	0.127	0.112	0.099	0.095	0.087	0.086	3.3E-3

Table 8. Aerosol optical depths for a clear day (20160413) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a lognormal distribution.

AOD	**Wavelength (nm)**													
from:	412	441	463	479	500	520	556	610	675	750	778	870	1020	$\bar{\chi}^2$
PGS-100	0.704	0.633	0.611	0.590	0.564	0.540	0.503	0.452	0.347	0.274	0.250	0.205	0.146	–
Run 1	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 2	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 3	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 4	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 5	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 6	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 7	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 8	0.695	0.652	0.618	0.594	0.563	0.536	0.487	0.421	0.353	0.288	0.266	0.209	0.142	5.5E-4
Run 9	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4
Run 10	0.690	0.651	0.619	0.596	0.566	0.539	0.490	0.423	0.354	0.287	0.265	0.206	0.138	5.2E-4

Table 9. Aerosol optical depths for a foggy day (201600508) as inferred from the PGS-100 sun photometer and calculated for each run of the PBPSO algorithm using a lognormal distribution.

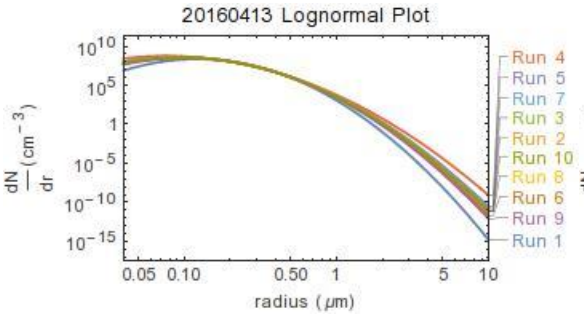

Figure 7.1

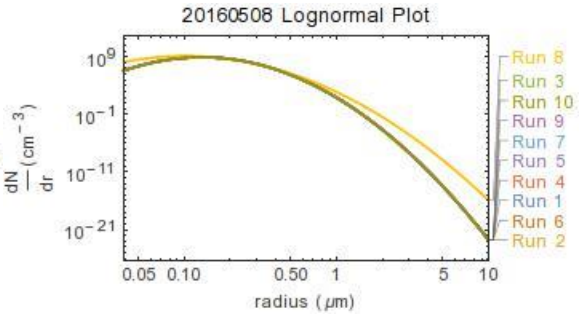

Figure 7.2

Figure 7. For the given days 20160413 in Figure 7.1 and 20160508 in Figure 7.2, all parameters are allowed to vary and there are multiple solutions that reproduce the aerosol optical depths. The solutions for the lognormal distribution are almost perfect and extremely uniform.

The lognormal results (**Figures 7.1-7.2**) show that once again the $\bar{\chi}^2$ values are almost identical; all being equal or off by 10^{-4} of each other for the clear day and 10^{-5} for the foggy day. Therefore all solutions are used to find the averages and SDOMs of the parameter values. Results for N_0, α, and β are shown in **Table 10.1** for the clear day (20160413) and **Table 10.2** for the foggy day (20160508).

	N_0 (cm^{-3})	α (μm)	β
Mean	4.98×10^{7}	0.122	0.211
SDOM	2.7×10^{6}	0.003	0.004

Table 10.1. Lognormal results 20160413 (clear day).

	N_0 (cm^{-3})	α (μm)	β
Mean	1.20×10^{8}	0.1453	0.159
SDOM	6×10^{6}	0.0026	0.003

Table 10.2. Lognormal results 20160508 (foggy day).

In general, the foggy day has smaller $\bar{\chi}^2$ values indicating that the PBPSO algorithm is better able to match AODs when there is a larger aerosol load in the atmosphere. **Tables 5** and **8** indicate that a bimodal distribution best describes the aerosol size distribution on the clear day (20160413) since the $\bar{\chi}^2$ values are lower in **Table 8**. **Tables 6** and **9** indicate that a lognormal distribution best describes the aerosol size distribution on the foggy day (20160508).

CONCLUSIONS

A PBPSO algorithm was used to analyze a synthetic set of atmospheric aerosol data in order to determine the reproducibility of results for four types of assumed size distributions. For Junge and lognormal distributions, there was little to no variation in the results from multiple runs of the algorithm, but the gamma and bimodal distributions the results indicated a large variability in size distribution parameters that would yield the same set of optical depths at 13 wavelengths. Running additional tests holding the α and/or β parameter(s) constant gave an indication of the sources of this variability. Inherent to the gamma distribution function, the β parameter directly relates to the rate of particle number increase as the radius increases to the peak radius $r_0 = \beta/\alpha$, while the α parameter determines the rate of particle number decrease for radii greater than the peak radius. Furthermore, the distribution is not symmetric in log space about r_0. Hence, the retrieved size distributions are more sensitive to variations in β because the contribution to AODs from small radii particles (where Q_{ext} has a maximum) becomes highly variable leading to large variations in N_0 and α. When retrieving bimodal distributions, the relatively flat Mie extinction profiles at larger radii for visible wavelengths of light make it difficult for the algorithm to pick out α_1, leading multiple valid solutions in which the relative contributions to the AODs from the small radius mode and the large radius mode are different. It should be noted that when α_0, α_1, β_0 and β_1 are all held constant, N_0 and N_1 values are exactly reproduced which is consistent with the findings in Yuan *et al.* 2011. In order to retrieve size distributions from AODs using PBPSO, multiple retrievals are performed and then a subset of solutions based on the criterion $\bar{\chi}^2 \leq 100\bar{\chi}^2_{min}$ is chosen to find averages and SDOMs of size distribution parameters.

