HETEROZYGOSITY OF A WOOD TURTLE POPULATION IN CENTRAL NY

Kyle Pursel
Department of Biological Sciences

This presentation will report on the work and results of a genetic study on the wood turtle, *Glyptemys insculpta*, at two sites in Eastern Oswego County, NY for the Student/Faculty Challenge Grant. This study looked at tandemly repeating segments of nuclear DNA, referred to as microsatellites, to assess heterozygosity and relatedness of marked turtles at both sites. This data is part of a larger project assessing the status of wood turtles at both sites as a part of my honors thesis.

I. Introduction

Wood turtles, *Glyptemys insculpta*, are riparian species of turtle native to New York State and much of the Northeastern US, Great Lakes, and Southeastern Canada (Ernst et al., 1994). Declines have been documented in populations throughout their range in recent decades (Daigle & Jutras, 2005; Garber & Burger 1995; Harding, 1991; Saumure & Bider, 1998). As a result of the these declines, many states and Canadian provinces are providing protective status to wood turtles by listing them as endangered, threatened, or a species of special concern (Bowen & Gillingham, 2004). In New York State, the wood turtle is a species of special concern (Al Breisch, personal communication).

In 2005, Dr. Peter Rosenbaum and Kyle Pursel initiated surveys to find and study wood turtles in eastern Oswego County in New York State to determine the natural history, ecology, and status of this species in this part of their range. In 2007, a genetic proponent was added to the study to assess the genetic variation of turtles found at the study sites and to aid in the assessment. A genetic component was added due to the valuable data genetics can provide about a population. It was determined that microsatellites, which measure tandemly repeating sequences of nuclear DNA and are generally considered to be good measures of heterozygosity, were chosen to calculate the genetic diversity of the population. Heterozygosity is the measure of heterozygotes in a population, and a heterozygote is an individual which has two separate alleles for a given gene. With microsatellites, an allele can be considered to be any variation in the number of tandem repeats (Frankham et al., 2004). For example, an individual that is heterozygous can have one allele for 30 repeats and another allele for 33 repeats, which a homozygous individual would have either two alleles with 30 or two alleles with 33 repeats. Aiding in the decision was also a study released in 2005 which calculated the heterozygosities of six populations

in Quebec, Canada, which provided the data and names of primers proven to work with wood turtles for microsatellite studies (Tessier et al., 2005).

II. Methods

Blood samples were taken from each turtle using a vacuum syringe. Samples of 1/10 to ½ cc were taken depending upon the size of the turtle and ease of blood taking. Samples were preserved in a clotting buffer and stored in a -40^{o}F freezer until use. Thawed samples were extracted using a PureGene DNA extraction kit and the corresponding manufacturer's protocol for clotted blood. The general process consists of lysing the blood cells of the clotted blood were and protein denaturation for subsequent protein removal using proteinase K and centrifugation. For each extraction, roughly 50µL of clotted blood was added to 550 µL of cell lysis solution and 3 µL of proteinase K solution in a sterile 1.5mL labeled centrifuge tube. This was inverted 25 times and incubated at 55^{o}C overnight. The sample was then cooled to room temperature and 200 µL of protein precipitation solution was added. The solution was then vortexed at high speed for 20 seconds and then placed on ice for 5 minutes before being centrifuged on the highest setting for approximately 5 minutes. The resulting supernatant containing the DNA was then poured into a new sterile 1.5mL labeled centrifuge tube containing 600 µL of 100% isopropanol and was mixed by inversion 50 times. The sample was then centrifuged again on high for approximately 1.5 minutes. The new supernatant was poured out and the tube allowed to dry while inverted. Then, 600 µL of 70% ethanol was added and then centrifuged for another 1.5 minutes. The supernatant was again drained and inverted for about 10 minutes to dry. Once dry, 20 µL of DNA hydration solution was added to the DNA pellet and incubated at 65^{o}C for approximately one hour. Once hydrated, the solution was stored at 4^{o}C until needed. DNA was then quantified using a fluorometer.

Once quantified, DNA extractions were diluted for PCR. Frozen PCR materials were allowed to thaw and then mixed to make a master solution that could be added to the DNA samples. Added materials include 10X PCR buffer, dNTP, $MgCl_2$, and Taq polymerase. During this time, the forward and reverse primers for one of the five microsatellite primers was added to the master solution. The appropriate amounts of master mix and DNA solution were added to a sterile labeled 0.2mL centrifuge tube. This was centrifuged briefly before being placed in a thermocycler. The sequence of the program used for the thermocycler was adapted from that used by Tessier et al. (2005) and goes as follows: 2 minutes at 95^{o}C for one cycle, 35 cycles of 94^{o}C for 45 seconds, 54^{o}C for 45 seconds, 72^{o}C for one minute.

Samples were then prepared and ran on a Beckman-Coulter CEQ 8000 Genetic Analyzer. Samples were placed in trays and ran in acrylamide gels through capillaries. Fragments would travel at different speeds through the gel in the capillaries and, when they reached a certain point, a laser would detect the fluorescent tag in the fragments and determine the base pair composition. Once complete, the data came out in graphs as a series of one or two peaks, depending on if that individual was heterozygous or homozygous for that loci. From these peaks, the amount of base pairs or nucleotide

compliments, was taken and analyzed using the freeware program PopGene to determine the heterozygosity, allele numbers and frequencies, and Hardy-Weinberg Equilibrium for both sites. The freeware program Structure was used to determine if both sites compromised one large population or separate populations.

III. Results

Three of the five loci tested yielded recordable results. The loci which worked were GmuB21, GmuD16, and GmuD93. Not all turtle blood samples worked, with twenty-three different individuals working with one to all of the three loci. Base pair data for loci GmuB21 was obtained for 14 individuals from Sloperville and 3 from Little Grindstone, 13 in Sloperville and 4 for Little Grindstone for loci GmuD16, and 11 for Sloperville and all samples for Little Grindstone with loci GmuD93.

Analysis using the program Structure could not differentiate individuals form either site into separate populations. Fig. 1 clearly shows the close relatedness of turtles from each site to one another, with each turtle consistently having approximately 50% likelihood of being assigned to one site or the other during analysis. This data is consistent with both sites being part of a single larger population.

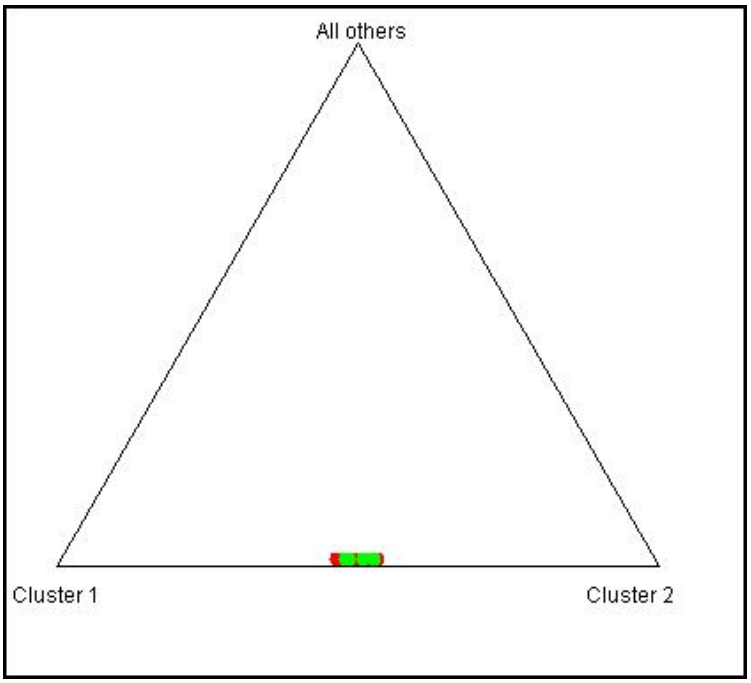


Fig. 1 Triangle Plot of the likelihood of turtle being in separate populations. Note that all points lie roughly in the center (50%) range of the plot. Red dots are individuals from Sloperville, green dots are individuals from Little Grindstone

Analysis using PopGene shows a total of 35 different alleles total from both sites together. Totals of 13, 9, and 13 different alleles were found for each loci (GmuB21, GmuD16, and GmuD93) respectively. Little Grindstone had fewer alleles per loci than Sloperville, but did contain 2 private alleles for GmuD93. Table 1 summarizes allele data for all loci and both site pooled together and separated. Sloperville consistently had larger numbers of alleles than Little Grindstone.

Table 1. Observed Allele Numbers per Loci

Locus	Diploid # of alleles	Observed allele total	Sloperville	LittleGrindstone
GmuB21	34	13	13	3
GmuD16	34	9	9	2
GmuD93	32	13	11	8
Mean	33	11.6667		
St. Dev		2.3094		

In all three loci, the observed heterozygosity was lower than the expected heterozygosity (Table 2). Tests of Hardy-Weinberg equilibrium within both sites pooled together as one population show that loci GmuB21 and GmuD93 were significantly out of equilibrium (B21 p=<0.001 Chi-Square=142.7; D93 p=0.004, Chi-Square=114.2), whereas loci GmuD16 was not significantly different from equilibrium (p=0.741, Chi-Square=30.2).

Table 2. Observed and Expected Heterozygosity of Both Sites

Locus	Diploid Sample Size	Observed Heterozygosity	Expected Heterozygosity
B21	34	0.5294	0.8663
D16	34	0.6471	0.7968
D93	32	0.6875	0.8851
Mean	33	0.6213	0.8494
St. Dev		0.0821	0.0465

Data for the individual loci and turtle were not obtained from the Quebec investigators (Tessier et al., 2005). A simple eye-ball comparison of data shows a similar trend amongst all populations (Table 3), in which all populations appear to have relatively high numbers of alleles and heterozygosity. However, thorough statistical analysis could not be conducted since only 3 of the 5 loci used in the Quebec studies yielded data for the NY population.

Table 3. Heterozygosity and Allele Comparisons Between Quebec & NY Populations

Site	Sample Size (turtles)	Observed Allele #	Observed Heterozygosity
Quebec FA*	46	50	0.561
Quebec SH*	39	46	0.673
Quebec SU*	8	40	0.886
Quebec MI*	19	45	0.804
Quebec TO*	6	36	0.767
Quebec DC*	18	54	0.837
Central NY	~23	35	0.621

*See Tessier et al., 2005. for in-depth site data and name meanings.

IV. Discussion

Turtles from both appear to be or recently were part of a larger population that has and continues to be fragmented by human development of the land. Although 2 out of the 5 loci failed to yield reliable data, a relatively large number of alleles were found for Sloperville and the overall population. This data indicates that turtles from Sloperville and the overall population are possibly genetically stable. However, Little Grindstone had lower allele numbers for 2 of the 3 loci. This may be an indication that Little Grindstone, through recent isolation and human impacts, may be facing an ecological and genetic bottleneck, and might be losing some of the genetic diversity that is present at other sites. This is consistent with the turtle collection data, which only yielded 5 turtles over 3 years. Since one loci, GmuD93, did yield a relatively decent number of alleles, it is possible that the addition of other loci may show that Little Grindstone turtles are more genetically variable than the current data suggests.

The heterozygosity and Hardy-Weinberg data suggests that the populations have fewer heterozygotes than expected for the number of loci present. However, Hardy-Weinberg is based off of many assumptions of nonrandom mating, no selection pressures (for that gene), no mutation, no migration, no genetic drift, and large populations. It is possible that sample sizes are not large enough to provide enough data and that the population is within equilibrium. However, it is also possible that there have been declines within the population, as appears to have happened at the Little Grindstone site, which could also lead to the elimination of some alleles and changes in allele frequency through genetic drift, since there should be little to no selection occurring for these genes, unless they are linked to selected genes.

V. Acknowledgements

I would like to thank the numerous volunteers who have helped me collect field data and turtles and the landowners who have graciously given me permission to conduct the field research and turtle surveys that led to this project. Funding for this work was provided by a Student/Faculty Challenge Grant. I would especially like to thank my two advisors, Dr. Peter Rosenbaum and Dr. Amy Welsh, both of whom have provided much help and advice.

VI. Literature Cited

Bowen, K. D. & Gillingham, J. C. (2004). R9 Species conservation assessment for wood turtle –*Glyptemys insculpta* (LeConte, 1830). U. S. Forest Service, Eastern Region, Milwaukee, WI.

Daigle, C & Jutras, J. (2005). Quantitative evidence of decline in a southern Quebec wood turtle (*Glyptemys insculpta*) population. *Journal of Herpetology* 39(1): 130-132.

Ernst, C., Jeffrey, E., Lovich, E., and Barbour, R. W. (1994). *Turtles of the United States and Canada*. Smithsonian Institute Press, Washington, D. C.

Garber, S. & Burger, J. (1995). A 20-yr study documenting the relationship between turtle decline and human recreation. *Ecological Applications* 5(4): 1151-1162.

Harding, J. H. (1991). A twenty year wood turtle study in Michigan: implications for conservation. *First International Symposium on Turtles & Tortoises: Conservation and Captive Husbandry*: 31-35.

Saumure, R. A. & Bider, J. R. (1998). Impact of agricultural development on a population of wood turtles (*Clemmys insculpta*) in southern Quebec, Canada. *Chelonian Conservation and Biology* 3(1): 37-45.

Tessier N, Paquette, S. R., & Lapointe, F. J. (2005). Conservation genetics of the wood turtle (*Glyptemys insculpta*) in Quebec, Canada. *Canadian Journal of Zoology* 83(6): 765-772.

LONGEVITY OF THE NORTHERN SHORT-TAIL SHREW (*BLARINA BREVICAUDA*) AT RICE CREEK FIELD STATION, SUNY OSWEGO

Sara Ressing
Department of Biological Sciences

Long-term longitudinal studies can provide information on habitat quality over time. The Short-tailed shrew (*Blarina brevicauda*) is carnivorous, abundant in most habitats and lives 18 months or less. Ageing criteria based on weight or reproductive condition provide unreliable results since shrews are nearly adult weight at emergence from the nest and can begin reproduction shortly thereafter. The only reliable ageing criterion is microscopic evaluation of wear on cheek teeth of preserved specimens: not applicable to live animals. We used field measurements of incisors of marked individuals trapped repeatedly over a 7 month period to see if incisor length could be reliably related to known age. We live trapped short-tailed shrews (*Blarina brevicauda*) at SUNY Oswego's Rice Creek Field Station every 1-2 weeks, May- November 2007. Individuals were marked with PIT tags and released. We trapped 62 individuals a total of 263 times (max. captures/individual = 14; max. interval between first and last capture: 193 days). Tooth wear is measurable over longer intervals of time, and incisor length can be used to place individuals into 3 age classes.

I. Introduction

The northern short-tailed shrew is the most abundant small mammal at Rice Creek Field Station. They are known to be relatively short lived with starvation, exposure, predation, and old age the primary causes of mortality. Individuals in captivity have lived for over 30 months and wild caught specimens have been shown to live at least 15 months (Pearson, 1945).

Most long term longitudinal studies of shrew populations (e.g. Getz et al., 2004) lack aging criteria, a fundamental demographic component, due to the difficulty of accurately determining the age of living shrews. General criteria for an ideal age determination technique for short-tailed shrews include: independence from irregular nutritional and physiological variations; clear separations into age classes without subjective judgment; suitability for living animals of all ages; and easy application (Larson & Taber, 1980). Criteria of this caliber are thus far absent from the literature regarding age identification in the northern short-tailed shrew (*Blarina brevicauda*). The most reliable criterion requires microscopic evaluation of wear on the cheek teeth of preserved specimens: a method not applicable to live animals (Pearson, 1945; Pruitt, 1954; Choate 1968; Strait & Smith 2006).

Here we discuss the value of using length of the pigmented portion of the upper incisor, a measurement that is easy to take in the field, as an appropriate aging criterion. We test the

reliability of incisor pigment length as an indicator of age, and apply this measurement to a living population of *B. brevicauda* at Rice Creek Field Station, Oswego, New York.

II. Materials and Methods

A museum study of preserved material was first conducted to assess the accuracy of using incisor pigmentation as an indicator of age. The skulls of specimens (n=102) used in this study were collected throughout the central New York area, preserved by Dr. J. A. Lackey, and held at SUNY Oswego's Rice Creek Biological Field Station, Oswego, New York. The upper mandibles of these skulls were inspected under a dissecting microscope, examining the occlusion of the three molars and third premolar (Fig. 1). Wear classes were established using criteria published by Pruitt (1954) and Pearson (1945) based on wear on the occlusion surfaces of the molars. Measurements of the length of the pigmented portion of the right upper incisor were taken using calipers accurate to 0.01 mm without reference to molar wear class.

Incisors were measured twice, from beginning of pigmentation to the tip of the tooth. The repeated measures of incisor pigmentation were analyzed for experimenter consistency using Pearson's correlation.

A representative photographic record of the dentition of the upper mandible, as well as a lateral view of the upper incisor was taken for each wear class (Fig. 2).

After establishing and testing the aging criteria, we applied the aging method in a mark and recapture study of living individuals. Shrews were live trapped at SUNY Oswego's Rice Creek Field Station every 1-2 weeks from May-November 2007. Due to the mainly nocturnal summer activity cycle of *B. brevicauda*, Sherman live traps were set at dusk and checked 3-4 hours later. Individuals were marked with Biomark 8.5mm X 2.12 mm passive integrated transponders (PIT tags) and released. Animals were handled at night by headlamp, and immobilized in rip-stop nylon cones. Basic data were recorded on each animal at the time of each capture, including weight, reproductive status, total body length, tail length, and right hind foot length. Incisor pigment length was measured using a set of calipers.