Generally when analyzing atmospheric aerosols, a bimodal or lognormal distribution is preferred. When using the bimodal distribution for inferred AODs from clear and foggy days, the uncertainty in the radius for the small aerosols is lower than for the large aerosols (**Tables 7.1** and **7.2**). In addition, the foggy day has smaller $\bar{\chi}^2$ values indicating that the PBPSO algorithm is better able to match AODs when there is a larger aerosol load in the atmosphere. To choose whether a bimodal or lognormal distribution best characterizes the atmospheric aerosols, both were used with the PBPSO algorithm and the distribution with the lowest $\bar{\chi}^2$ was identified. A bimodal distribution best describes the aerosols on the clear day (20160413) and a lognormal distribution best describes the aerosols on the foggy day (20160508).

ACKNOWLEDGEMENTS

The author thanks the College of Saint Benedict/Saint John's University for excellent work facilities and general support of undergraduate research.

REFERENCES

1. Tao, W.-K., Chen, J.-P., Li, Z., Wang, C., and Zhang, C. (2012) Impact of Aerosols on Convective Clouds and Precipitation, *Rev. Geophys.* 50, RG2001, doi:10.1029/2011RG000369.
2. Schwartz, J. and Neas, L. M. (2000) Fine Particles Are More Strongly Associated than Coarse Particles with Acute Respiratory Health Effects in Schoolchildren, *Epidemiology* 11(1), 6–10.
3. Fleagle, R. G. and Businger, J. A. (1980) *An Introduction to Atmospheric Physics*, 2nd ed., Academic Press, Orlando.
4. U.S. Department of Energy (DOE)/National Renewable Energy Laboratory (NREL)/Alliance for Sustainable Energy, LLC (ALLIANCE). "Solar Spectra: Standard Air Mass Zero." E490_00a_AM0.xls spreadsheet downloaded 9 May 2015. http://rredc.nrel.gov/solar/spectra/am0/ASTM2000.html

5. King, M. D., Byrne, D. M., Herman, B. M., Reagan, J. A. (1978) Aerosol size distributions obtained by inversion of spectral optical depth measurements, *J Atmos Sci* 35, 2153–2167.
6. Kocifaj, M. and Gueymard, C. (2012) Aerosol size distribution retrievals from sunphotometer measurements: Theoretical evaluation of errors due to circumsolar and related effects, *Atmos Env* 51, 131–139.
7. He, Z., Qi, H., Lew, Z., Ruan, L., Tan, H., and Luo, K. (2016) Application of the LSQR algorithm in non-parametric estimation of aerosol size distribution, *Opt Comms* 366, 154–162.
8. Yamamoto, G. and Tanaka, M. (1969) Determination of aerosol size distribution from spectral attenuation measurements, *App Opt* 8(2), 447–453.
9. Grassl, H. (1971) Determination of aerosol size distributions from spectral attenuation measurements, *App Opt* 10(11), 2534–2538.
10. Wanf, Y., Fan, S., and Feng, X. (2007) Retrieval of the aerosol particle size distribution function by incorporating a priori information, *J Aer Sci* 38, 885–901.
11. Wang, Y. (2008) An efficient gradient method for maximum entropy regularizing retrieval of atmospheric aerosol particle size distribution function, *J Aer Sci* 39(4), 305–322.
12. Garey, M. R., and Johnson, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*, San Francisco, W.H. Freeman.
13. Eberhart, R. C. and Shi, Y. H. (2001) Particle swarm optimization: Developments, applications and resources, Proceedings of the 2001 Congress on Evolutionary Computation, Vols 1 and 2, 81–86 IEEE, New York.
14. Poli, R., Kennedy, J., and Blackwell, T. (2007) Particle swarm optimization: An overview, Swarm Intell 1, 33–57.
15. Settles, M. (2005) An Introduction to Particle Swarm Optimization. Research Gate. Web. 25 May 2016, https://www.researchgate.net/publication/242463151_An_Introduction_to_Particle_Swarm_Optimization.
16. Trelea, I. C., (2003) The particle swarm optimization algorithm: convergence analysis and parameter selection, *Inf Process Let* 85, 317–325.
17. Yuan, Y., Yi, H.-L., Shuai, Y., Wang, F.-Q., and Tan, H.-P. (2010) Inverse Problem for Particle Size Distributions of Atmospheric Aerosols Using Stochastic Particle Swarm Optimization, *JQSRT* 111(14), 2106–2114.
18. Yuan, Y., Yi, H.-L., Shuai, Y., Liu, B., and Tan, H.-P. (2011) Inverse Problem for Aerosol Particle Size Distribution Using SPSO Associated with Multi-lognormal Distribution Model, *Atmos Env* 45(28), 4892–4897.
19. Mao, J. and Li, J. (2015) Retrieval of Particle Size Distribution from Aerosol Optical Thickness Using an Improved Particle Swarm Optimization Algorithm, *Opt Rev* 22(5), 809–818.
20. Cimel Electronique. "Sun sky lunar multiband photometer CE318T." www.cimel.fr, downloaded 30 March 2017.
21. Che, H., Uchiyama, A., Yamazaki, A., Chen, H., Goloub, P., and Zhang, X. (2008) Intercomparison between aerosol optical properties by a PREDE skyradiometer and a CIMEL sunphotometer over Beijing, China, *Atmos Chem Phys* 8, 3199–3214.
22. GAW Report No. 227 (2016), *WMO/GAW Aerosol Measurement Procedures, Guidelines and Recommendations, 2nd Edition*, World Meteorological Organization, Switzerland.
23. Bohren, C. F. and Huffman, D. R. (2007) Appendix A: Homogeneous Spheres in *Absorption and Scattering of Light by Small Particles*, Wiley-VCH Verlag GmbH.