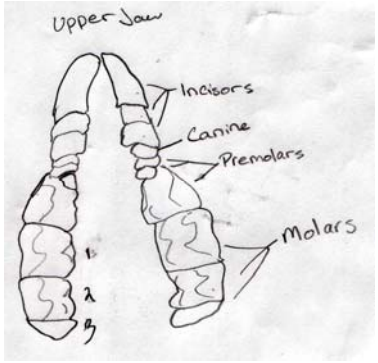


Fig. 1: Dorsal view of *B. brevicauda* upper jaw. Dental formula P3M 1M2M3

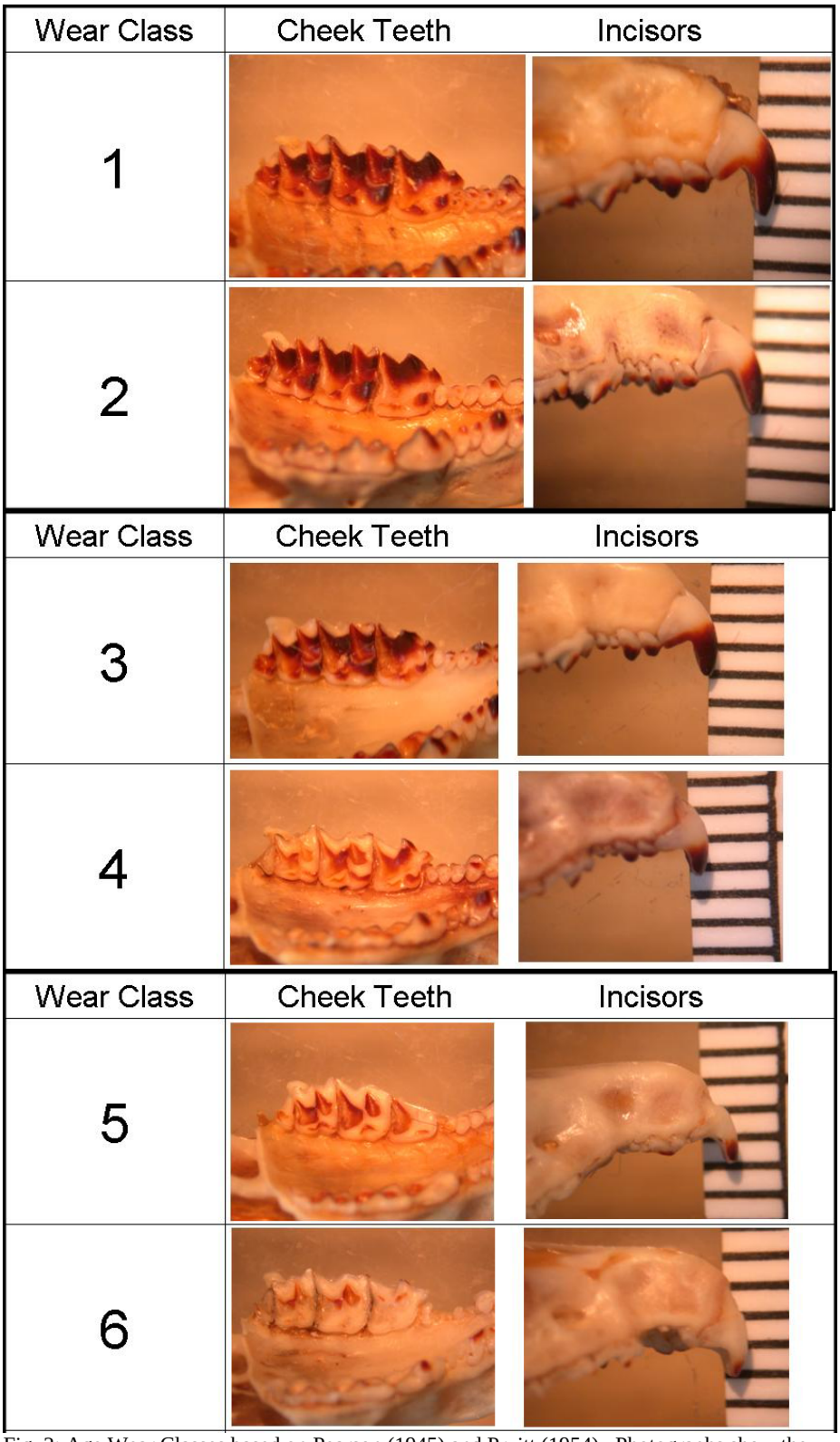


Fig. 2: Age Wear Classes based on Pearson (1945) and Pruitt (1954). Photographs show the dentition of the upper mandible and lateral view of the upper incisor at each wear class.

III. Results

The relationship between incisor pigment length and wear class based on wear of the occlusal surface of the molars was highly significant (r= - 0.768, p= 0.000, n=102, Pearson's Correlation); however there was some overlap between wear classes (Fig. 3). Observer consistency was very high for incisor pigment length remeasure (r = .950, p = 0.000, n=102, Pearson's Correlation).

Table 1 shows range of incisor pigment length for age groups. Age classes were assigned based on Pearson's (1945) established wear classes with young adults encompassing classes one and two; adults expected to have teeth consistent with wear classes three and four, and old adults would be considered in wear class five and six.

In the field study, 62 individual shrews were trapped a total of 263 times. A regression of tooth wear over time using only the first and last measurements showed significant change for the shrews of the upper field (P = 0.0001; R^2 = 37%) but not for the lower field (P = 0.903; R^2 = 0%) (Fig. 4). A comparison of total change in incisor pigment length and days between first and last capture (Fig. 5) shows considerable variability over shorter periods of time, but measurable decrease in incisor length over longer periods of time. Average change in one month was less than one millimeter, however after two months, the average change, 0.21 mm was substantial enough to be measured using calipers Standard deviation in changes in incisor pigment length (SD=0.3 mm) is not large enough to move individuals from one age class to another.

In Fig. 6, age profiles over the trapping period are shown. On average, 64% of individuals captured were adults, 29.6% were young adults, and 6.4% were old adult animals based on the criteria outlined in Table 1. Fig. 7 shows the longevity of animals marked at the upper and lower field. The maximum number of captures per individual was 14, and the largest interval between first and last capture was 193 days. Forty-eight percent of marked animals were not subsequently recaptured. The average number of captures per individual was 3.5.

Efforts were made to reduce trap mortality in the shrews. The use of cardboard inserts reduced trap mortality from 40% to 30.4% (Chi-squared, P=0.0104); and using animal protein as bait in the trap reduced trap mortality from 28.8% to 12.9% (Chi-squared, P=0.0010); but only very frequent trap checks (<4 hours) reduced trap mortality to an acceptable level of 3%. These results were reported in a poster presented by Chepko-Sade and Gerald Zoanetti at the American Society of Mammalogists meetings in 2001,

Table 1: Age categories based on incisor pigment length in *B. brevicauda*.

Age Category	Pigmented Part of Incisors (mm)
Young Adult	>2.0
Adult	1.0 – 2.0
Old Adult	<1.0

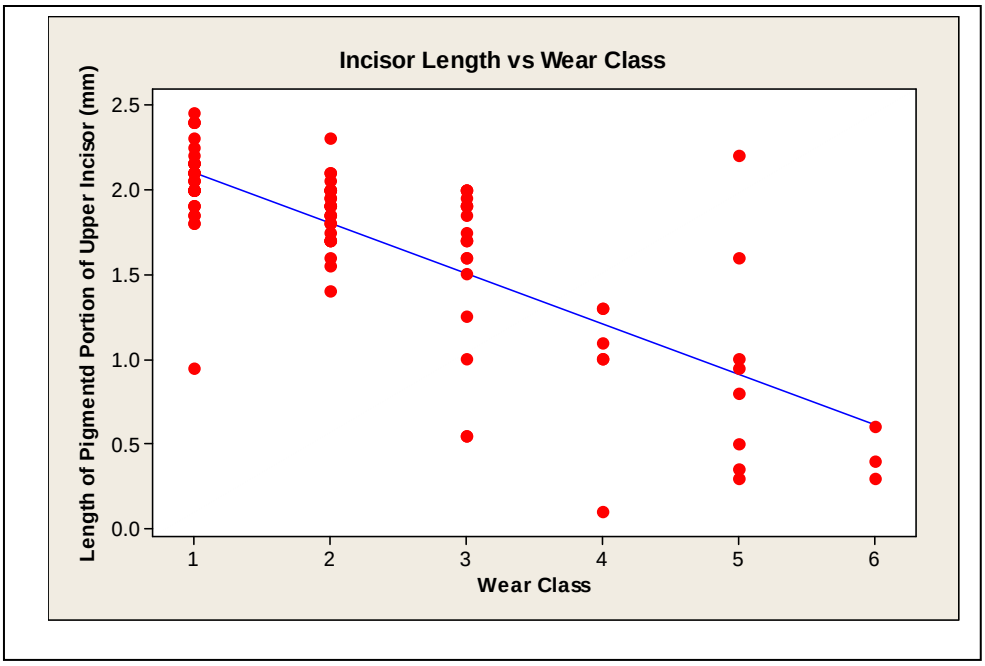


Fig. 3: Degree of incisor wear is highly correlated with degree of wear on the cheek teeth (r = -0.768, p = 0.000, n=102).

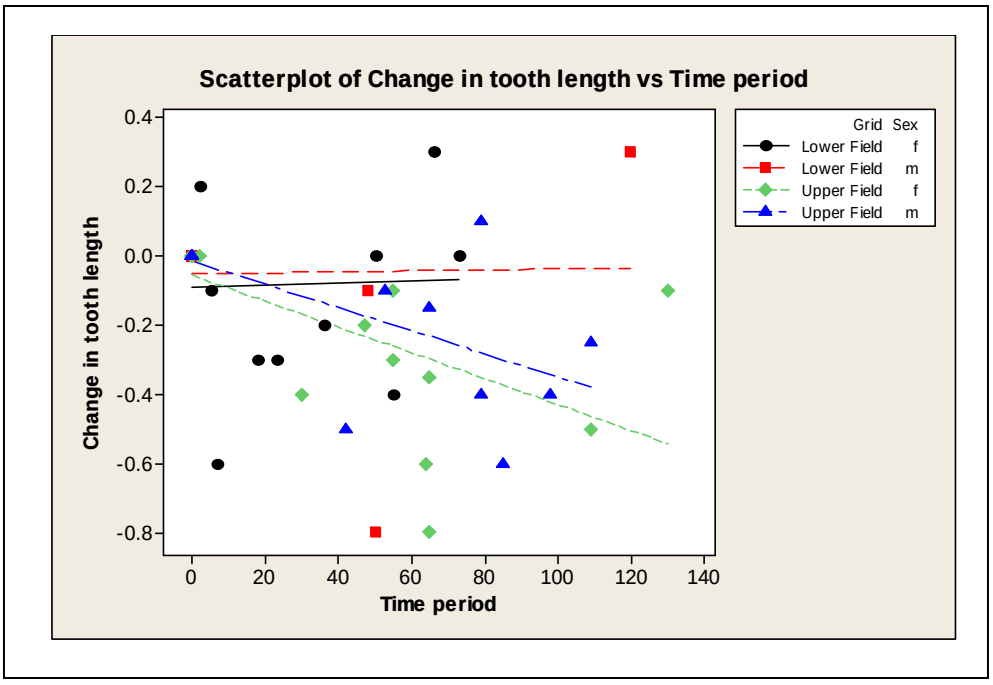


Fig. 4: Regression of tooth length over time shows a significant change for individuals in the upper field (Upper Field: P = 0.0001; R^2 = 37%, n=37), but not for the lower field (Lower Field: P = 0.903; R^2 = 0%; n=28).

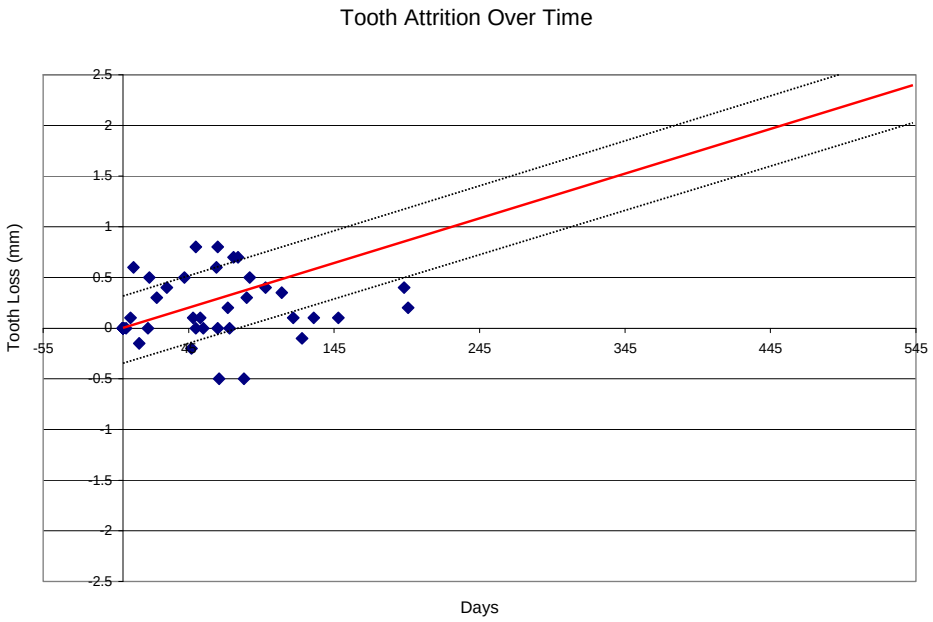


Fig. 5: Reliable measures of wear may take 18+ months. Total tooth loss over capture interval is depicted. The red line shows the projected change in tooth length, with ±0.3mm standard deviation shown by dotted lines.

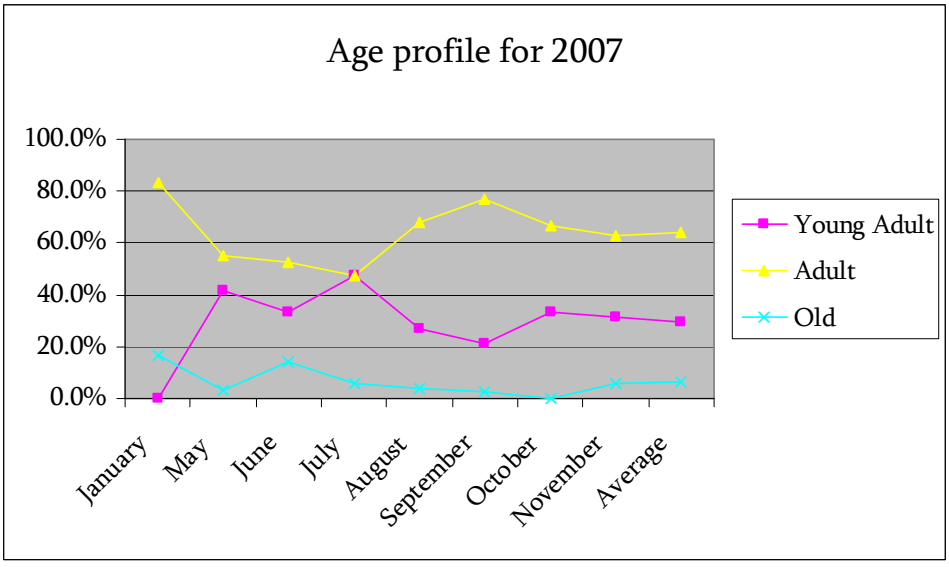


Fig. 6: Graphed age profile indicates the progression of aging as spring young mature, and adults from the previous spring die off. Over the trapping period 64% of individuals captured were adult, 29.6% were young adult, and 6.4% were old adult animals.

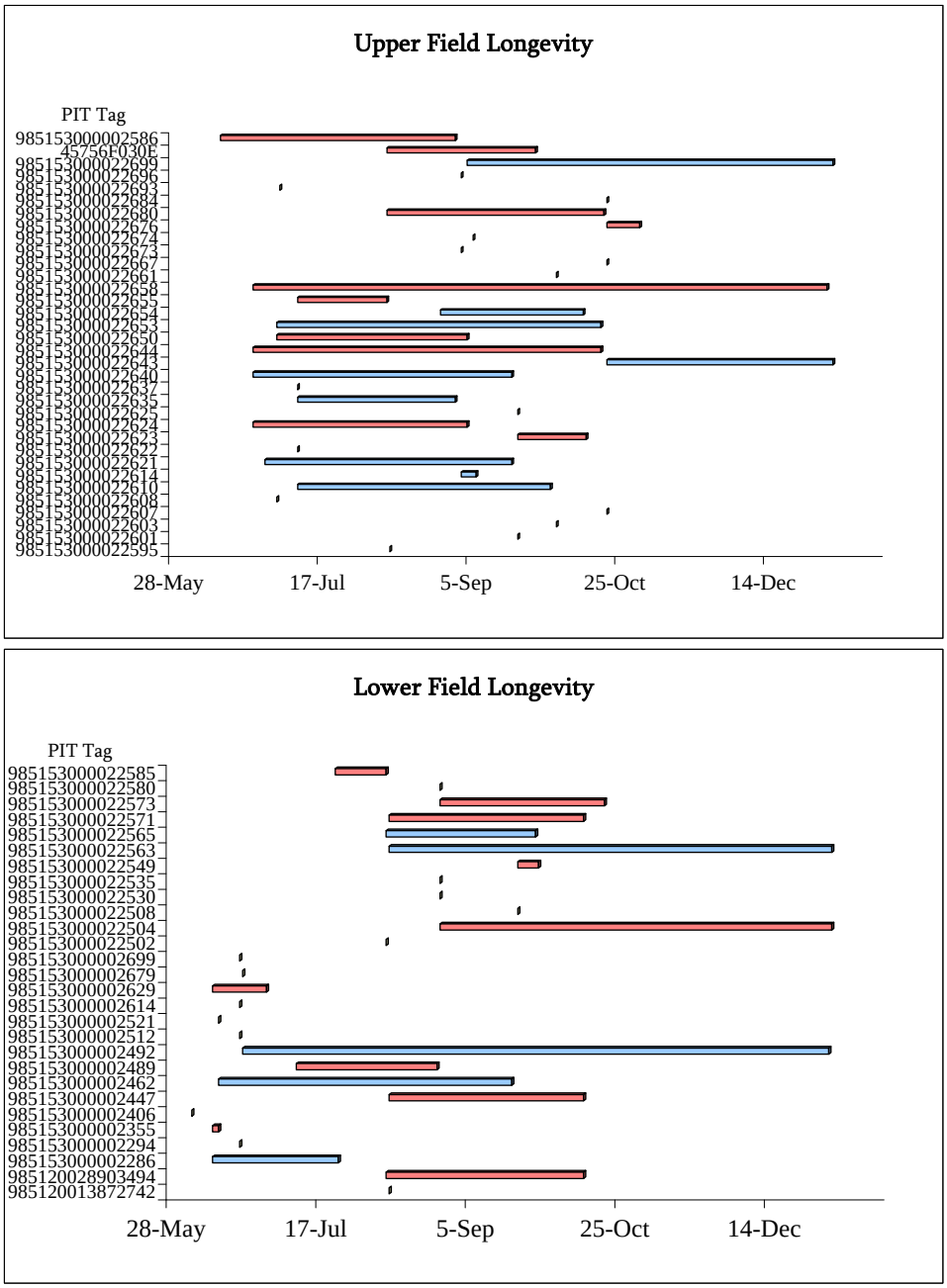


Fig. 7: Persistence of short tailed shrew on the upper and lower field grids. Date of the first and most recent capture for each individual is shown. The largest interval between first and last capture was 193 days. Blue bars are males, pink are females.

IV. Discussion

Incisor pigment length is an accurate and reliable indicator of age in short-tailed shrews and can be used to place individuals into three age classes as shown in Table 1. As individuals

age, their teeth wear down which is confirmed by decreases in the amount of pigment measured on the upper incisors. The practical application of using these criteria for aging animals in the field is clear. Individuals can be placed in age groups (Fig. 6) to create a demographic profile for a living population for comparison with results of previous studies based on dead individuals (Pearson, 1945). This allows the periodic examination of the age profile of the same population at different times without the impact of removing large numbers of individuals from the population to ascertain age profiles based on cheek teeth attrition.

Mortality for this trapping session was greatly reduced compared to mortality recorded in previous seasons (Chepko-Sade, unpublished data). *B. brevicauda* weaken and die quickly in Sherman traps due to their fast metabolism and the stress of captivity. Frequent trap checks reduced mortality the greatest amount, but the inclusion of cardboard inserts and animal protein in the traps also had an effect.

Due to small sample sizes and high variation between successive measurements of the same individual, further trapping is needed to alleviate the spread in the data. The project will be continued over the summer in order to further quantify the rate of tooth attrition over time so that individuals may be more precisely aged rather than just assigned to age classes. This information will allow us to study the demographics of this population in more detail.

V. Acknowledgements

Dr. B. Diane Chepko-Sade, Dr. J. A. Lackey, and Dr. Andrew Nelson. Field Assistants: Matthew Volny, Wendy Paterson, and Jenna McAdoo Funding received from Student Faculty Collaborative challenge Grant and Rice Creek Associates Grant.

VI. References

Choate, J. R. (1968). Dental Abnormalities in the Short-tailed Shrew, *Blarina brevicauda*. *Journal of Mammalogy*, 49(2), 251-258.

Getz, L. L., Hofmann, J. E., McGuire, B., Oli, M. K. (2004). Population Dynamics of the Northern Short-tailed Shrew, *Blarina brevicauda*, Insights from a 25-year study. *Canadian Journal of Zoology*, 82, 1679-1686.

Larson, J. S. & Taber, R. D. (1980). Criteria of Sex and Age. In S. D. Schemnitz (Ed.), *Wildlife Management Techniques Manual* (4th ed., pp. 143-202). Washington: The Wildlife Society, Inc.

Pearson, O. P. (1945). Longevity of the Short-tailed Shrew. *American Midland Naturalist, 34*(2), 531-546.

Pruitt, W. O., Jr. (1954). Notes on the Shorttail Shrew (Blarina brevicauda kirtlandi) in Northern Lower Michigan. *American Midland Naturalist, 52*(1), 236-241.

Strait, S. G. & Smith, S. C. (2006). Elemental Analysis of Soricine Enamel: Pigmentation Variation and Distribution in Molars of Blarina brevicauda. *Journal of Mammalogy, 87*(4), 700-705.

EVALUATION OF PERSISTENT CONTAMINANT LEVELS IN GREAT LAKES SALMON EGGS

Colleen M. Alexander & James J. Pagano
Environmental Research Center, Department of Chemistry, 310 Piez Hall,

Chinook, coho, and steelhead eggs collected in 2006 from Lake Ontario, L. Erie, L. Michigan, and L. Superior were analyzed by gas chromatography for congener-specific polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (PBDEs), and select organochlorine pesticides (OCs). Across lake comparisons for total PCB, total DDT, and total PBDE indicated a consistent trend in contaminant concentrations in which Lake Michigan > Lake Ontario > Lake Erie > Lake Superior. The contaminants Mirex and octachlorostyrene (OCS) were found not to be consistent with this trend, but were instead found to exist in higher concentrations in Lake Ontario in relation to the other Great Lakes. PCB chlorination patterns, as indicated by the average number of chlorines per biphenyl (Avg Cl/BP) were compared within species. The PCB chlorination pattern in chinook eggs was found to be similar for both Lake Michigan (5.47) and Lake Ontario (5.43). The PCB chlorination pattern in coho eggs was notably different in Lake Superior (5.78) than in Lake Ontario (5.41).

I. Introduction

The USEPA Great Lakes Fish Monitoring Program (GLFMP) is designed to evaluate contaminant concentrations in Great Lakes sport and top predator fish, evaluate human exposure to contaminants, and obtain new information on persistent and emerging contaminants. As a part of the GLFMP sport fish program, the purpose of this preliminary research was to investigate the utilization of salmon eggs as bioindicators of persistent and emerging contaminants in the Great Lakes.

The utilization of bioindicators in persistent contaminant research is widely discussed in the literature. Species such as snapping turtles and tree swallows have been widely used as reliable bioindicators of persistent contaminants in the environment (Pagano et al., 1999; Nichols et al., 1995). The utilization of fish eggs as biomonitors of persistent contaminants has advantages including: ease and cost effectiveness of collection and analysis; large composite egg batch samples reflecting entire spawning populations; and reduced adverse effects to the general salmon population (Miller, 1993; Miller, 1994; Pagano, 2005; Pagano, 2007).

II. Materials and Methods

The Salmon River Fish Hatchery located in the Village of Altmar, Oswego County is operated by the New York State Department of Environmental Conservation (NYSDEC) to meet the sport fishery stocking needs of Lake Ontario and tributaries. Lake Ontario coho (*Oncorhynchus kisutch*) and chinook (*Oncorhynchus tshawytscha*) salmon eggs were randomly sampled by NYSDEC personnel at the Salmon River Fish Hatchery. Sampling locations for Lake Superior, Lake Michigan, and Lake Erie eggs are noted in Fig. 1. Egg samples were collected during the Fall 2006 spawning run. For the Lake Ontario samples, paired egg and muscle fillet (skin off) samples were collected from individual coho and chinook salmon before fertilization as previously described (Pagano, 2005b).

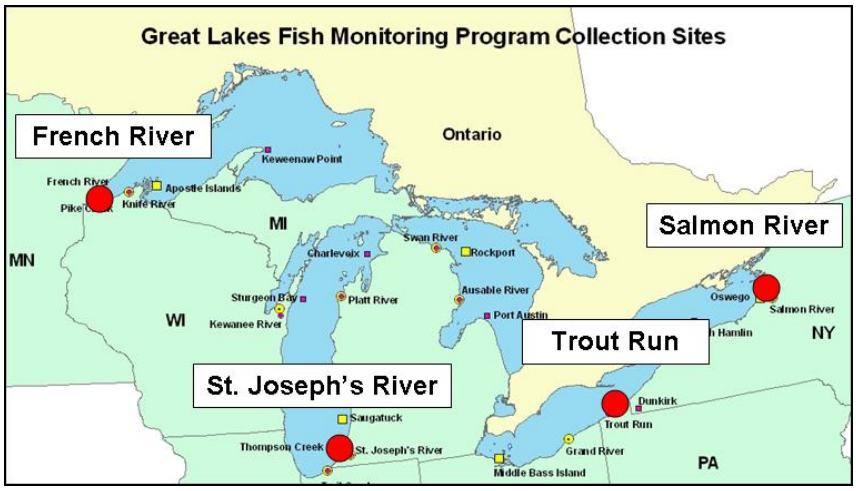


Fig. 1: Great Lakes 2006 fish egg collection sites. Map obtained from USEPA GLNPO, 2008.

Egg samples were extracted at Clarkson University by Accelerated Solvent Extraction (ASE) using a Dionex ASE 300 and dichloromethane solvent. The sample extract was condensed to 2 mL in a TurboVap II for Gel Permeation Chromatography (GPC) cleanup. Sample cleanup followed EPA method 3640A (GPC cleanup - pesticide option) using a Waters GPC system (binary pump, Envirogel® column, UV detector, and fraction collector) followed by silica gel column for separation of PCBs/OCs/PBDEs from other interferences. Adsorption column chromatography clean-up utilized 5.5 grams of 4% deactivated silica (Sigma-Aldrich, grade 923, 100-200 mesh) placed in a chromatography column and sample extract sequentially eluted with 25 mL of hexane (fraction 1) and 25 mL hexane:dichloromethane (1:1, fraction 2). Each fraction was then concentrated in a TurboVap II to 2 mL for gas chromatographic analysis.

Congener-specific PCB, hexachlorobenzene, p-p' DDE, and Mirex analyses were conducted based on capillary column procedures previously described (Pagano et al., 1995; Pagano, 2005a; Pagano, 2007b). Briefly, analytical instruments were recalibrated every five samples, with a system blank, instrument blank, and mid-level calibration check

solution analyzed during each analytical run. A Hewlett-Packard (HP) Model 5890II GC with an electron capture detector (ECD - Ni^{63}) and autosampler were used for primary data acquisition. The capillary column utilized was a HP Ultra II, 25 meter with 0.22 mm id and 0.33 *u*m film thickness. The calibration standard used was a 1:1:1:1 mixture of Aroclors 1221, 1016, 1254, and 1260 each at 200 pg/uL, hexachlorobenzene (HCB) at 5 pg/uL, and p-p' DDE and Mirex each at 10 pg/uL (Custom Mixed Fraction #3, AccuStandard, Inc.), which allows for the analysis of 99 chromatographic zones of 132 congeners/co-eluters. PCB analyses were confirmed with a HP Model 5890 II gas chromatograph with an electron capture detector (Ni^{63}) and autosampler using a 60 meter DB-XLB capillary column with 0.25 mm id and 0.25 *u*m film thickness. The calibration standard used was a 1:1:1:1:1 mixture of congener mixture sets C-CSQ-SET 1-5; 10 pg/uL per individual congener (AccuStandard, Inc., New Haven, CT). PCB data was processed using HP ChemStation software and Microsoft Excel spreadsheet procedures such that the mole percent (congener specific and homolog), and average chlorine/biphenyl (Avg Cl/BP) values were generated.

Congener-specific PBDE analyses were conducted with HP Model 5890 II gas chromatograph with an electron capture detector (Ni^{63}) and autosampler using a 60 meter DB-XLB capillary column with 0.25 mm id and 0.25 *u*m film thickness. The PBDE calibration standard used was an 800 pg/uL (total PBDE = $\sum$ 12 components) solution using the original Great Lakes Chemical Corporation (Great Lakes DE-71, CAS 32534-81-9) technical formulation. The DE-71 technical formulation mass fractions and congener identifications were confirmed with pure PBDE congener standards (BDE-MXE) purchased from Wellington Laboratories, Guelph, ON, Canada, and by mass spectrometric confirmation.

II. Results and Discussion

A comparison of select contaminant and percent lipid data for five egg sample groups is presented in Table 1. As noted, total PCB concentrations ranged from 863.1 ng/g in Lake Michigan chinook eggs to 105.3 ng/g in Lake Superior coho eggs. Similar across-lake concentration trends are noted for total PBDE and total DDT. Concentrations in these three contaminants groups were found to be greatest in Lake Michigan, followed by L. Ontario, L. Erie, and L. Superior. In comparison, lake trout whole fish data from the USEPA Great Lakes National Program Office (GLNPO) Great Lakes Environmental Database (GLENDA) is reported in Fig. 2. Similar to the fish egg trend noted in Table 1, total PCB concentrations for 2003 lake trout whole fish were found to be greatest in Lake Michigan, followed by L. Ontario, L. Huron, L. Erie, and L. Superior. The similarity of fish egg and lake trout whole fish trend data appears to indicate that fish eggs have the potential to act as bioindicators of contaminants in Great Lakes fish.

Table 1. Select contaminant data for five egg sample groups (concentrations expressed in ng/g – wet weight).

	LM Chinook (N=10)	**LO Chinook (N=3)**	**LO Coho (N=3)**	**LE Steelhead (N=7)**	**LS Coho (N=3)**
Total PCB	863.1	663.7	428.8	311.8	105.3
Ave Cl/BP (PCB)	5.47	5.43	5.41	5.49	5.78
% Lipid	10.3%	5.2%	3.7%	11.5%	8.9%
Total PBDE	55.2	40.3	25.6	23.2	23.2
Total DDT	291.9	267.3	154.9	40.8	34.7
Mirex	1.2	26.0	19.4	0.6	0.6
OCS	0.4	4.9	3.4	0.4	0.4

Contrary to total PCB, PBDE, and DDT across-lake trends, Mirex is present in significantly higher concentrations in Lake Ontario than any of the other Great Lakes. For example, L. Ontario chinook egg Mirex concentrations (26.0 ng/g wet) are significantly higher than chinook eggs from L. Michigan (1.2 ng/g wet). Mirex was released into Lake Ontario by the Hooker Chemical company from the Niagara River and, to a lesser extent, the Oswego River (Pagano, 2004). Our fish egg results are consistent with the established fact that Mirex is a contaminant associated with Lake Ontario and not significantly present in the other Great Lakes.

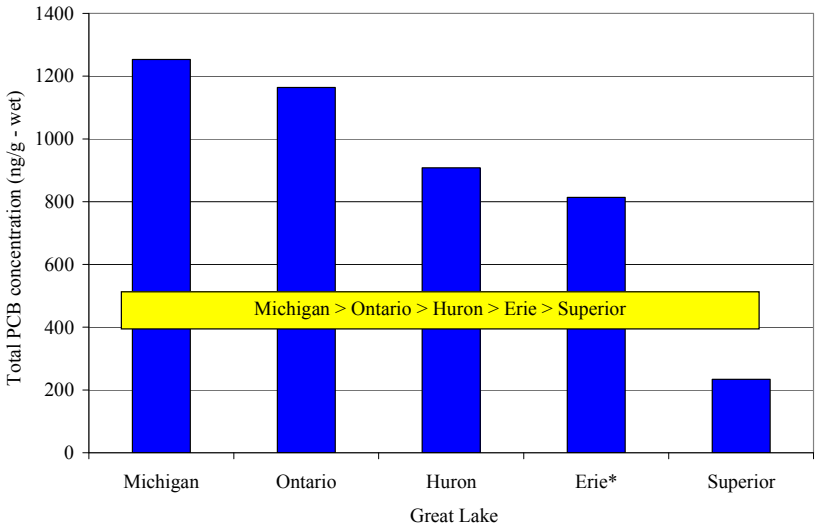


Fig. 2. Across-lake total PCB concentrations in 2003 lake trout whole fish. Data obtained from EPA GLNPO GLENDA database. * Lake Erie samples are walleye.

In order to validate use of eggs as environmental biomonitors, it is important to demonstrate if quantitative and qualitative relationships exist between contaminants found in eggs and contaminants found in muscles. In Fig. 3, a linear regression analysis of PCB congener concentrations in 2006 eggs versus muscle tissue was performed. Results indicate (R^2 = 0.97 for coho, and R^2 = 0.98 for chinook) a very strong quantitative fit

between egg and muscle PCB congener concentrations for both species. Although the egg to muscle relationships are different for coho and chinook, both demonstrate a very strong quantitative relationship indicating that egg PCB congener concentrations have the potential to be highly predictive of muscle PCB congener concentrations.