ABOUT THE STUDENT AUTHOR

Kaitlin DuPaul graduated in May 2017 with a B.S. in Mathematics. After spending her summer researching atmospheric aerosols, she hopes that this exceptional research experience will give her an advantage as she plans to apply to graduate programs in Material Science this fall. In the long term she hopes to develop new ways to implement renewable energy resources.

PRESS SUMMARY

The parameter based particle swarm optimization (PBPSO) algorithm is introduced in order to retrieve aerosol size distributions from aerosol optical depth calculations without holding any size distribution parameters constant. PBSPO is tested against standard sets of aerosol optical depths to determine the reproducibility of results for Junge, gamma, lognormal, and bimodal size distributions. The PBPSO algorithm is applied to aerosol optical depth calculated from irradiance measurements from a Kipp-Zonen PGS-100 solar spectrometer. Results indicate that a lognormal size distribution best describes the aerosols on a foggy day (May 8, 2016) and a bimodal distribution best describes the aerosols on a clear day (April 13, 2016).

Strategy Abandonment Effects in Cued Recall

Stephanie A. Robinson[*a], Amy A. Overman[a], & Joseph D.W. Stephens[b]*

[a]Department of Psychology, Elon University, NC
[b]Department of Psychology, North Carolina A&T State University, NC

Student: srobins1@brandeis.edu
Mentor: aoverman@elon.edu

ABSTRACT
Decades of research have investigated the effects of encoding strategies in the formation of associations in memory. Despite this, it is not known whether or how changes in the use of strategies within a brief time span may affect memory. For example, what is the effect on memory of abandoning a recent strategy or switching to a different strategy? The present study systematically varied the strategies used by participants in two closely-spaced associative memory tasks. Results indicated that intentional abandonment of a verbal (sentence-generation) strategy had disproportionately negative consequences on memory for semantically unrelated word pairs. The findings suggest that memory encoding is affected by differences in strategy use across recent memory tasks, and have implications for effective use of memory strategies in practical settings.

KEYWORDS
Cued Recall; Encoding Strategies; Inhibition

INTRODUCTION

The intentional use of appropriate encoding strategies to improve memory for associations between items has been well-documented[1]. Much of the initial work on memory encoding strategies compared the relative effectiveness of specific strategies, such as mental imagery and sentence formation[2,3]. Other work has focused on examining factors that influence individual differences in strategy selection and use[4,5,6,7], as well as the effects of strategy training, often in special populations[8,9,10,11].

The existing literature has provided important insights into the effectiveness of memory encoding strategies within the context of individual memory tasks. However, memory tasks in everyday life often do not exist in isolation. For example, one may need to perform more than one memorization task within a short time frame, such as studying for an exam or orienting to a new work environment. There is currently little research that has investigated strategy use across more than one closely-spaced memory task. Thus, it is not known whether abandoning an encoding strategy, switching between encoding strategies, or practicing use of an encoding strategy confers any benefits or costs across associative memory tasks.

Successive memory tasks with the same or different strategies may be approached as a type of task-switching, in which the particular encoding strategy instructions constitute the participant's task. Therefore, the task-switching literature provides a framework for understanding the ways in which cognitive processing at a given moment is influenced by recent processing[12]. For example, performance on a current task can be impaired by inhibition of a prior task with similar processing demands[13]. These findings in the task-switching literature are relevant to the question of whether a current memory encoding task is influenced by an encoding task that preceded it.

With regard to switching of memory strategies or abandonment of a memory strategy, the previously-used strategy must be inhibited during the use of the current strategy. The inhibition of prior encoding strategies could impair current encoding if some of the same cognitive processes are contributing to both tasks. For example, if an explicitly verbal encoding strategy is being inhibited, the inhibition of the strategy may also inhibit verbal encoding processes that are not specific to that strategy. This could result in performance that is worse than if the prior strategy had not been employed at all (and therefore does not need to be inhibited). In addition to the possible direct inhibition of relevant encoding processes, the attentional and cognitive control resources involved in the inhibition may reduce the general processing resources available for encoding.