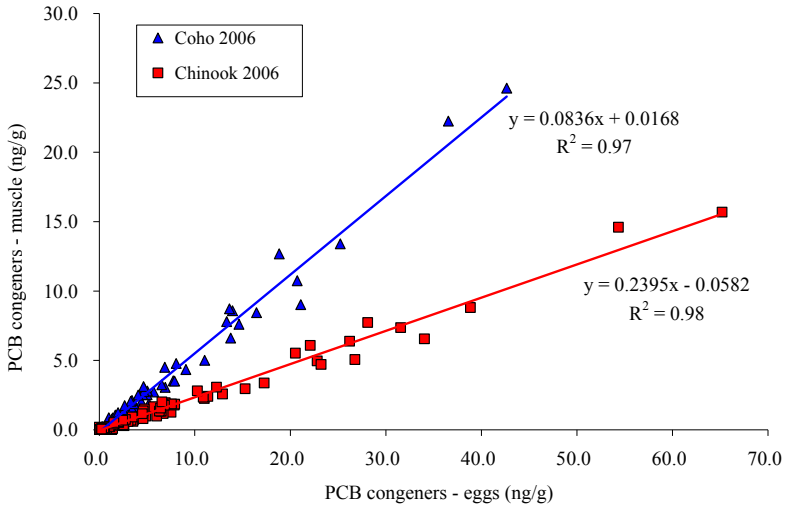
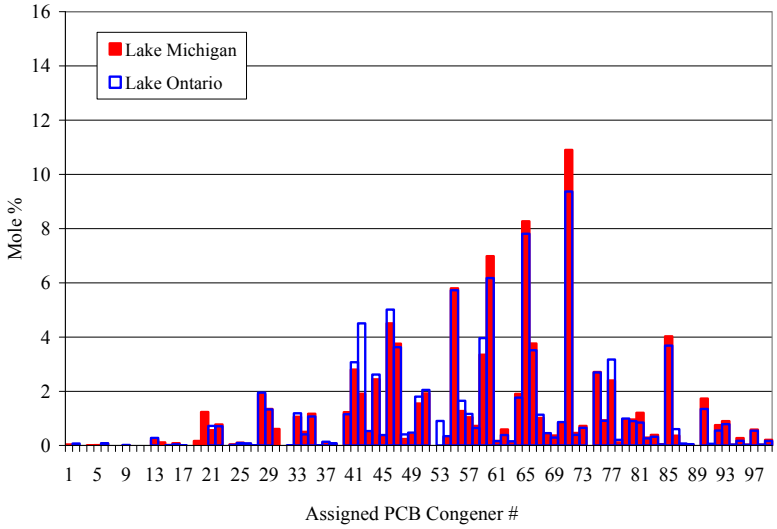


Fig. 3. Comparative linear regression analysis for 2006 coho and chinook average muscle and average egg PCB congener concentrations. Data adapted from Pagano, 2007a.

PCBs in the environment are a complex mixture of 209 possible congeners, in which the number of chlorines substituted per biphenyl may vary from 1-10. In Fig. 4, PCB assigned congener numbers are displayed by mole percent in order to illustrate PCB chlorination patterns, independent of concentration in each fish-egg group. The lower assigned PCB congener numbers correspond to lower-chlorinated PCB congeners, and the higher assigned PCB congener numbers correspond to higher-chlorinated PCB congeners. In Fig. 4a, visual observation of L. Ontario and L. Michigan chinook egg data demonstrates very similar PCB chlorination patterns. As contrasted in Fig. 4b, a comparison of the L. Ontario and L. Superior coho data demonstrates a markedly different (and higher-chlorinated) PCB chlorination pattern in L. Superior eggs.

(4a.)



(4b.)

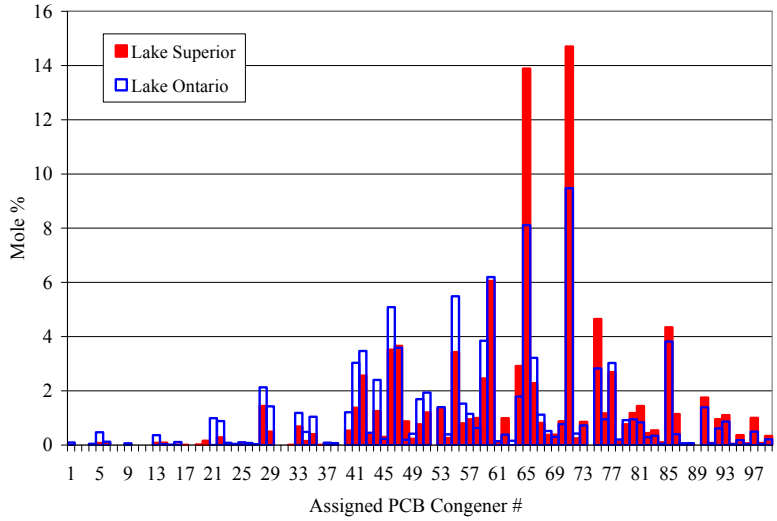


Fig. 4. Congener-specific PCB mole percent comparison of L. Ontario versus L. Michigan chinook eggs (4a.) and L. Ontario versus L. Superior coho eggs (4b.).

In Fig. 5, PCB homologs are displayed by mole percent and average number of chlorines per biphenyl in order to represent the PCB chlorination for each sample group. Homologs are subcategories of PCB congeners having equal numbers of chlorine substituents. As noted in Fig. 5a, similar mid-chlorinated homolog patterns and Avg Cl/BP values are observed in both L. Ontario (5.43) and L. Michigan (5.47) chinook eggs. As contrasted in Fig. 5b, notably different homolog chlorination patterns and Avg Cl/BP values are observed for L. Superior (5.78) and L. Ontario (5.41).

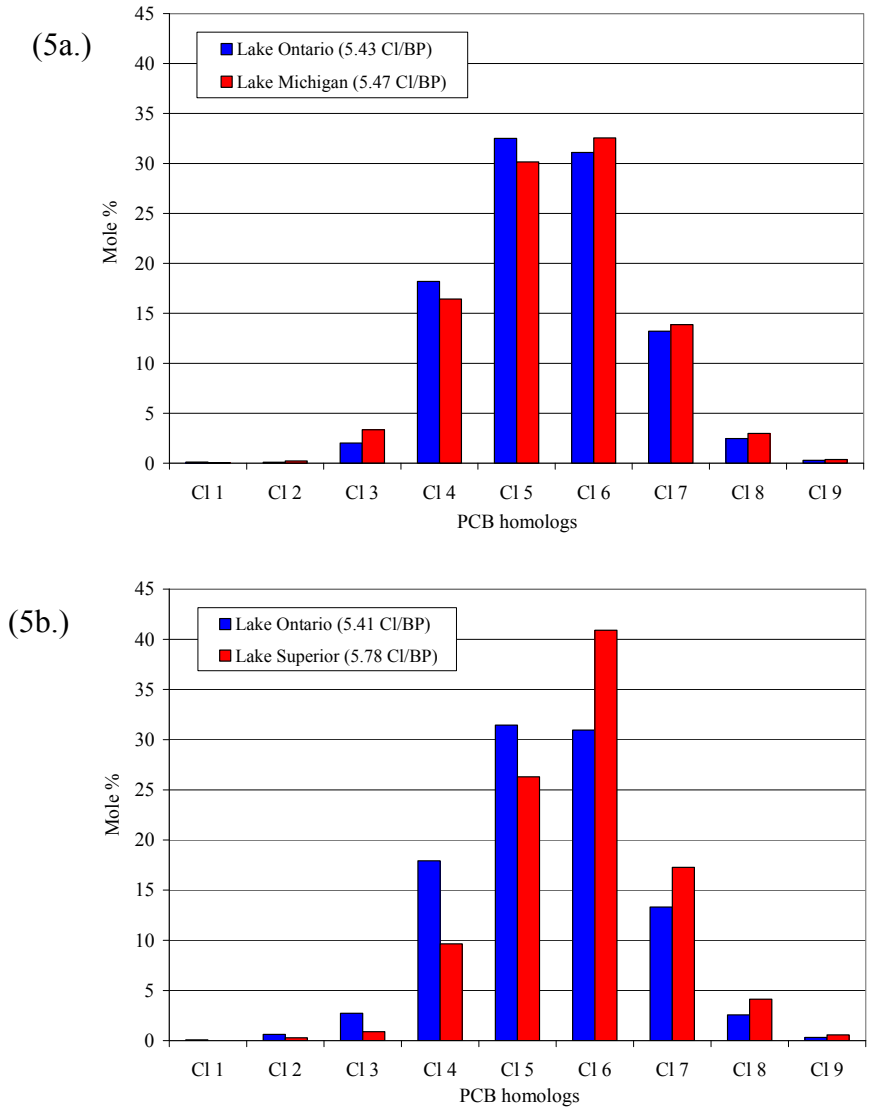


Fig. 5. Mole percent PCB homolog comparison for L. Ontario versus L. Michigan chinook eggs (5a.) and for L. Ontario versus L. Superior coho eggs (5b.).

Determination of the qualititative relationship between PCB congeners found in fish eggs and muscle tissue is important for evaluating the reliability of fish eggs as biomonitors. In Fig. 6, congener-specific PCB mole percent values for 2006 coho and chinook were submitted to a linear regression analysis for egg and muscle tissue data. A strong qualitative fit between egg and muscle tissue is indicated for both coho (R^2 = 0.98) and chinook (R^2 = 0.98). These results qualitatively indicate that the contaminants found salmon eggs are highly predictive of the contaminants found in muscle tissue.

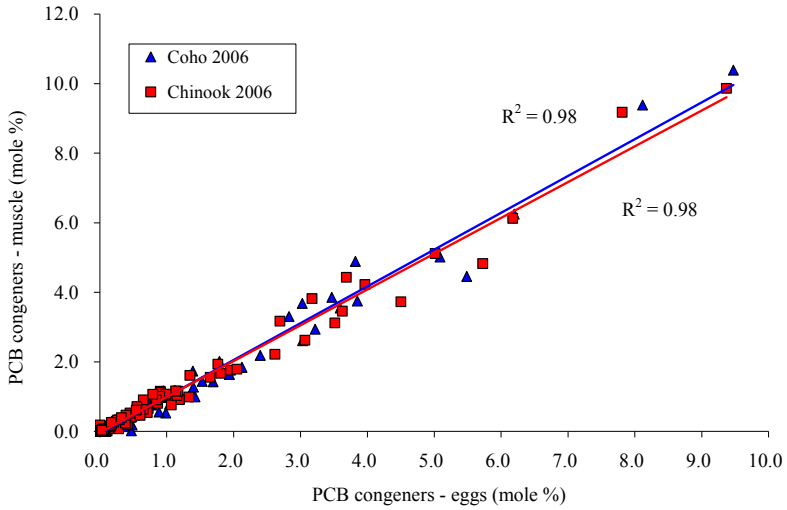


Fig. 6. Comparative linear regression analysis for 2006 coho and chinook average muscle and average egg PCB mole percent. Data adapted from Pagano, 2007a.

IV. Conclusions

Results from preliminary data indicate an across-lake contaminant concentration trend for total PCB, PBDE, and DDT in which concentrations are highest in Lake Michigan, followed by L. Ontario, L. Erie, and L. Superior. Our data suggests that fish eggs are highly predictive both quantitatively and qualitatively of congener-specific PCBs found in muscle tissue. Although preliminary, our results to date suggest that fish eggs have the potential to be utilized as bioindicators of persistent contaminants in the Great Lakes.

V. Acknowledgements

The authors would like to thank the USEPA Great Lakes Fish Monitoring Program for funding this research. We also thank Phil Hulbert and Andy Gruelich, NYS Department of Environmental Conservation for providing samples and technical advice. We would also like to thank Beth Murphy – EPA Project Manager, Greg Sumner – SUNY Oswego ERC for assistance in preparing samples, and Bernie Crimmins – Clarkson University for extracting samples.

VI. References

Miller, M. A. (1993). Maternal transfer of organochlorine compounds in salmonines to their eggs. *Can J. Fish. Aquat. Sci.*, 50:1405-1413.

Miller, M. A. (1994). Organochlorine concentration dynamics in Lake Michigan chinook salmon. *Arch. Environ. Contam. Toxicol.*, 27: 367-374.

Nichols, J. W., Larsen, C. P., McDonald, M. E., Niemi, G. J., and Ankley, G.T. (1995). Bioenergetics based model for accumulation of polychlorinated biphenyls by nestling tree swallows, *Tachycineta bicolor*. *Environ. Sci. Technol*. 29(3): 604-612.

Pagano, J., Scrudato, R., Roberts, R. and Bemis, J. (1995). Reductive dechlorination of PCB contaminated sediments in an anaerobic bioreactor system. *Environmental Science & Technology*, 29:10 2584-2589.

Pagano, J. J., Rosenbaum, P., Roberts, R., Sumner, G., and Williamson, L. (1999). Assessment of maternal contaminant burden by analysis of snapping turtle eggs. *Journal of Great Lakes Research*, 25(4):950-962.

Pagano, J., (2004). Identification of a source of Aroclor 1268 and mirex to the Oswego River Area of Concern. In: *Proceedings of the 228th National Meeting American Chemical Society*, Philadelphia, PA, USA, Vol. 44 No. 2: 525-530.

Pagano, J. (2005a). Deposition and ambient concentrations of PBTs (PCBs, OCs, PBDEs, and Dioxins/Furans) in support of the Lake Ontario air deposition study (LOADS), Quality Assurance Project Plan (QAPP). Great Lakes Commission-Great Lakes Air Deposition (GLAD) Program, Ann Arbor, MI. June 28, 2005 – Version 01.

Pagano, J. (2005b). Utilization of salmonid eggs as bioindicators of organohalogen pollutants in Lake Ontario. *Proceedings of the 230th National Meeting American Chemical Society*, Washington, DC, USA, Vol. 45 No. 2: 337-342.

Pagano, J. (2007a). Polybrominated diphenyl ethers in Lake Ontario salmonid fillets and eggs. *Proceedings of the 234th National Meeting American Chemical Society*, Boston, MA, USA, Vol. 47 No. 2: 223-229.

Pagano, J. (2007b). Deposition and ambient concentrations of PBTs (PCBs, OCs, PBDEs, and Dioxins/Furans) in support of the Lake Ontario air deposition study (LOADS). Final Report: Great Lakes Commission-Great Lakes Air Deposition (GLAD) Program, Ann Arbor, MI. December 20, 2007.

United States Environmental Protection Agency GLNPO, (2008). Great Lakes Monitoring. http://www.epa.gov/glnpo/monitoring/fish/index.html, accessed March 11, 2008.

A GUI FOR EVOLVE ZAMS

D. R. Schlegel
Computer Science Department

Here the early work on a new user interface for the Evolve ZAMS stellar evolution code is presented. The initial goal of this project is to provide an easy way to generate and visualize stellar evolution tracks for use in the classroom or lab. Eventually the code will allow for more complex analysis looking at star models as they evolve along an evolutionary track, becoming more useful for the researcher.

I. Introduction

A study of stellar evolution intends to discuss how a star changes throughout its lifetime. This can be a very difficult subject governed by so many equations that it would be nearly impossible for someone to accurately determine the evolution of a star without the use of a computer program. A computer program can often make the process much less painful but an experiment can still be difficult to set up requiring a lot of scripting and in depth knowledge of both stellar evolution and the computer code you intend to use.

The code which was used for this project – Evolve ZAMS, or often just called EZ – is no different than this. Like most research software it shows signs of being written for a specific purpose, and trying to do anything outside of that is very difficult. Therefore it is not very user friendly and the learning curve is very steep. In order to run a simulation you must of course know the parameters of the star to wish to simulate, but you must also understand how EZ works.

After one has figured out how to use EZ and they have produced output (in the form of a tab-delimited table printed to the console), the task then becomes analyzing the data. The tables generated are often 20 columns wide and 50 lines or so, intermixed with debug statements about the status of the star on its evolutionary track. This is great if perhaps you are analyzing the age of a star at which it leaves the main sequence which is only one point, but many times a graph is desired.

My goal with this project was to make this process as a whole something which is easy for the user to do. My aim was to provide an end-to-end solution to the user, beginning with a graphical interface for setting variables about the star, and ending with a tabular look at the two variables the user chose with the option of generating a GNUPlot script which can be easily loaded.

II. Evolve ZAMS

Evolve ZAMS itself is a stellar evolution code written by Bill Paxton from the University of California at Santa Barbara. The evolution code itself is derived from a one dimensional code written by Peter Eggleton. As the name of the code implies, the program evolves stars from the Zero Age Main Sequence (ZAMS) to a specified stopping point (Paxton, 2007).

The application uses simulation files which are compiled FORTRAN code calling each module required individually. This is more difficult to use than a scripting language since there are preconditions and postconditions to take into account with each module. Not limiting the user to a predefined set of commands provides a lot of flexibility for the user if they are willing to become acquainted with both FORTRAN and the modules available in the code.

The problem with this though is that the documentation available for EZ is slim to none. The main form of documentation is in the two demo simulations distributed with the code. These can be easily modified to change the mass of a star and some other basic variables but they do not do everything a user might be looking for, and they don't make it clear how they do much of anything. In order to produce results the user really wants they need to dig into the FORTRAN code, for which there is a browser supplied on the web site.

III. Basic Stellar Plots

Someone studying stellar evolution very often does not use just numerical values to do their analysis, graphs are very common. As a star evolves a great deal of variables need to be taken into account – from the temperature at the very center of a star, to the concentration of Helium in the star. There are a lot of very interesting plots possible, but perhaps the most common one is plotting a stars evolution on a HR – or Hertzsprung-Russell - diagram.

In its most basic form, an HR diagram plots the logarithm of the luminosity of a star (measured in terms of the luminosity of our sun) versus the log of the temperature of a star in a reversed axis. An example of this is provided in Fig. 1.

This figure shows this relationship for many stars during a single snapshot of time. From the image you can see where the main sequence lies, and where some stars which have evolved off of the main sequence now reside. While this plot can come largely from observational data, on these same axes a plot can be made of a single stars evolution from the main sequence through its different evolutionary phases.

Shown in Fig. 2 is an example of one such plot. A plot like this is impossible to arrive at observationally since stars exist for billions of years, but it can be arrived at mathematically and therefore, computationally. A plot like this shows much of how a stars appearance to the outside world will change as it evolves (since luminosity is related to brightness, and surface temperature is related to color).