There may also be immediate effects of repeating the use of a particular strategy. For example, when memorizing two sets of associates within a single study session, it might be beneficial to use the same strategy for both, to take advantage of practice effects. For example, extensive training on a memory strategy has been shown to improve use of the strategy across multiple sessions[8]. On the other hand, immediate strategy repetition within an experimental session might lead to detrimental effects such as interference or cognitive fatigue[14]. If associative memory benefits from using a strategy due to familiarity with an encoding

technique, then the benefit would be expected to increase with repeated use (practice/training) of the encoding strategy. Thus, it remains unclear how memory performance is affected by strategy practice and strategy switching. This study aimed to compare the effects, if any, of strategy addition or abandonment, strategy switching, and strategy repetition on cued recall.

METHODS AND PROCEDURES

Participants

One hundred eighty-nine participants were recruited from university psychology courses (mean age = 19.12 years; range = 18-23 years; 155 females, 34 males). They either received course credit or were entered into a lottery to win a gift card to a local store. All participants were native English speakers and reported no history of neurological or psychiatric disorders. Prior to the experiment, all participants provided written informed consent to participate using consent forms approved by the Institutional Review Board of the university. After study participation was complete, all participants were given a verbal and written debriefing.

Materials

The experimental stimuli consisted of 128 words selected from the University of South Florida word association database[15]. The words were selected to produce a heterogeneous assortment in terms of length, word frequency, part of speech, concreteness, and imageability, with the constraint that half of the words were selected to produce semantically related pairs according to the association norms (mean relatedness = .77; range = .65-.94), and the other half were selected to ensure no semantic relationship (mean relatedness = 0.00; range = 0.00). The pairs were separated into two study lists for use in the cued-recall tasks. The lists were counterbalanced so that each pair was equally likely to be used in the first or second study-test phase for each participant, in order to prevent any systematic variation in the individual word properties across conditions[16].

Design

Three encoding strategies were selected for this experiment because of prior research that has established that the three basic categories of self-initiated strategies are visual, semantic, and shallow[4]. The use of a visual word-learning task allowed us to employ each of these strategies: visual, semantic, no strategy. In order to examine any effects of changing between encoding strategies, we used each possible pairing of three encoding strategy instructions: imagery, sentence, and no strategy, resulting in a 9 (encoding strategy order; between-subjects) X 2 (semantic relatedness; within-subjects) mixed model design.

For all conditions, care was taken to thoroughly discuss with the participant what was expected. Instructions were first presented on the computer screen for the participant to read. Then, participants were asked to explain the task back to the experimenter and the experimenter emphasized task expectations and corrected any errors in understanding. If there were errors, participants were asked again to explain the instructions to the experimenter. Then, prior to using a prescribed strategy, participants completed a practice trial of the task. After the practice trial they described what they were doing during the practice trials to the experimenter. If participants were not correctly employing a strategy after the first practice trial or reported using a strategy when they were instructed not to do so, the instructions were re-explained and discussed with the participant and then the participant attempted a second practice trial. On the second practice trial, no participants reported using an incorrect strategy or using a strategy in the no strategy task. Participants were randomly assigned to one of nine possible strategy sequences resulting from the combination of the three sets of strategy instructions across two encoding phases (See **Table 1**). There were an average of 21 participants per condition, although the exact number in each condition varied slightly due to random assignment.

	Second List		
First List	**No Strategy**	**Sentence**	**Image**
No Strategy	*Repetition* (n=19)	*Addition* (n=21)	*Addition* (n=22)
Sentence	*Abandonment* (n=21)	*Repetition* (n=22)	*Switching* (n=20)
Image	*Abandonment* (n=21)	*Switching* (n=20)	*Repetition* (n=23)

Table 1. Strategy Sequences Across Study Lists
Note. Each participant was given one of nine possible sequences of encoding strategy instructions across two cued-recall tasks. Effects of strategy addition, strategy abandonment, strategy switching, and strategy repetition were analysed in terms of the performance differences observed between the two cued-recall tasks in each of the nine sequences.

For the imagery encoding strategy, participants were asked to visualize the words in a scene: for example, the pair SALT-PEPPER (semantically related) might be visualized as salt and pepper shakers on a table; the pair TOUCAN-FORK (semantically unrelated) might be visualized as a toucan using a fork to eat. For the sentence formation strategy, participants were asked to think of a

sentence using both words in the word pair. Using the same word pairs above, for the semantically-related pair, the participant could form the sentence "Could you please put the salt and pepper on the dinner table?", and for the semantically-unrelated pair, the participant could form the sentence "The toucan stole my fork during our picnic". For no strategy, participants were instructed not to use any strategy and examples were given of strategies that should not be used (*e.g.,* sentence formation, imagery, rote repetition).