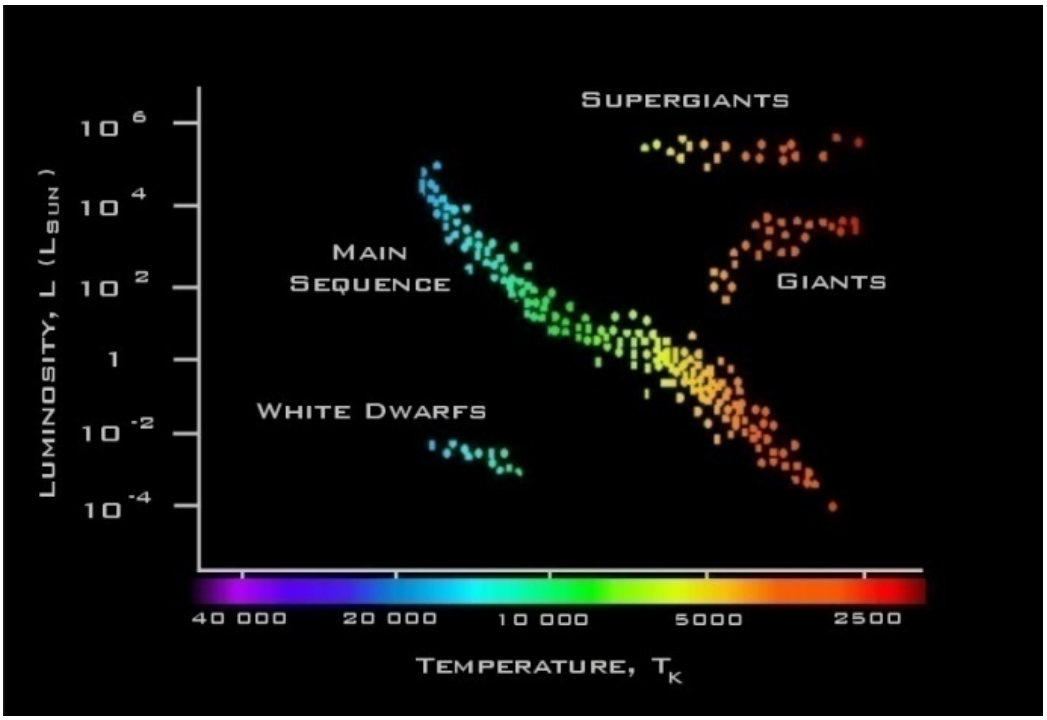


Fig. 1: HR Diagram (Brinkworth & Thomas, 2001)

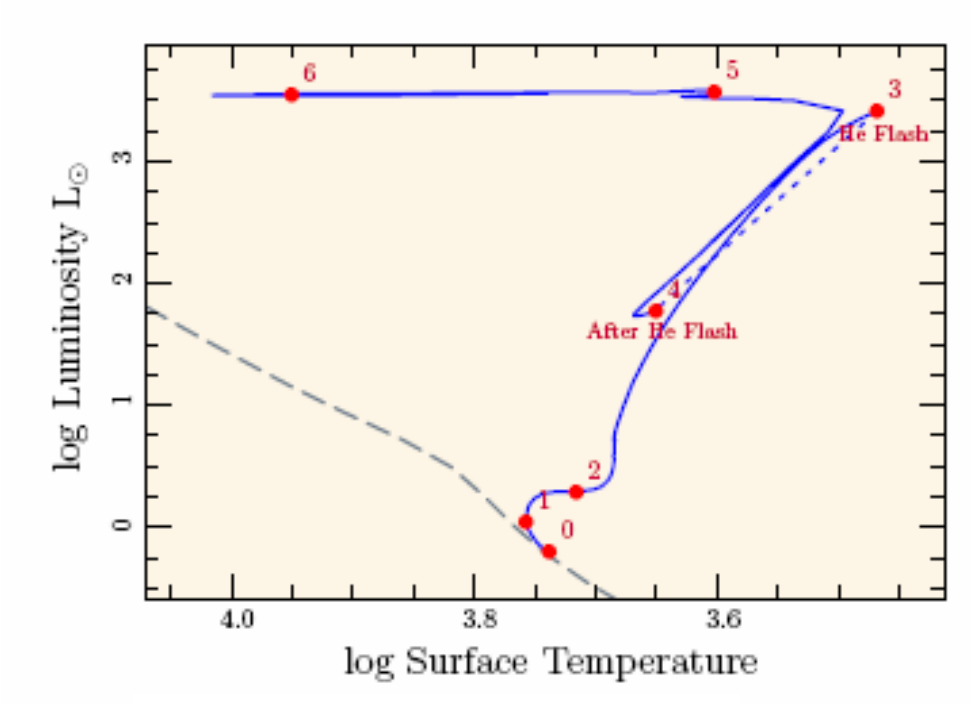


Fig. 2: Star Evolution (M=1.0, Z=0.3) (Paxton, 2007)

IV. The Project

The concepts behind stellar evolution can be very difficult to understand, but I feel that the software to simulate it doesn't have to be. The goal of this software is for it to be able to be used by students, professors, and researchers in their generation and analysis of stellar evolution graphs. The project design in general consists of three phases – generating/running the simulation (input), allowing the user to choose the data which is important to them, and the output phase in which a graph is created.

The first phase is that which abstracts the user from having to deal with the intricacies of the EZ stellar evolution package itself. The interface provides to the user options so that they can set the mass, metalicity, etc of the star they wish to evolve. Based on the user's selections the program automatically generates the FORTRAN code that EZ requires to run a simulation. This is compiled and run by the GUI, never requiring the user to touch a command prompt.

One the simulation has finished (after a couple of minutes usually), the tabular format of the data is broken down into its individual parts automatically. EZ tries to make itself more user friendly by dispersing throughout the output table comments about when the star is leaving the main sequence, when the helium flash occurs, etc. The program therefore has to clean up this data before it is presented to the user for use in graphing.

Most of these messages dispersed throughout the output can actually be turned off by editing the source code sufficiently, but this is not necessarily desirable. As the project evolves it is very likely that it will be desirable to plot these events on the graph, so keeping them around is in our best interest.

Once the data has been parsed it is presented to the user. The user chooses from drop down boxes which data they would like to plot (the column headers are preserved to make this easier). As this is done a preview of what the results look like graphically are shown to the user in a preview pane. This can help reduce errors choosing X and Y properly.

Finally once the user has selected what data they are interested in and selected any plotting preferences (such as reversing an axis or setting the ranges manually), the data can be output to a GNUPlot script. This includes both the data file and a script to load it and set the preferences the user selected.

V. Guaging Success

In order to call this project a success it was decided that it needed to plot some basic stellar evolution paths without forcing the user to touch any command line tools. While the project is in desperate need of usability testing, these goals appear to have been met. Fig. 3 is an evolution path created completely using this user interface for a star similar to our own sun.

Creating a graph such as this is very straight forward for anyone, even someone who knows nothing about how stellar evolution actually works. The goal has been to provide as many options to the user as possible while not asking the user to fill in many fields manually. This makes it so the user always knows what the possible values are and doesn't let the user "mess up." As time goes on and the amount of options increases one goal of the application should be maintaining this ease of use.

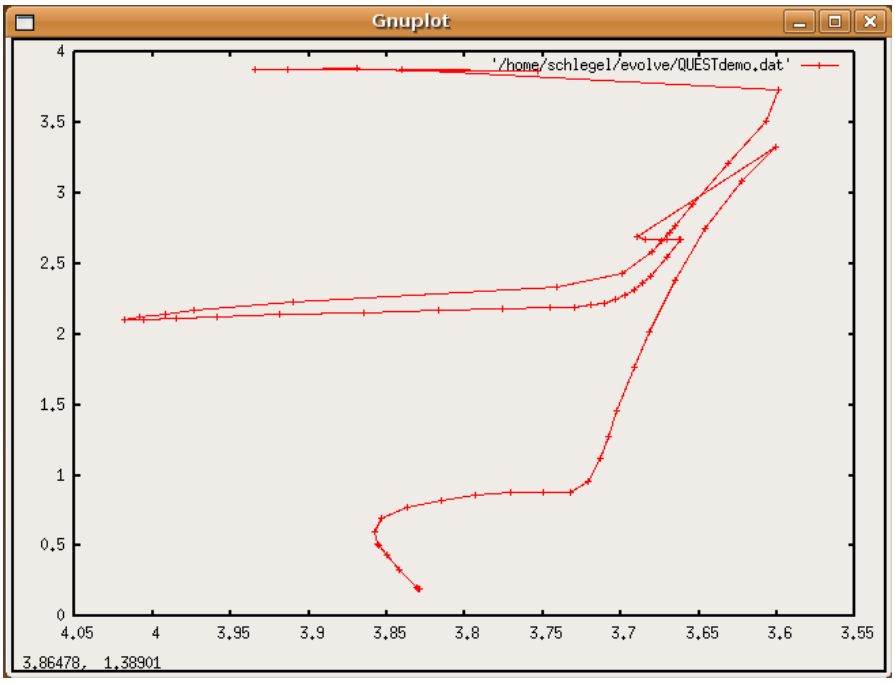


Fig. 3: Evolution Path (M=1.0, Z=0.2)

In addition to this, the application also has to provide active feedback to the user about what they are doing. To do this the program has a part of its display known as the "graph preview area." As the user is selecting what they want to plot this area actively updates showing the points to be plotted. Obviously it is not as sophisticated as a tool like GNUPlot, but it lets the user see what they are doing before they spend the time generating a GNUPlot script, which the program also automates.

VI. Future Work

The system presented here is still in its infancy but already could prove useful in perhaps an introductory astronomy or astrophysics course as an instructional tool. That said though, there are much more ambitious goals for this project as time goes on. These range from simple additions to the plotting abilities, to much more complex analysis of the models of stars as they evolve.

The first of these features I would like to see added is the ability to note important points along the evolutionary curve much like those in Fig. 2. This would increase the utility of the diagrams for physicists interested in a specific phase of evolution specifically.

The much larger and more significant long term goal for the project is examining the model of the star at points along the evolutionary path. EZ allows the user to specify in the

design of an experiment at what points they would like to save model data of the star, but this function is not currently used by this application.

This would obviously allow us to create much more advanced plots. An example of one such plot which one might generate is a plot of log ($[M - M_R]/M$) *vs T*. Specifically of interest in this plot is the photosphere (denoted by the black dots in Fig. 4) and determining its interaction with the hydrogen ionization front as the star evolves.

The real challenge here will not be the actual code generation itself I think, it will be providing a consistent and easy to use interface. The more functions and features that are added that don't necessarily fit to the same purpose, the more cluttered and complex the interface will be come.

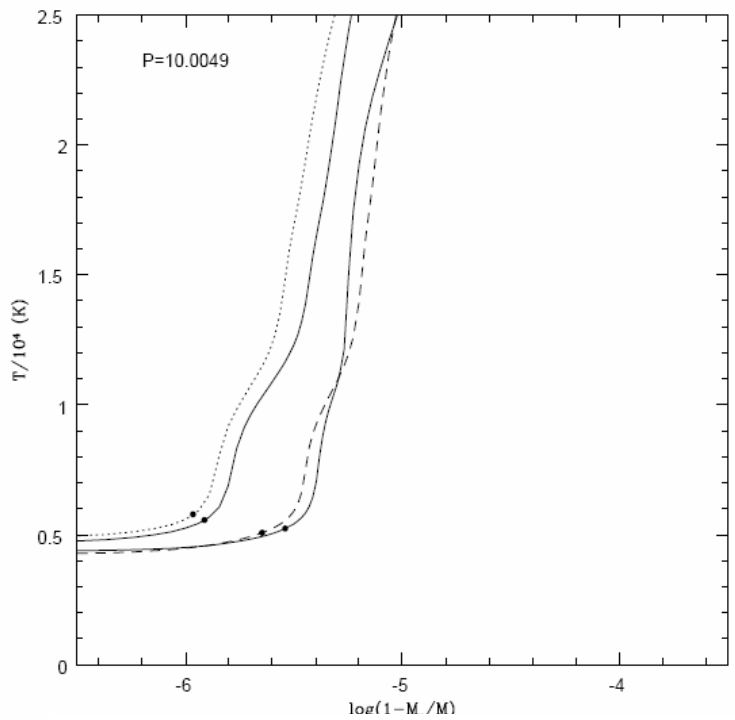


Fig. 4: Analysis of Mass at a Radius (Kanbur, Ngeow, & Buchler, (n.d.))

VII. Conclusion

Scientific research tools have a reputation for being extremely difficult to use. This learning curve makes it very difficult for an undergraduate student who perhaps does not have the UNIX background to use these tools. In turn, most research tasks are given to a graduate student with more experiance. Text based systems do have their appeal for power users, but for the casual user or for a student who is interested in only a certain function of the software a user interface is a much nicer presentation.

While still in its early form this project is already a useful tool for generating evolution paths of stars very quickly. I hope that this project has a long life ahead of it and that it will inspire others to create easy to use interfaces for research software. The amount of development time spent creating something like this will in the end save much more time for the scientific community as a whole and produce better researchers earlier in their college career.

VIII. References

Brinkworth, C., & Thomas, C. (2001). *Stars - An Overview*. Retrieved April 20, 2008, from http://www.le.ac.uk/ph/faulkes/web/stars/o_st_overview.html

Kanbur, S. M., Ngeow, C.-C., & Buchler, J. R. (n.d.). Period-colour and amplitude-colour relations in classical Cepheid variables II: the Galactic Cepheid models.

Paxton, B. (2007). *EZ Stellar Evolution*. Retrieved April 20, 2008, from http://www.kitp.ucsb.edu/~paxton/EZ-intro.html

DESKTOP RELATIONAL SEARCH OR ONE WAY AI COULD SAVE US FROM OUR FILES

D. R. Schlegel
Computer Science Department

Herein a new system for the automatic organization of user files based on metadata is presented. The system is based on a new derivative of Kohonen's Self-Organizing Map algorithm which is also discussed in detail. The system overall is one which aims to provide an end to end user experience for their file and metadata management.

I. Introduction

Consider the chore of searching for files on your computer. Consider the usual metaphors for search. You go to your computer, realize you've lost some document, open up a search tool, type in some sort of search string and wait as the computer finds the document you are looking for. What if you don't remember the file name? What if all you remember is that you worked on it with your roommate in college? With today's search paradigm you are out of luck.

The goal of this paper is to outline a new system which has been created to solve this kind of problem. This system is not based on searching for file names, it is based on the looser and less well defined search for relationships. Now imagine you go to your computer and type in a search string such as "Paper with Alex" and the system parses that and searches for it in some sort of relational store of your data. You would be returned a lot of things having to do with Alex himself, papers of his, pictures of him, etc… but tangentially related visibly could be a collection of things which you did together. This is the grand vision of search – and while the work shown here doesn't quite reach that goal, it does get us much closer.

Herein I will chronicle my attempt at a search system such as this. First discussed will be the problem with searching using today's file organization techniques and the concept of metadata. Next I will present the method I have used to solve this problem and its origins. Finally the search algorithm will be considered in context as part of a larger system from the point of view a user.

II. File Organization Overview

We are all familiar with today's methods of storing files on our computers. Our files are stored on drives, the drives are split into a folder hierarchy, and files are stored throughout this hierarchy. When this paradigm was developed disk size was measured in kilobytes, not

gigabytes or terabytes. In those days we used to label our disks with what specifically they contained. The main organization difficulty was losing the physical disks themselves, not losing our files.

Today our operating systems try to recreate this paradigm through the use of special folders like “My Documents” in Windows XP. Windows Vista takes this one step further with folders such as “Pictures” and the concept of stacks which are actively updated with all files meeting a certain criteria. These files are just deposited into these folders willy-nilly with no attempt to organize by content. Users don’t like to spend their time actively organizing their files, so there must be another way to find what they are looking for. Some applications such as Windows Media Player will manage your music folders for you, creating new ones when you rip CDs, but this does not suit every user. We are all different and like to access data in different ways.

Enter search. Since nearly the beginning of computing we have had search systems that would find something based on its file name or type. In the old days of DOS filenames using the “8.3” format (8 letters for the file name, and three for the extension) this made practically no sense. UNIX traditionally has fared better due to less strict limitations for file names, but file name doesn’t always convey what a file is really about. These search systems would manually traverse a directory tree looking for matches. This requires a complete disk traversal which is very costly.

Today we enjoy desktop search systems from Google, Microsoft, and Apple which catalog our files and even our e-mail in a centralized database. Desktop Search tries to emulate the internet search experience on the desktop. When we search on the internet we navigate to a search engine, enter our query and are returned a list of relevant pages from a pool of billions of documents. Therein lies the problem as yet unsolved – documents are not in standard easily searchable formats (a la HTML), and on the computers of today documents are no longer the principal file format.

Lately we seem to have a bit of an infatuation with the ease of internet search engines such as Google. When we search on the internet we look for a topic based on a search string, click on the results, and see if they contain what we are looking for. This paradigm does not easily extend to the desktop - search for our own files is a much more difficult job. We have to search for image and mp3 files directly without the benefit of related data that could be stored on a web page. Therefore extending this “search-and-browse” paradigm to the desktop cannot succeed.

Google has tried to change this internet search paradigm by adding video and image search to their engine, but just as on the desktop this requires a lot of manual tagging work, and the results are often faulty (Could Google News be Sued for Libel?, n.d.). They have spent countless man-hours on this technology for these two file types and it is useful but as it currently stands is not transferable to the desktop, and of course it only represents a small subset of files on a user’s machine anyway.

So how to do we fix this? The amount of files on our computers is growing all the time. Combine this with the laws of entropy and the ability of any organization system we devise becomes limited. Therefore we must seek out a new paradigm, a new, automatic, strategy

for organizing data based on what the data really means, based on the things we as users would use to organize our data.

III. Metadata

Contained within our data, without our even knowing it, is a resource just waiting to be tapped. In some cases this data is already there just waiting for us to look at it. In other cases it requires a bit more work to get at, but is still helpful. This resource I refer to of course is metadata. Loware defines metadata as "...our knowledge of data that we have interpreted as information in a particular decision-making situation (Laware, 2005)." So what really is metadata then? More simply, it is knowledge about our data.

When we talk about knowledge about data it follows that this knowledge needs to at some point either be explicitly defined or automatically generated. In the case of IDv3 tags for mp3 files it is explicitly defined by someone, and then propagated using a tool over the internet to our media players when we rip our CDs or sync our music with an external player. For a textual document on the other hand, the metadata is generally not going to be explicitly defined and requires some sort of a generation tool.

As time goes on it will be more important to discover ways to extract important features from our documents – whether it be image tagging as in social network sites like Facebook, or in rhythmic identification in our music files (Tenbergen, 2007). Once this new metadata is extracted there will be the need to organize it in a relational way, in a way which makes sense to the user.

If metadata is determined properly it should describe the data it represents. This highly filtered form of the data is much easier for our computers to process and compare versus whole documents of unknown type. There have been many attempts to solve this problem. On the web a search engine can employ algorithms to generate "snippets" of a web site. These snippets are essentially an abstract or the main idea of the article. Microsoft Word allows another way to auto-generate the idea of a document through its summary feature. So for at least a few document formats we have already started to solve the problem of generating the metadata. There of course is more work needed in this area but after we have the metadata putting it to use is all that is left.

So far we have discussed that we have a problem with organizing our data because we have so much of it. We have also discussed the fact that there is knowledge about this data we have available to us. Combining this knowledge with some maturing AI techniques a new, more natural paradigm can be developed. We have to combine metadata, search, and artificial intelligence techniques to help create a new paradigm for search.

IV. Self-Organizing Map

The Self Organizing Map, developed by Teuvo Kohonen, is an artificial neural network (ANN) based technique for organizing data in an unsupervised manner. This means that there is no user interaction while the algorithm runs, denoting a “black-box” type algorithm. It is generally represented by a two dimensional array of nodes (which in ANNs

have the biological counterpart neurons). These nodes contain some numeric data which represents what is being organized.The result of the map is similar to multidimensional scaling techniques which aim to recover the “underlying structure among stimuli which is ‘hidden’ in the data (Schiffman, Reynolds, and Young, 1981)."

When the algorithm is initially started the map is created using the specified size and is initialized with random data. This random data is stored in what is called a *weight vector*, and is of the same dimensionality of the input to the map.

Once the training of the map begins, input vectors in a training set are exposed to the network. When this happens, each node has a level of activation and that which is activated the most is called the Best Matching Unit. Generally to determine this level of activation Euclidean distance is used, and the one which generates the smallest distance is the BMU. This unit and its surrounding *neighborhood* nodes are then made to be closer to the input vector (Zavrel, n.d.) (How much closer is determined by a Gaussian function and a learning rate usually). The neighborhood is defined by a radius which shrinks over time.

This process is iterative, happening hundreds or thousands of times until the network is converged and smooth. This happens since over time the neighborhood shrinks. At the start of the run it may be as wide as the lattice, but by the end it is usually very small in relation to the map size. Some SOM algorithms split this into two phases. During the first, longer, phase the neighborhood shrinks very rapidly. During the second phase which lasts much longer, the neighborhood shrinks much more slowly until it is finally converged (Kaski n.d.). Kohonen mentions in his book on SOMs that since there are generally far less training data than numbers of iterations of training necessary "the samples may be applied cyclically or in a randomly permuted order, or picked up at random from the basic set (so-called *bootstrap learning*). It has turned out in practice that ordered cyclic application is not noticeably worse than the other mathematically better justifiable methods (Kohonen, 2001).”

The speed of convergence of the network is very important. If the network converges too quickly it will end up mosaic-like, with blocks of related data but not very smooth. If it converges too slowly then each node in the network will contain largely the same data. Either of these cases are not optimal and will only result in results which are not representative of the data.

The typical toy example for self organizing maps is to categorize colors based on an array containing their RGB values. This case is very simple since the color can already be represented with numbers without any extra abstraction and there are a finite number of values which fall into definite categories. The results from these maps are very successful and have been shown to work with any random initialization of the map (Honkela et al., 1998).

In addition to toy examples, there have been several real world applications of SOMs. SOMs have been used extensively in Biology to sort organisms into their respective families. Kohonen himself has spent many years working to organize large document collections using what is called the WEBSOM method (Honkela, et al., 1998). Other researchers have worked to create virtual library systems which involve creating “shelves” of similar data to be browed by a user on the web (Merkel & Rauber, 1999).

All is not rosy for the SOM though; there are a great deal of real world limitations when it comes to the algorithm itself. The self organizing map depends on a very structured model for feature vectors. It is easily applicable to systems where feature vectors are easily represented through numerical values. For textual values though, things are more difficult. For example, Kohonen mentions in his SOM book that most ways for organizing documents involves to some extent a word histogram. Methods are employed to simplify it, but it is still a histogram of thousands of words, with lots of empty space, and breaking any biological representation the SOM might have.

V. Abstract Self-Organizing Map

The ASOM or Abstract Self Organizing Map is different than a normal SOM in how it handles its features. The method for comparing feature vectors is not as obvious as it is with a word histogram or some other method since the input vectors may be only tangentially related (or not at all) to the training set. To deal with this the actual computation comparing vectors is done outside of the map itself where a weight showing the correlation between the input and the node on the map is generated. This abstraction is the reason for the algorithm being called the Abstract Self Organizing Map.

The ASOM (Fig. 1) has been designed to work with any set of data. The only requirement is that the author of that specific implementation devise a way to make the data comparable. A node in a self organizing map contains the numeric weights on which the algorithm will operate. In an ASOM the node is basically a container for the *Feature Vector* and its coordinates on the map.

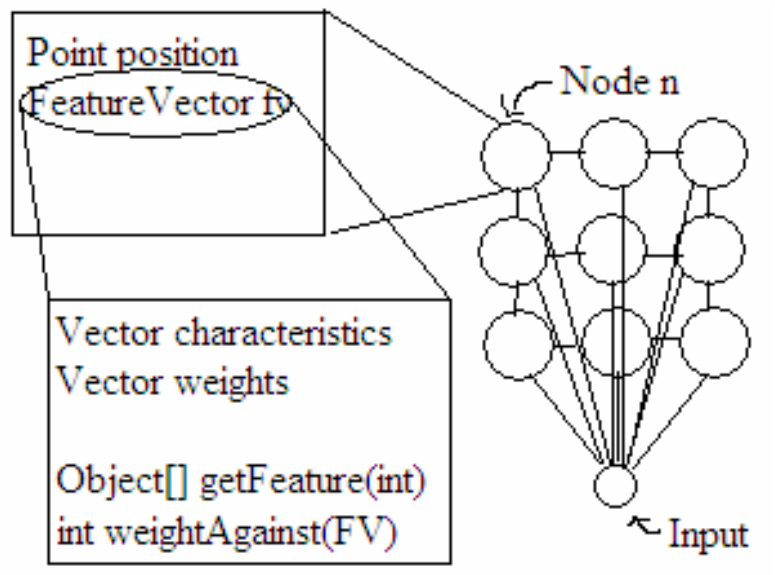


Fig. 1: ASOM Algorithm

The feature vector in an ASOM is really two vectors which are side by side. The first of these is a vector of *characteristics* and the second is a vector of *weights*. The feature vector also contains some method of comparing itself to another feature vector and returning a single value, this is called the *relative weight*. This relative weight function replaces the Euclidean distance method for determining similarity in Kohonen's SOM algorithm. It is important to mention though at this point, that an argument could be made against this algorithm being a neural network. It contains many of the same features of a neural network, but ideologically it breaks from the tradition of neural networks by requiring the weights to be computed and not stored concretely. Also, in the same way the SOM algorithm does, this algorithm goes against the highly interconnected view of a neural network. Indeed without an input vector, no connections exist between the nodes, which obviously is not biologically accurate and is not perhaps conducive to maintaining its status as a neural network based algorithm.

A characteristic is a defining aspect of whatever the implementer is trying to organize. It can be of any data type as long as the implementation can compare that type for likeness. The weights are a numeric value which say how important that specific characteristic is. These values are unbounded and should be tweaked to fit the needs of the specific implementation. Finally the relative weight is the outcome of the comparison. In general higher values denote a higher level of likeness, and lower values denote the opposite. This too is unbounded and should be taken as a relative scale in comparison to the entire map and never on its own.

Since the ASOM shares its main architecture with Kohonen's feature map, in many of the same ways that SOMs fail to represent biology, so does the ASOM. Even though this is the case, the ASOM may be more akin to how we consciously think about information. The SOM fails in mimicking our thought processes by forcing everything to be represented through weights with no real data attached to them. The ASOM though creates relative weights by actually comparing the data and not just a weight associated with it. This is more similar to how we draw connections between information cognitively, but is perhaps more an issue of semantics than an actual likeness.

Altohough this seems to be an improvement conceptually, there are still issues. It appears from initial testing that the rate of learning must be very carefully tuned to the amount of iterations of learning (and consequently the radius of the neighborhood – this is especially important in a small network). If either of these are disproportionate then the system over-converges, meaning that each node becomes nearly identical to every other. Obviously search breaks down at this point and results are nearly random.

V.a *Implementation*

In my test of the map I chose to use 100 Word 2007 documents which I had local on my machine as a sort of real world test. Based on the metric of ¼ the number of nodes as documents, the result was a 5x5 map. Looking at examples from other SOM work this seems to be a fairly natural value – Kohonen used 1/6 in his WEBSOM experiment but there is no tried and true method for determining this value. The advantage though to a

larger map is that there is more room for unrelated documents to separate themselves. A 5x5 map is fairly useful in analyzing whether or not the map works since there are several inner nodes and several edge cases.

After creating the map the next step was to define what the feature vectors were going to contain. For my feature vectors I extracted all of the words which are capitalized and aren't extremely common sentence starters (such as For, It, The, Then, etc…) and set the weight for each of these characteristics to the number of occurrences of that word. This is obviously a very crude method but for the purpose of testing the maps ability to organize data it is sufficient.

In order to compare these feature vectors the algorithm first finds all of the characteristics which are in common between the two FVs being compared. The weights of each in common characteristic are multiplied together and these values are summed. This determines the relative weight of these two feature vectors.

The implementation which I designed uses a Gaussian function for the purpose of converging neighbor nodes with the BMU. The Gaussian function is useful for this since it allows for the natural graded-ness of the map which we are seeking. The learning rate is two, and the terminal learning rate is 1/2. I ran the map for 100 iterations. Testing at different values determined this was optimal for this size map, but there is no known algorithm for determining what this value should be. At values much smaller than this (~50) the map would not have any gradient, and at values much higher (~300) the maps gradient was too large (as expected and discussed earlier).

V.b. *Algorithm Analysis*

To test the ability of the ASOM to organize data I used a technique which is used often with traditional SOMs. This method tries to determine whether the topology of the map has converged at the end of the training. This involves using "input samples to determine how continuous the mapping from the input space to the map grid is. If the two reference vectors that are closest to a given data item are not neighbors on the map lattice, the mapping is not continuous near that item(Kaski, n.d)." Testing that two nodes are neighbors is not enough though – the map should be smooth. If the map has a large spike when an input is presented and nearly no activation around it, then the map has not been trained over enough iterations.

To do this I applied some sample search strings to the converged map. In each case the node for which the BMU was the document I was looking for was found, and one node removed from it showed lower activation levels. Depending on the search string usually two nodes removed produced an activation level of zero which is expected for a map of this size. A sample of this is shown in Table 1. This has been tried with maps smaller than this, and the results seem to scale well from the limited tests I have performed thus far.

0	0	0	0	0
10	60	10	0	0
15	10	0	0	0
10	10	0	0	0
0	0	0	0	0

Table 1: A search gradient

Neural networks have a reputation of not being particularly speedy. The combined time of training the map (including getting the generated metadata from persistant storage) averaged around three minutes on an older Pentium 4M laptop. Previous tests have shown that smaller maps take far less time to compute. There is a nearly exponential increase in time as the training set grows larger, but there are several techniques for increasing the speed of SOM-like algorithms which have not been explored yet. In comparison, standard desktop search systems can take many hours or days to catalog just the document data on someone's PC.

The ASOM algorithm and method provide a great deal of needed improvement over the original SOM algorithm. The major improvements are reiterated below in a more easily digestible tabular form.

Abstract SOM	"Classic" SOM
Flexible design – Doesn't require changing the algorithm to sort different data types	The SOM is crafted and optimized for the target data type as needed
Can sort multiple types of data together as long as a method is defined to make them mutually comparable	Maps are created with a specific purpose in mind, there are few (if any) "in general" type solutions.
Uses feature vectors to store comparable data. Weights are relative, on no particular scale.	Does not have a defined method, sparse word histogram arrays are often used for text. Must be made to correspond to a 0-1 weight value

VI. The User Perspective

Making search relational is not enough, the user needs an intuitive way to interact with their files. I have thought hard about what types of applications this metadata-driven system might allow us to take advantage that we cannot today. One example may be a system to tag images. As the user adds tags to an image they influence the way search works with those files. This adds an extra dimension to the idea of search – essentially giving the user control over their search results. It is conceivable that in time this could be expanded with functionality to tag a group of related images together, or to propagate changes to "like" images. This flexibility will be appreciated by the user later when they search for their files.

This allows a user to impact search before they do it, but all that we have done here is so that the user can relationally search for their files, and an interface is needed for that. In the SOM Explorer interface (Fig. 2) there is no need for an "advanced search" interface – the user can simply type in what they are looking for and hit search. Since the SOM has already been created and has converged the search results are returned instantly. The nodes in the center of the image are very closely related to the search string, and the top ranked feature in that node is displayed. As you look further out in the image the items are less related. This is simply a view of the underlying neural network which the SOM algorithm has already ordered in this way for us.

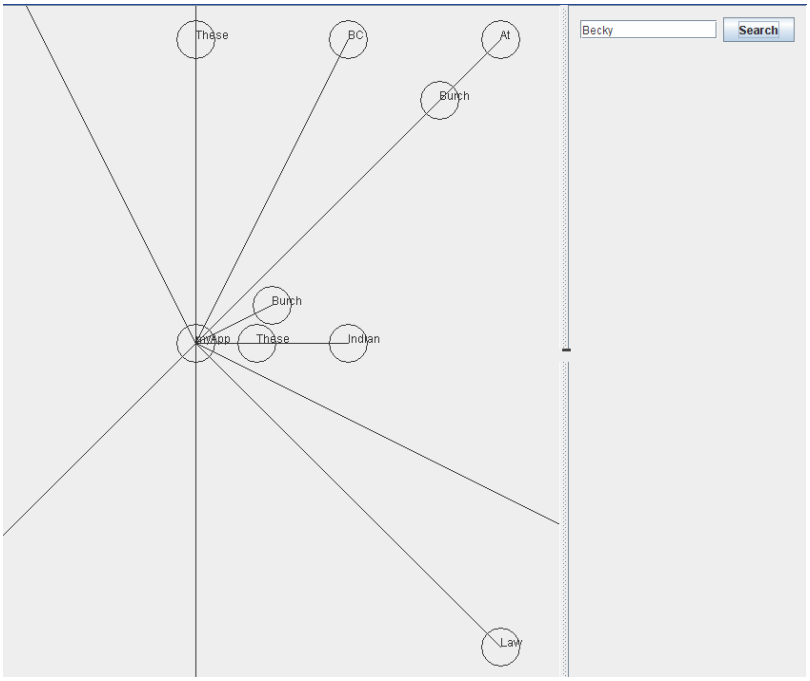


Fig. 2: SOM Explorer

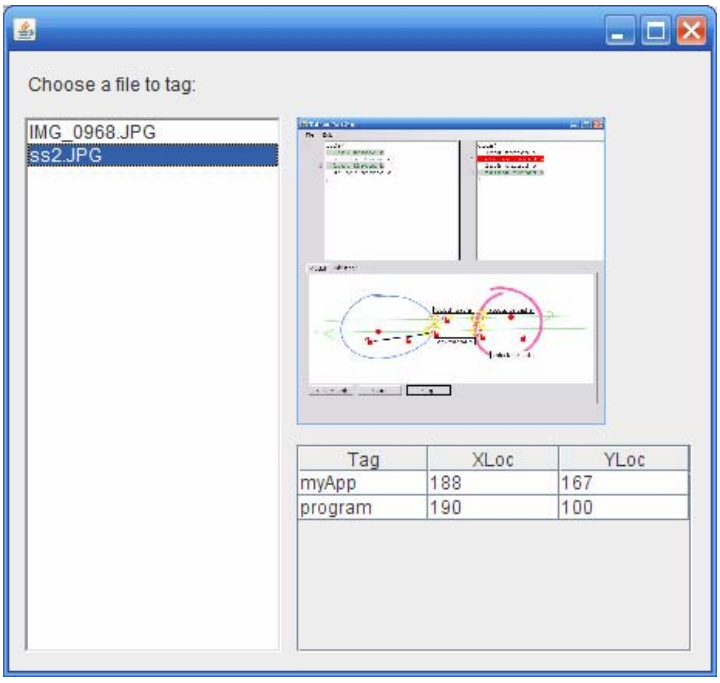


Fig. 3: Image Tagger

VII. Thoughts for Future Work

The list of further work I would like to do on this project easily is far more expansive than any of that which has been completed thus far. These changes range from finding methods to enhance the speed of convergence to finding better results through more useful metadata generation techniques.

One way to speed up the convergence could be to take advantage of multicore processors. Kohonen mentions some examples in which he uses many small maps of data to estimate the layout of the larger map. These smaller maps could each be computed separately in dedicated processors or processor cores which can then be combined and calculated as we are now.