Procedure

The experiment was conducted using E-Prime software (Psychology Software Tools, Sharpsburg, PA). The stimuli were presented side-by-side in black Courier New 16-point font on a white background. Participants were seated at a computer and told that they would be seeing a list of word pairs and would later be asked to remember what they saw. Each study list contained 32 pairs of words. Sixteen related pairs and 16 unrelated pairs were randomly assigned to each study list. Half were pairs of words that were semantically related and half were pairs of words that were arbitrarily assigned. Pairs were presented for seven seconds. At the end of the list, participants completed a three-minute simple addition task before proceeding to the cued recall task. For the cued recall task, one of the words from each pair in the list was presented on a sheet of paper and the participant was instructed to fill in the corresponding word in the pair. All the pairs from the study list were tested. Thus, there were 16 cue words from related pairs and 16 cue words from unrelated pairs presented on the cued recall test. After participating in the first study-test phase, the participants waited for two minutes and then proceeded to the second study-test phase, which followed the same format: written and verbal instructions, study, distractor task, and then test.

RESULTS

Of particular interest for the current study was whether the use of a strategy (or no strategy) on a second cued recall test would be influenced by the use of a strategy (or no strategy) on a prior cued recall test. Thus, the focus was not on the relative efficacies of the encoding strategies but rather on the differences in performance observed with changes in strategy use across successive tasks. We considered three possibilities for such influences: 1) whether performance changed when the same strategy was used twice (*strategy repetition*); 2) whether the difference between a strategy and no strategy depended on whether the no-strategy task came first (*strategy addition*) versus second (*strategy abandonment*); and 3) whether the difference between the two strategies depended on the order of the two strategies (*strategy switching*). Each of these possibilities was examined by analysing the groups of participants relevant to each comparison. Because each group of participants completed two tests, we particularly focused the analyses on the *change* in performance from one test to another, thereby controlling for variations in overall performance across individual participants and groups.

Strategy repetition

For the question of whether the effect of any strategy changed with repetition, we compared the three groups that had the same strategy (or lack thereof) twice – i.e., No Strategy (N-N group), Sentence (S-S group), and Imagery (I-I group). We analysed performance in a 2 (test) X 2 (relatedness) X 3 (strategy) ANOVA. When the same strategy was used twice, there was no change in performance between the first and second tests, $F(1,61) = 1.74$, $MSE = .028$, $p = .19$. There was also no difference between strategy types, $F(2,61) = .934$, $MSE = .098$, $p = .40$, and the were no interactions (all $p > .25$). There was a main effect of relatedness, $F(1,61) = 238.6$, $MSE = 6.18$, $p < .001$, $\eta_p^2 = .80$, such that cued recall was better for target words that were semantically related to their cues (see below for an overall analysis of the effect of relatedness across all strategy sequences). Thus, performance did not significantly change with strategy repetition.

Strategy addition versus abandonment

For the question of whether the effect of a strategy versus no strategy depended on whether the no-strategy task occurred first or second, we compared the group that had No Strategy in the first study list, followed by Imagery in the second study list (i.e., the N-I group), with the group that had Imagery in the first study list, then No Strategy in the second study list (i.e., the I-N group). We also compared the group that had No Strategy in the first study list, followed by Sentence in the second study list (i.e., the N-S group), with the group that had Sentence in the first study list, followed by No Strategy in the second study list (i.e., the S-N group).

The comparison between the N-I and I-N groups is displayed in Figure 1. The difference in performance between the two cued-recall tests was computed for each participant as p(Correct|Imagery) – p(Correct|No Strategy). Thus, the data represent the relative benefit of Imagery across the two strategy orders. These data were analysed in a 2 (relatedness) X 2 (order) ANOVA with relatedness as a within-subjects factor and order as a between-subjects factor. The effect of relatedness was significant, $F(1,41) = 20.6$, $MSE = .394$, $p < .001$, $\eta_p^2 = .34$, indicating that the Imagery strategy benefitted recall of target words that were semantically unrelated to their cue words more than it benefitted cued recall of semantically-related targets. There was no significant effect of order, $F(1,41) = 1.15$, $MSE = .063$, $p = .29$, and no significant interaction, $F(1,41) = 3.05$, $MSE = .058$, $p = .09$, suggesting that the benefit to recall performance caused by adding the Imagery strategy was similar to the harm caused by abandoning it.