In addition to the improvements in speed, there is a need for better data to test on. The feature vectors from the documents are formed in a very rudimentary way (choosing all of the “important” capitalized words). I can see several simple ways of improving this. Chopping off a words prefixes and suffixes would help the weighting be more accurate (Etzioni and Zamir, 1998), and grouping words that are capitalized that appear consecutively may also help since they are likely related.

I would also like to explore how we determine characteristics of documents more carefully. In addition to using proper nouns, I think the structure of documents should be considered. For example, the size of text, or whether or not text is in a bulleted list can provide a sense of importance to words or phrases. Using this kind of analysis we could not only compare documents with similar important words or phrases, but also documents with a similar structural format.

On top of these improvements more tests will be needed, on a larger data set. Obviously a 5x5 network is not appropriate to determine if an algorithm is a success or failure. Just because it works at a small scale doesn’t mean it will not break down or become far too complex at larger sizes. I aim next to test on a network at least an order of magnitude larger, with perhaps a thousand documents and two hundred nodes.

As I’ve mentioned multiple times throughout this paper, documents do not define all or perhaps even a large part of the files on a user’s computer. The project contains an early tagging system for images now, but there is much of room for improvement. Current AI techniques for facial recognition and comparison could be used to automatically tag images. In addition to images, work has been done to categorize music not only by tag but by the actual musical patterns. This type of system could be included into this system with relative ease.

VIII. Conclusion

Organization of user data has over the last few years become an exponentially more difficult problem as the ubiquity of the internet and downloading of files becomes obvious, and several attempts have been made to try to solve the problem. These solutions though are just extensions of research that has already been done for search on the web and do not translate particularly well to the desktop paradigm. There have been novel ideas such as

Microsoft's proposed relational file store (WinFS), but it was ultimately cancelled in 2006 (Clark, 2006).

In order to succeed researchers and companies alike need to begin thinking about metadata as a viable method for ascertaining the meaning of user data. There have been many problems trying to use metadata in the past because it is largely non-standard. We can no longer ignore what is in reality self evident though, we need to bite the bullet and write code for these non-standard constructs or work together to ensure our data has proper metadata and that it is properly used.

Algorithms such as the SOM have shown promise for document organization, but have been quite clumsy and slow. Over time our AI techniques will improve and I think therein lies the next paradigm of search. AI has consistently gotten a bad rap as consisting of a collection of very complex programs which don't work, but slowly AI based applications are making their way to the desktop. The solution I have presented here is just one example of this evolution. Although I don't think there will ever be a permanent solution for our user data woes, I hope the technique presented here will help to fuel the next generation research.

IX. References

Clark, Q. (2007). *Update to the Update – WinFS Team Blog.* June 26, 2006. http://blogs.msdn.com/winfs/archive/2006/06/26/648075.aspx (accessed November 13, 2007).

Could Google News be Sued for Libel? http://www.cyberjournalist.net/news/001346.php (accessed November 3, 2007).

Etzioni, O. and Zamir, O. (1998). "Web document clustering: a feasibility demonstration." *Proceedings of the 21st annual international ACM SIGIR conference on research and devoeopment in information retrieval*, 46-54.

Honkela, T, Kaski, S., Kohonen, T, and Lagus, K. (1998). "WEBSOM - Self-organizing maps of document collections." *Neurocomputing* 21, 101-117.

Kaski, S. (n.d.). (*Data Exploration Using Self-Organizing Maps.* PhD Thesis, Helsinki University.

Kohonen, T. (2001). *Self-Organizing Maps.* Springer.

Laware, G. W. (2005). "Metadata Management: A Requirement for Web Warehousing and Knowledge Management." In *Web Mining: Applications and Techniques*, edited by Anthony Scime.

Merkel, D., and Rauber. A. (1999). "The SOMLib Digital Library System." *Third uropean Conference on Research and Advanced Technology for Digital Libraries*, 323-342.

Schiffman, S. S., Reynolds, M. L., and Young, F. W. (1981). *Introduction to Multidimensional Scaling : Theory, Methods, and Applications.* Academic Press.

Tenbergen, B. (2007). "Similarity Clustering of Music Files According to User Preference in Proceedings." *MICIA*, 182-192.

Zavrel, J. . *Neural Information Retrieval: An Experimental Study of Clustering and Browsing Document Collections with Neural Networks.* PhD Thesis, University of Amsterdam.

SKETCH UML:
A TABLET PC-BASED E-LEARNING TOOL FOR UML SYNTAX USING A MINIMALISTIC INTERFACE

Bastian Tenbergen, Colleen Grieshaber, Lisa Lazzaro, & Rick Buck
Human-Computer Interaction MA Program

The Interactive Learning Technology Laboratory of the State University of New York at Oswego has developed a Tablet PC-based, ink-input enabled learning utility that will aid students in learning the UML syntax as well as teachers in critiquing UML diagrams designed by students. This is intended to enhance the learning and teaching experience regarding UML using non-traditional interfaces. As a merely ink-based application, SketchUML is a novel approach in pen-based user interfaces. As such, the usability of it has not been assessed. This study addresses the question of usability of non-hybrid, ink-only, non-traditional interfaces by conducting usability tests. The collected data shows that ink-only interfaces facilitate learning in non-expert users and do not hinder the usability and learnability of these interfaces and allow for an enhanced learning experience.

I. Introduction

SketchUML was conceived as an interactive learning utility that allows for learning the syntax of the UML modeling language using Tablet PCs. UML – or Unified Markup Language – is a general-purpose standardized specification modeling language, used in Software Engineering to model object diagrams and abstract specifications of software systems. UML has a wide range of symbols and syntactical rules according to which valid diagrams are designed (Fowler, 2003). Learning the entire UML syntax can be tedious and time consuming, due to its error-driven nature.

Computer-aided UML learning has a number of advantages over ordinary approaches using pen and paper. The fact that it allows for immediate feedback to the user if a certain UML symbol is incorrect in the current context is an invaluable tool that cannot be accomplished by pen and paper approaches. Furthermore, computer-aided UML design enables the user to modify an existing diagram quickly without the need to redesign a specific section or even the entire diagram. Another advantage that is provided by incorporating software into UML diagram design is that from an existing class diagram, code can be exported, resulting in class and method stubs in a programming language of choice that obey to the class hierarchy depicted in the diagram and only have to be filled with content.

However, using software to design UML schemes has a negative impact on the learning curve of UML. When a student is asked to learn the UML syntax by using a certain software, this generally results in the fact that the software has to be learned beforehand. It is very uncommon for a user to know how to use a software that performs a specific task without having in-depth knowledge of the task itself. Many

UML softwares that are currently available do not tackle this problem. They are not designed to facilitate the learning experience of a user to learn the syntax of UML. Instead, these programs are designed to incorporate more and more features, which results in amazingly complex software that is hard to use for novices.

This problem is tackled by SketchUML. Previous publications by Qiu (2007) have shown that this software is a very usable and versatile tool for learning and designing UML diagrams. The product has been massively augmented since Qiu (2007) and incorporates many new features. SketchUML is a completely user centered software. It is designed to facilitate the learning of computer-aided UML design, so that it is virtually entirely naturalistic. The interface is designed with simplicity being the driving force, offering only a paper-like white canvas and a menu bar. Users can simply open the application and start drawing UML diagrams, as if they would use pen and paper. Designed for Tablet PCs, SketchUML accepts input using the Tablet PC's pen or stylus and converts the drawn ink into UML components, if the ink input corresponds to a valid symbol in UML. This allows for immediate feedback to the user when an ink gesture is not valid UML syntax component. The user is thereby forced to design UML diagrams accurately which aids in learning the UML syntax. Since SketchUML is designed to be as versatile as possible, it allows to fully edit the entire diagram at any point in the design process, just as users would expect to use pen and paper. Also, SketchUML is designed to be robust against different drawing styles of ink gestures that represent a UML symbol – for instance, it does not matter if a user draws a square by drawing one side at a time, lifting the pen tip every time a stroke has been performed, or if the user draws a square by using only one stroke from start to end, with the start and end point being identical.

Previous studies have shown that using Tablet PC technology is an effective tool for teaching and collaborative learning (Bull et. al., 2004; Berque et. al., 2004; Simon et al., 2004). SketchUML is intended to allow for an augmented teaching experience for faculty teaching UML to students. This is achieved by the ability to import a student's diagram that has previously been saved to hard disk, and critiquing it using the Teacher Mode. Critiquing the students' diagrams allows the teachers to give feedback on the student's work online, without the need of paper. Also, this reduces the amount of time needed to correct students' solutions to a given assignment. More significantly, it aids the teacher in a way that there is no need for an actual teacher's model solution. It is conceivable that the system can create a model solution or a solution template by analyzing a number of student solutions that have been critiqued by the teacher.

Overall speaking, SketchUML offers a good learning platform for students and teachers with regard to UML diagram design and syntax learning. Its wide variety of user centered features and the incorporated design metaphors create a well developed conceptual model that allows users to learn the syntax of UML rather than learning the software that is supposed to aid them to do so first. This paper discusses the research that has been done to augment the usability of SketchUML. A number of usability tests have been conducted in order to assess the pitfalls of ink-only, non-traditional interfaces and to understand how these types of interfaces can aid in e-Learning. The results of these experiments directly influenced the understanding of how learners interact with learning software and how the proposed software SketchUML needs to be augmented in order to provide for a decreased learning curve and a more naturalistic approach in Advanced Learning Technologies.

II. Methods

The main purpose of the tests was to assess the limitations of ink-enabled user interfaces. SketchUML is not only an e-learning tool that facilitates learning of the UML syntax, it also a novel approach for only ink-input accepting computer interfaces. The testing session was intended to identify the limitations of ink-only input and evaluate the usability of non-hybrid interfaces (rather than hybrid ink & point-and-click-based interfaces) and the effectiveness of SketchUML as an e-Learning tool for UML diagrams.

II.a. Participants

A number of undergraduate students (7 seniors, 3 juniors and 1 sophomore) that were enrolled in a software engineering class have been tested in this study. The students were all novices to UML or had intermediate knowledge, but no student was an UML expert. The participants' experience with computers and user interfaces ranged from 5 to 15 years (M = 12,8, SD = 2.5). No student received any reward for participating in the study and their participation was voluntary. All but two students were new to Tablet PCs and the SketchUML software.

II.b. Apparatus

The environment was a classroom setting with other students, desks and typical supplies. The lighting was also that of a typical classroom. The equipment being used was 6 Hewlett Packard Compaq TC4200 Tablet PC Laptops, with 1.7GHz Intel Centrino Processors, 512MB Ram, running Windows XP Tablet PC Edition. The test software was the latest stable release of SketchUML. For testing, the laptops have been switched to Tablet Mode and participants were asked to use the Pen as the input device.

II.c. Procedure

After the participants signed informed consent forms and filled out a pre-test demographic questionnaire, they were given a short introduction to the interaction with Tablet PCs and the software SketchUML. The introduction consisted of a brief overview of the features of SketchUML and the features of the Tablet PC hardware that was used and was presented using a standard classroom projector. All participants were introduced at the same time so every participant received the same introduction. An effort was made to randomly divide the participants into two groups, so that 6 participants were in group one and 5 participants were in group two. Each participant was given a task sheet with a task description; the task the students had to complete was different for both groups. The members of group one were given a class hierarchy description and were asked to draw the class hierarchy using the Tablet PCs, on which SketchUML was running. Group two was given the same class hierarchy description and a pre-loaded class diagram that was drawn incorrectly and does not meet the hierarchy description. The erroneous parts of the diagram were labeled accordingly (e.g. The class name “wrong class” for a class that must be deleted), with exception of incorrectly placed connectors, as connector labeling support is currently not fully supported in SketchUML. Participants in group two were asked to critique the pre-

loaded diagram using the Teacher Mode that is built into SketchUML. Fig. 1 shows the experimenters ideal solution to the task of the first group. Fig. 2 shows the pre-loaded diagram the participants in the second group were asked to critique and Fig. 3 shows the ideal solution to the task of group two.