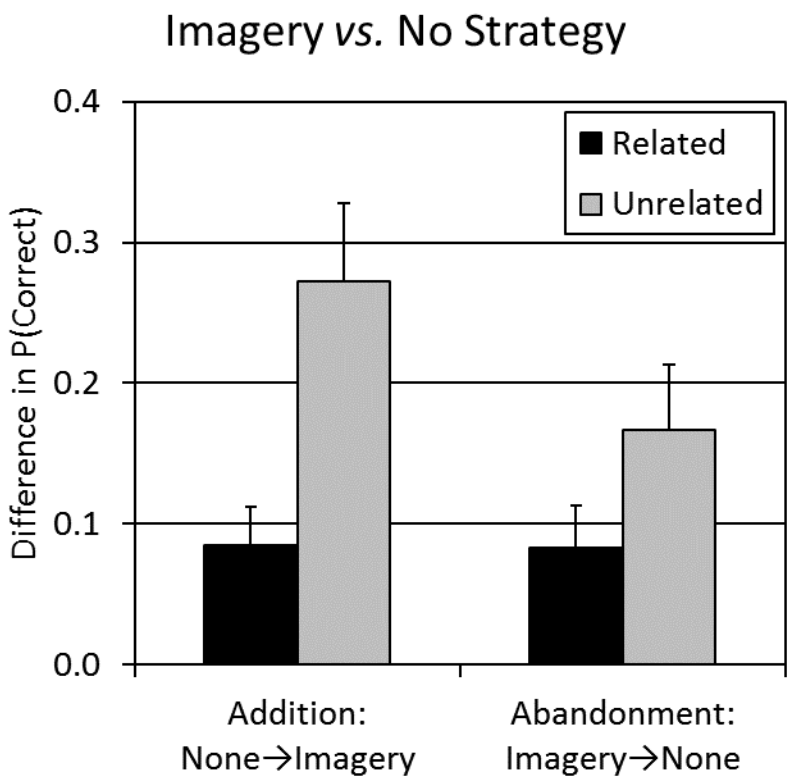

Figure 1. Cued-recall performance with imagery versus no strategy, across strategy addition and abandonment conditions. Difference scores were computed as proportion correct with imagery minus proportion correct with no strategy, for each participant. Error bars represent standard error of the mean.

The comparison between the N-S and S-N groups is displayed in Figure 2. As with N-I and I-N, the difference in performance between strategy and no strategy was computed as the relative benefit of Sentence, p(Correct|Sentence) – p(Correct|No Strategy). The data were analysed in a 2 (relatedness) X 2 (order) ANOVA with relatedness as a within-subjects factor and order as a between-subjects factor. The main effect of relatedness was not significant, $F(1,40) = 2.03$, $MSE = .086$, $p = .16$, and there was no significant effect of order, $F(1,40) = .03$, $MSE = .002$, $p = .87$. However, there was a significant interaction of relatedness X order, $F(1,40) = 11.2$, $MSE = .474$, $p = .002$, $\eta_p^2 = .22$. Inspection of the data in Figure 2 indicates that the addition of the Sentence strategy mainly caused a benefit in cued recall of target words that were semantically related to their cue words, whereas abandonment of the Sentence strategy mainly caused a cost to the cued recall of semantically-unrelated words. These observations were examined in detail using one-sample *t*-tests comparing the means in each condition to zero. For the N-S group, the change in performance was significantly different than zero for related words, $t(20) = 3.85$, $p = .001$, $d = .84$, and unrelated words, $t(20) = 2.09$, $p = .049$, $d = .46$. For the S-N group, the change in performance between strategies was significant only for the unrelated words, $t(20) = 4.33$, $p < .001$, $d = .94$, and not for related words, $t(20) = .36$, $p = .73$.

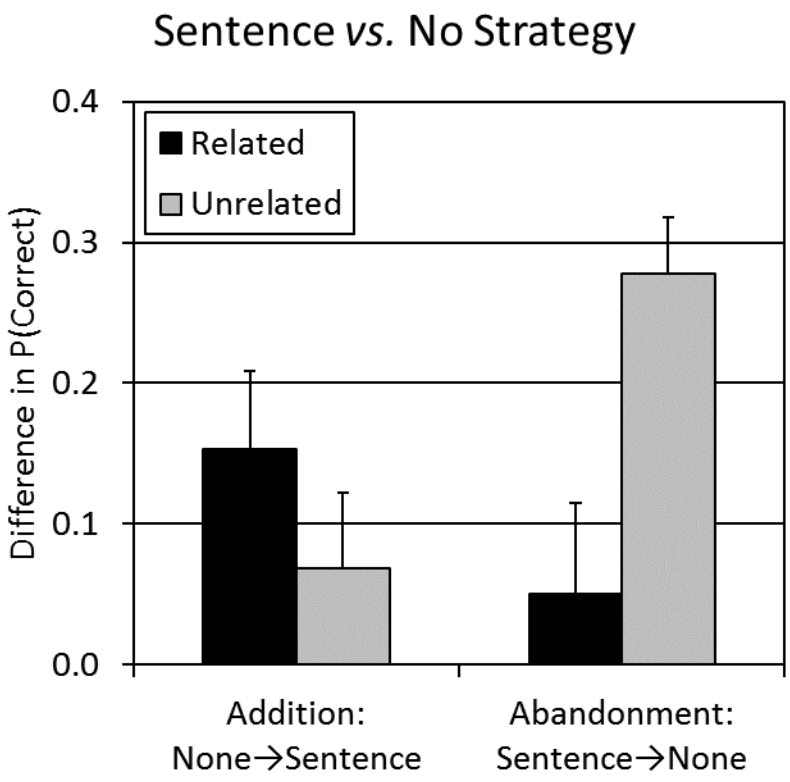

Figure 2. Cued-recall performance with sentence generation versus no strategy, across strategy addition and abandonment conditions. Difference scores were computed as proportion correct with sentence generation minus proportion correct with no strategy, for each participant. Error bars represent standard error of the mean.

Strategy switching

For the question of whether the difference in performance between the two strategies depended on strategy order, we compared the group that had the Sentence strategy first, followed by Imagery (S-I group), with the group that had Imagery first, then Sentence (I-S group). Similar to the analyses of strategy addition versus abandonment, we examined the change in performance between strategies as the benefit of Imagery over Sentence, p(Correct|Imagery) – p(Correct|Sentence). We then analysed these data in a 2 (relatedness) X 2 (order) ANOVA. There was no main effect of relatedness, $F(1,38) = 3.48$, $MSE = .127$, $p = .07$, and there was no main effect of order, $F(1,38) = .07$, $MSE = .002$, $p = .79$. There was also no interaction of relatedness X order, $F(1,38) = .03$, $MSE = .001$, $p = .86$. Thus, the order in which the two strategies were used did not affect cued-recall performance.

Overall effect of relatedness

In addition to the questions of interest analysed above, the results were analysed with respect to the overall effect of relatedness. Across both tests, target words that were semantically related to their cue words were recalled more often than those that were unrelated. A 2 (test list) X 2 (relatedness) ANOVA on number of words correctly recalled found a significant effect of relatedness, $F(1,188) = 704.3$, $MSE = 18.2$, $p < .001$, $\eta_p^2 = .79$, no effect of list, $F(1,188) = .573$, $MSE = .019$, $p = .45$, and no interaction, $F(1,188) = 1.56$, $MSE = .032$, $p = .24$. This effect of relatedness is in line with prior studies (e.g., Jenkins & Russell, 1952).

Overall effects of strategy type

Although the main questions of interest for the current study had to do with differences in performance between tests according to strategy use, we also compared the strategies to each other overall, for comparison to prior work on study strategies. Because different participants received different strategy sequences, the overall effectiveness of each strategy was examined separately within each test list. Thus, strategy was analysed as a between-subjects factor in two ANOVAs, one for each test list. For the first test list, the effect of strategy was significant, $F(2,186) = 14.8$, $MSE = .380$, $p < .001$, $\eta_p^2 = .14$. Pairwise Bonferroni-corrected post-hoc comparisons indicated that Imagery was better than no strategy, $p < .001$, Sentence was better than no strategy, $p < .001$, but Imagery and Sentence did not differ, $p = 1.0$. For the second test list, the effect of strategy was significant, $F(2,186) = 11.4$,

MSE = .294, $p < .001$, $\eta_p^2 = .11$. Pairwise Bonferroni-corrected post-hoc comparisons indicated that Imagery was better than no strategy, $p < .001$, and Imagery was better than Sentence, $p = .037$. Due to the correction for multiple comparisons, sentence was not significantly better than no strategy, $p = .076$. The overall greater effectiveness of Imagery was in line with the well-established literature on memory strategies[2,3]. The fact that the statistical differences between strategies were not equivalent between the two tests may be attributable to the order effects of N-S versus S-N described above.

DISCUSSION & CONCLUSIONS

The purpose of the current study was to examine the effects, if any, of changes in memory strategy use across two cued-recall tasks performed within a single experimental session. Little prior research has investigated how the immediate addition, abandonment, switching, or repetition of memory strategies affects the usefulness of those strategies in boosting memory performance.

Interestingly, we found that the effect of abandoning the sentence-formation strategy did not mirror the effect of adding the sentence-formation strategy. Unrelated words were disproportionately affected by abandonment of the sentence strategy, in comparison to how they were affected by addition of the sentence strategy. This finding was in contrast to the results for the mental imagery strategy. Abandonment of the imagery strategy on the second task had nearly the exact opposite effect as adding the strategy on the second task after using no strategy on the first.

Although it may seem counterintuitive that abandonment of a sentence formation strategy would have different effects than abandonment of an imagery strategy, the results may be interpreted in light of the relationship between the task and the strategy. The memory tasks in this study were verbal in nature (i.e., encoding of word pairs). Participants who used a sentence strategy on the first task, and were then instructed to stop using the strategy, likely did so by inhibiting some of their verbal processing during the second task. Thus, participants who abandoned the sentence strategy had fewer verbal processing resources available for encoding word pairs during the subsequent no-strategy task. In contrast, participants who used an imagery strategy on the first task, and then were instructed to stop using the strategy, did not need to inhibit any verbal processing because the first strategy they used did not require any verbal processing.

It may be argued that we do not know exactly *what* participants were doing in the no-strategy task, and thus strategy abandonment might be considered another form of strategy switching, in which participants simply switched to idiosyncratic strategies of their own choosing. Nonetheless, the important result remains that this particular form of switching led to a different result in the case of Sentence to No Strategy switching than in the other forms of switching considered here. Based on this finding, we argue that the instruction to abandon the sentence strategy, without providing an alternative strategy such as imagery, leads to a subsequent inhibition of verbal processing in whatever idiosyncratic encoding processes are subsequently employed in the next task. An additional possibility to consider is whether encoding strategy instructions might also affect retrieval processing during the subsequent cued recall task. For example, participants who are instructed to abandon a sentence strategy might also inhibit verbal processes used in recollection of the target word at test. Participants who are abandoning an imagery strategy might not inhibit such retrieval processes as they seem unrelated to the strategy being abandoned.

Other than the effects of strategy abandonment, our study indicated that the effect of a strategy on the second cued-recall task was independent of what strategy was used on the previous cued-recall task. Immediate repetition did not lead to significant improvements, and the effect of changing strategies did not depend on the order in which the strategies were used. It is important to note that, in the conditions that involved immediate repetition of a strategy, we were specifically interested in possible practice effects, rather than in training effects. Thus, participants were not instructed or encouraged to become more proficient at the use of the strategy as in some prior studies[8].

This study is an important first step toward understanding how the abandonment of prior strategies may lead to inhibitory effects on subsequent strategies, which result in a detriment to memory performance. These results also have implications for real-world scenarios in which people engage in more than one study task within a brief period of time, such as studying for examinations. Further research should build on this study both in terms of examining the cognitive mechanisms involved in strategy abandonment as well as how to optimize strategy use in learning contexts.

ACKNOWLEDGMENTS

Joseph Stephens was supported by Air Force Research Laboratory and OSD under agreement number FA8750-15-2-0116. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory and OSD or the U.S. Government.

REFERENCES

1. Richardson, J. T. E. (1998). The availability and effectiveness of reported mediators in associative learning: A historical review and an experimental investigation. *Psychonomic Bulletin & Review,* 5, 597–614.
2. Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review,* 76, 241-263.
3. Bower, G. H., & Winzenz, D. (1970). Comparison of associative learning strategies. *Psychonomic Science,* 20, 119-120.
4. Dunlosky, J., & Hertzog, C. (1998). Aging and deficits in associative memory: what is the role of strategy production? *Psychology & Aging,* 13, 597–607.
5. Dunlosky, J., & Hertzog, C. (2001). Measuring strategy production during associative learning: The relative utility of concurrent versus retrospective reports. *Memory & Cognition,* 29, 247-253.
6. Kirchhoff, B. A. (2009). Individual differences in episodic memory: The role of self-initiated encoding strategies. *The Neuroscientist,* 15, 166-179.
7. Kuhlmann, B. G., & Touron, D. R. (2012). Mediator-based encoding strategies in source monitoring in young and older adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 38, 1352.
8. Brehmer, Y., Shing, Y. L., Heekeren, H. R., Lindenberger, U., & Bäckman, L. (2016). Training-induced changes in subsequent-memory effects: No major differences among children, younger adults, and older adults. *NeuroImage*, 131, 214-225.
9. Hampstead, B. M., Stringer, A. Y., Stilla, R. F., Giddens, M., & Sathian, K. (2012). Mnemonic strategy training partially restores hippocampal activity in patients with mild cognitive impairment. *Hippocampus,* 22, 1652-1658.
10. Patrick, J., Morgan, P. L., Smy, V., Tiley, L., Seeby, H., Patrick, T., & Evans, J. (2015). The influence of training and experience on memory strategy. *Memory & Cognition,* 43, 775-787.
11. Rebok, G. W., Carlson, M. C., & Langbaum, J. B. (2007). Training and maintaining memory abilities in healthy older adults: traditional and novel approaches. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences,* 62, 53-61.
12. Kiesel, A., Steinhauser, M., Wendt, M., Falkenstein, M., Jost, K., Philipp, A. M., & Koch, I. (2010). Control and interference in task switching—A review. *Psychological Bulletin,* 136, 849-874.
13. Philipp, A. M., & Koch, I. (2006). Task inhibition and task repetition in task switching. *European Journal of Cognitive Psychology,* 18, 624-639.
14. Van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: Effects on perseveration and planning. *Acta Psychologica,* 113, 45-65.
15. Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers,* 36, 402-407.
16. Madan, C., Glaholt, M., & Caplan, J. (2010). The influence of item properties on association-memory. *Journal of Memory and Language,* 63, 46–63.
17. Jenkins, J. J., & Russell, W. A. (1952). Associative clustering during recall. *Journal of Abnormal and Social Psychology,* 47, 818-821.

ABOUT THE STUDENT AUTHOR

Stephanie Robinson graduated from Elon University with a B.A. in Psychology. Stephanie completed a Master's in Psychology at the Catholic University of America where she continued to explore memory and learning strategies. She is currently completing a Ph.D. in Psychology at Brandeis University in Waltham, MA with a focus on behavioural and psychosocial differences in cognition across the lifespan and investigating intervention methods to ameliorate age-related cognitive decline.

PRESS SUMMARY

For decades, researchers have investigated strategies to facilitate memory; however, it not yet known if continuous use of the same strategy is more useful than using a novel strategy. That is, if an individual was studying for two exams, would it be helpful to use the same strategies to memorize the material (strategy repetition; i.e., practice), or is it more helpful to use a new strategy for new material (strategy abandonment)? This study demonstrated that the effects of switching strategies was dependent qualified by the relatedness of the material. This study provides an important first step to understanding how switching study strategies could negatively impact memory.

Print Edition ISSN 1536-4585
Online Edition ISSN 2375-8732