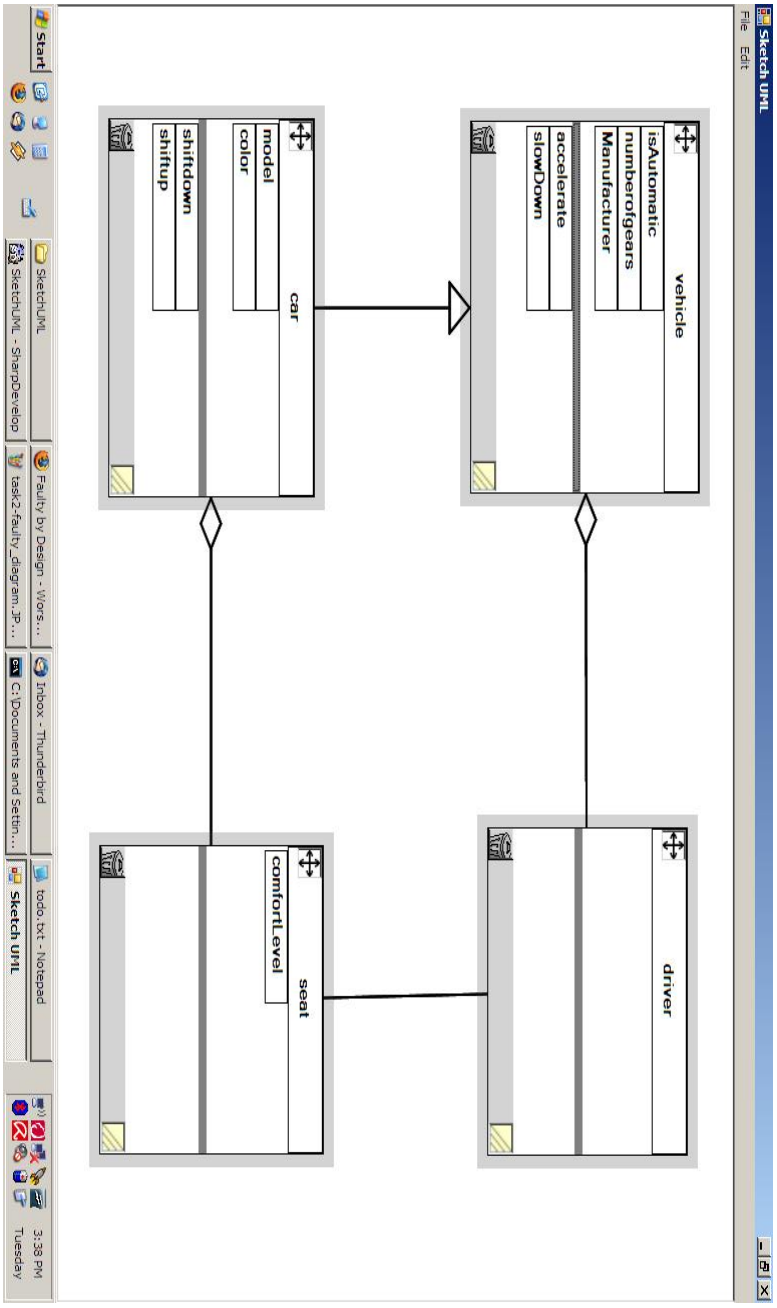


Fig. 1. Ideal Solution to the Task of Group One

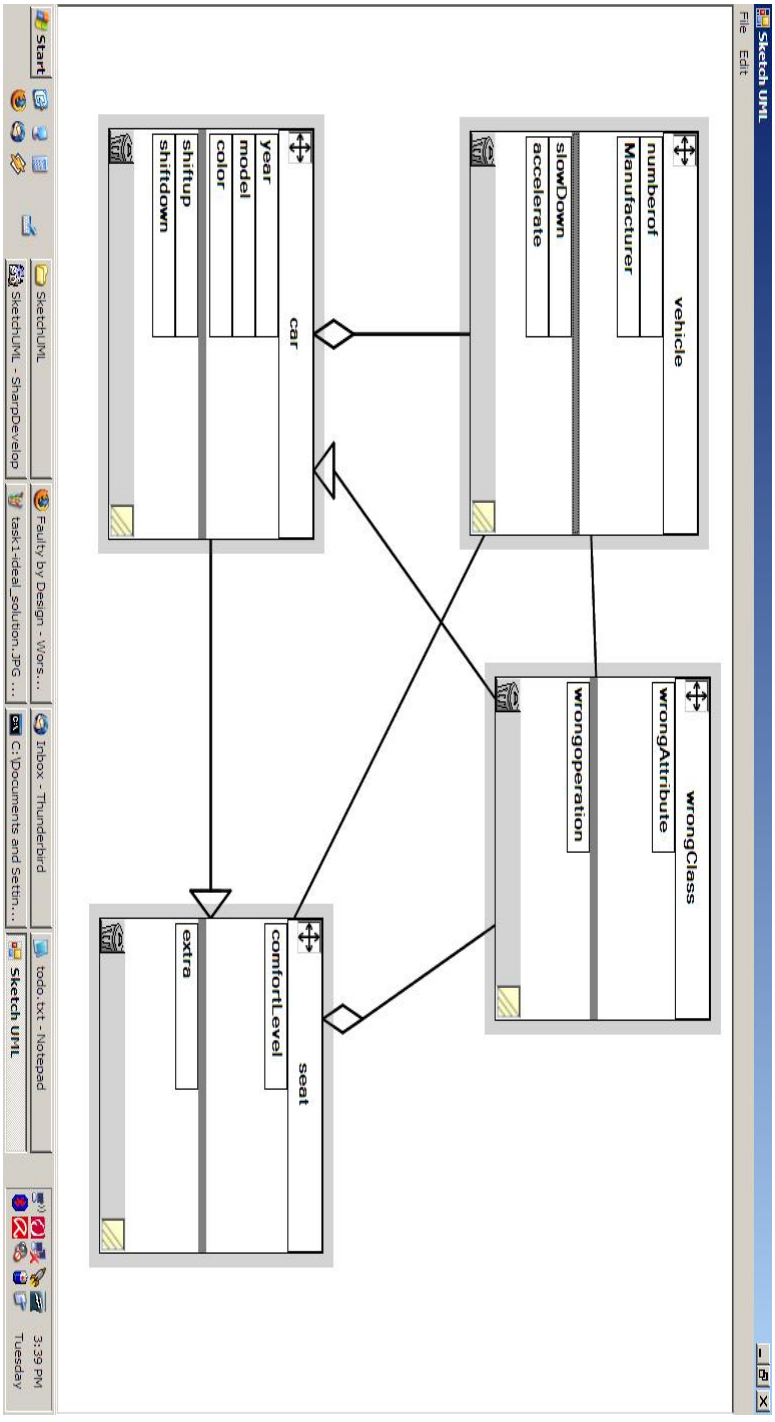


Fig. 2. Pre-Loaded Diagram to the Task of Group Two

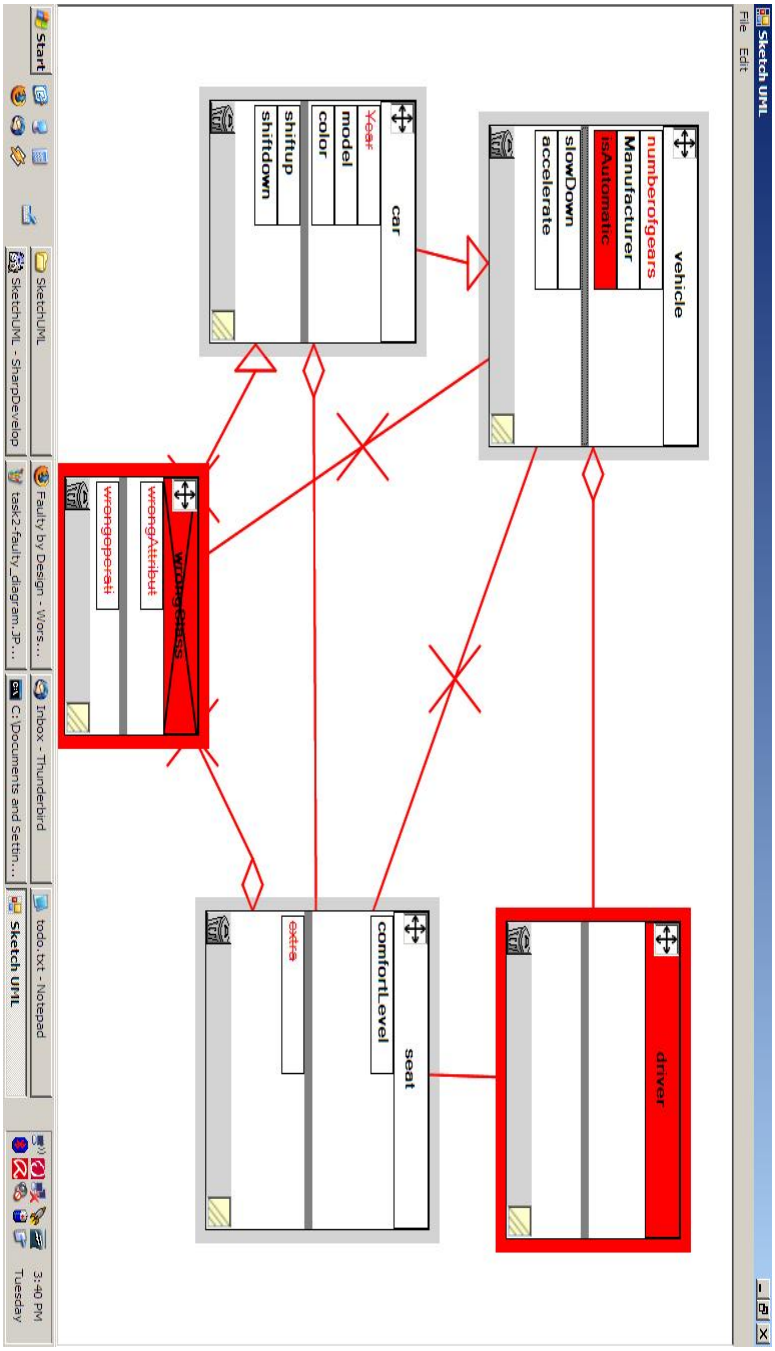


Fig. 3. Ideal Solution to the Task of Group Two. Note, that some elements are shown in a red background in the original program

The tasks were designed so that every feature of SketchUML was used at least once by at least seven participants throughout the experiment.

The experiment was not timed and the participants could modify their diagram or start all over again as often as they wanted. The completed diagram was collected from the participants individually when the participants felt like they were done. The

diagram was stored into the XML format using SketchUML's built-in Save function and later converted into JPEG format for easier analysis.

Each participant was asked to fill out a uniform post-test survey after they completed the individual task. The post-test survey contained questions regarding the subjective quality of SketchUML, ease of use, entertainment aspects and missing features.

The post-experiment data analysis was performed in two parts. The first parts regarded the qualitative answers from the post-test questionnaire. The second part focused on the quantitative performance of the participants, i.e. how well the interface allowed them to fulfill their individual tasks. Since the participants were explained in detail how to use the interface, and since the participants were not unfamiliar with UML, a failure to achieve a certain goal (for instance, a connector could not be drawn or an attribute field was added incorrectly) is a result of a usability issue with the interface, which we seek to uncover in this study. The quantitative analysis could therefore be done by simply measuring the number of mistakes of a participants solution (i.e. the amount of diagram components, like connectors, or operation labels, that weren't drawn correctly in the diagram) and comparing the minimum number of steps needed to create an ideal solution (i.e. the number of gestures, actions and click-events minimally necessary to complete the task). The mistakes were divided into three groups: mistakes made during class symbol manipulation (i.e. when a class symbol was to be created, moved, or deleted), mistakes made during label manipulation (i.e. when a class symbol label or attribute or operation label was to be created, modified or deleted) and mistakes made during connector manipulation (i.e. when a connector like aggregation or generalization was to be created, modified or deleted).

III. Results

In the first condition, i.e. the task that was given to the first group, 9 mistakes were made overall. It took a minimum of 21 steps to complete the task. In total, no mistakes were made by any participant regarding class manipulation. Only one participant failed to correctly manipulate a connector – this participant created an aggregation instead of a generalization. The majority of 8 mistakes were done during label manipulation, with similar likelihood in every participant. These errors were exclusively errors in assignment of correct labels (e.g. two labels “is” and “Automatic” instead of one label “isAutomatic”). Each student made an average of 1.5 mistakes during label manipulation.

In the second condition, i.e. the task that was given to the second group, a total of 7 steps was minimally necessary to complete the task. A total of 16 mistakes were made, which corresponds to 3.2 mistakes per participant. Only two students incorrectly manipulated a class symbol, and a total of 5 mistakes were made during label manipulation. During connector manipulation, a total of 9 mistakes were made.

The qualitative post-test survey shows that the majority of the students ranked the overall quality of SketchUML as “medium to high”, which corresponds to a score of 4 on the 5-point Likert scale that was employed in the survey questionnaire. Open questions regarding the participants' subjective opinion of the program (their likes and dislikes) and their ideas for improvement uniformly contained criticism on

SketchUML's mode of feedback. Students mainly complained about difficulties to correctly spell names of attributes and operations. Positive feedback comprised the ease of creating connectors and class symbols. All students independently reported a low frustration level (despite the fact that they were not all able to complete the task correctly) and a high factor of enjoyment.

IV. Discussion and Future Work

The purpose of the Usability Testing with SketchUML was twofold: Firstly, the quality and usability of the newly developed software and the recently added features or SketchUML were to be assessed. Secondly, this study aims at understanding the advantage of Tablet PCs in non-collaborative educational work and in which way this technology can help to facilitate learning concepts in Software Engineering, exemplified at the UML syntax.

The first group completed a task that was designed to exploit all features of SketchUML's Student Mode, in which students have the ability to sketch UML components and thereby creating UML diagrams naturally in a pen-and-paper resembling minimalistic interface. The overall performance of the participants in this condition was very good, as only a minimal amount of mistakes were made by each student in drawing a diagram. The majority of the errors made were during label manipulation. Labels are deleted by using a scratch-out gesture and created or corrected by handwriting into the specific area. The post-test survey showed that no participants reported trouble when interacting with SketchUML using gestures, but the majority of criticism was directed towards poor recognition of the handwriting. This is consistent with the fact that no errors were made during class symbol manipulation (as this is mainly done by interaction through gestures) and with the fact that all label manipulation mistakes were made when assigning correct content to the label. The post-test survey further shows that the poor performance of the handwriting recognition engine is the mere reason for increased frustration, as it took some participants multiple attempts to correctly create a label. Although this significantly affects the efficiency of interacting with SketchUML, it did not affect the effectiveness, as all participants in the first group were able to create correct diagrams, compared to the target solution. The fact that one participant made one error during connector manipulation is assumed be caused by a lack of expertise with the UML syntax, as the student correctly created a connector – it was just the wrong type. Connector manipulation can therefore be considered both effective and efficient to use, as no student reported difficulties with such. Overall speaking, these findings suggest that the features of the Student Mode are naturalistic, effective, efficient and enjoyable. Improving the poor handwriting recognition performance will reduce the remaining frustration during label manipulation.

The second group was asked to complete a task that exploited all features of the Teacher Mode that differ from the Student Mode. The performance of the participants in this condition was not as high as in the first condition, yet remarkable. As the Teacher Mode gives slightly different visual feedback to certain gestures than the Student Mode, participants seemed to be somewhat confused. When creating a class or a label, it will occur with a red background color, to symbolize that this component was

created by a teacher. This however, was incorrectly interpreted by some students who made mistakes during class symbol and label manipulation. Both errors during class manipulation and three of the five errors during label manipulation occurred when a student intended to create a class or a label and was uncertain about the red background. As a result, the previously correctly created class or label was removed. The other two mistakes during label manipulation were errors due to poor handwriting recognition. Similar in the first condition, the students' handwriting was incorrectly recognized, therefore rendering wrong label names. The majority of mistakes were made during connector manipulation in the second condition. When a connector is created in Teacher Mode, it will appear in red, to symbolize that this connector was created by a teacher. In contrast, when a connector, that was created in Student Mode (therefore appearing in black) is deleted in Teacher Mode, this connector will appear in red, with a red cross in the middle, symbolizing that this connector was removed by a teacher. Just like the problems during class symbol and label manipulation in the second group, this seems to be a problem with the mapping and feedback of the system to user input. Although the participants correctly understood that the red ink color depicts changes made in Teacher Mode, the difference between teacher created connectors (plain red ink) and teacher deleted connectors (red ink with a cross in the middle) does not appear to be obvious. Exclusively all mistakes have been made due to confusion of these two different connector drawing styles.

The post-test questionnaire further shows that SketchUML as an ink-only non-hybrid, non-traditional interface is very capable of satisfying user needs. The ease and effectiveness of the participants' interaction with the software shows that simplistic paper-and-pen resembling interfaces suffice in providing a good interaction with the user. Participants generally did not miss the ability to use the mouse, in fact, preferred using the pen as a pointing device over the traditional pointing devices. Also, the use of a keyboard was not missed, except for when handwriting recognition was not performed correctly. As a suggestion, it was mentioned that a virtual keyboard can help to manually correct incorrectly recognized handwriting. These results are very promising. Combined with the idiom of learning software, ink-enabled software for Tablet PCs have a high and important value in computer-aided learning (as also described in Bull et al and Bergue et al, both 2004), also when exposing simplistic interfaces, as it can be seen from this study.

SketchUML also proves to be a competent tool to facilitate learning the UML syntax. On basis of the findings of Qiu (2007), the recognition accuracy of gestures could be improved by a great deal so that SketchUML can be considered a naturalistic, learning facilitating tool for UML diagrams.

SketchUML has not reached a state of completeness. This study aimed at adding further functionality to SketchUML and assessing its usability in a formal evaluation. Work still needs to be done to improve the recognition performance of handwriting. From Qiu's findings earlier in 2007, gesture recognition was already improved significantly, but SketchUML needs to be more reliable in terms of handwriting. Also, this study has shown that there is some confusion about feedback in the Teacher Mode. In order to disambiguate the meaning of certain complex symbols, especially with regard to connectors, future work has to find a way to implement extended feedback abilities and a better mapping between gestures and their representing symbols. One idea to accomplish this task would be to include a status bar that informs the user what

the last gesture was that the system recognized. Alternatively, tool tips that appear on the screen when the cursor is hovering over a symbol and explain what the underlying symbol represents can help to disambiguate the meaning of similar symbols. For the near future, further development is planned to add more UML-related features (i.e. symbols of the UML syntax that are currently not supported in SketchUML) as well as non-UML-related features, like undo/redo functionality, or a layout algorithm that supports the user to design uncluttered diagrams.

V. Acknowledgement

This project is made possible by the generous support of Dr. Gary Klatsky, Dr. Douglas Lea (both at SUNY Oswego) and Dr. Lin Qiu (formerly at SUNY Oswego) and, in part, The Research Foundation of SUNY through the generous funding of the Hewlett-Packard Development Company, L. P., HP Technology for Teaching Grant Number U06TFH0037C.

VI. References

Berque, D., Bonebright, T., and Whitesell, M. (2004). Using pen-based computers across the computer science curriculum. *Proceedings of the 35th SIGCSE Technical Symposium on Computer Science Education.* Norfolk, Virginia, USA, pp. 61-65.

Bull, S., Bridgefoot, L., Corlett, D., Kiddie, P., Marianczak, T., Mistry C., Sandle, N., Sharples, M., Williams, D. (2004). Interacitive Lagbook: the development of an application to enhance and facilitate collaborative working within groups in higher education. *A book of papers from MLEARN 2004.* pp. 39-43.

Fowler, M. (2003). *UML Distilled, Third Edition: A Brief Guide to the Standard Object Modeling Language.* Addison-Wesley, Pearson Education.

Heines, J. M., Liang, W. T. (2007). Combining, Storing, and Sharing Digital Ink. *Proceedings of SIGCSE 2007.* Covington, Kentucky, USA.

Qiu, L. (2007). SketchUML: The Design of a Sketch-based Tool for UML Class Diagrams. *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications.* Chesapeake, Virginia, USA, pp. 986- 994.

Simon, B., Anderson, R., Hoyer, C., and Su, J. (2004). Preliminary experiences with a tablet PC based system to support active learning in computer science courses. *SIGCSE Bull.* Vol. 36, No. 3, pp. 213-217.

LAKE-EFFECT STORM CLIMATOLOGY STUDY

Andrew Aizer, Ted Letcher, & Scott Steiger
Department of Earth Sciences

Lake-effect weather can affect millions of lives in upstate New York every winter. A major field project aimed at studying lake-effect weather is currently in the planning stages. In order to assist the planning and coordination of this project a climatology study of lake-effect weather is underway. The two main aspects of this study are focused on 1) snow band climatology and 2) lake-effect lightning climatology. In both cases there was little differences between Lake Erie and Lake Ontario with respect to total events, however there was a much more significant difference when the lakes were compared on a monthly scale.

I. Introduction

Lake-effect snow can severely impact residents near the lower Great Lakes (Erie and Ontario). Lake-effect storms are relatively well-forecasted; however there are still significant errors with respect to predicted band location and snowfall rates (Ballentine, 2007). More research is needed to fully understand the dynamics and behavior of this phenomenon. The goal of this project is to determine the climatology of lake-effect events in order to support the planning of a major scientific field project aimed at studying lake-effect weather produced by the lower Great Lakes. This is a two part project. Part I of the project will focus on the type of lake-effect snow being produced, specifically the difference in occurrence between wind parallel and shore parallel bands off of both Lakes Ontario and Erie. The second part of the project will examine lightning associated with lake-effect weather events off of the lower Great Lakes.

II. Lake-effect Band Frequency

The first part of this project will answer two questions about lake-effect snow occurrence per lake: 1) how often per year do they occur? 2) which band type can be expected more often; wind parallel bands or shore parallel bands.

a. Data and methods

Lake-effect events for both Lake Ontario and Lake Erie were observed from 1996 to 2001 between October and March. Those months were chosen because the majority of lake-effect events tend to occur during the cool season.

For this time period, both radar and upper-air soundings were examined in order to search for lake-effect characteristics. They had to be used in conjunction with each other to

act as mutual fail-safes since the radar data used did not have very high spatial or temporal resolution, so reflectivity could have been misinterpreted undermining the credibility of this research. For example, there could be reflectivity over a lake which does not necessarily look like lake-effect snow band and could be associated with a low pressure system, but in reality it is a lake-effect snow event because the upper air soundings show the conditions were right for a lake-effect storm to form. Also, upper-air sounding data were only observed at 00Z and 12Z (7 pm & 7 am) where as radar data were available hourly most times throughout the day. So using upper-air soundings might have hinted a lake-effect band could form, but the radar data can be used as visual confirmation it did actually form.

Some radar data were easy to confirm lake-effect simply by whether or not there was reflectivity over the lake during the day. If there was no reflectivity over either lake, then the day was diagnosed as having no lake-effect. However, if there was reflectivity over the lake then the upper-air sounding data were analyzed in order to confirm the atmosphere met the conditions for lake-effect to occur. Since the radar data was only available every 4 hours, satellite data or higher resolution radar data would give a better idea whether or not lake effect occurred during those gaps, but this would only be used if certain days were questionable. These are the parameters that were examined to determine whether or not the atmosphere was conducive to lake-effect:

i. $T(\text{lake})-T(850) \geq 13°C$, where T = temperature

ii. No to weak low-level vertical wind shear (<30°), between the surface and 700mb

iii. No or weak low level capping inversion, capping inversion base above 800mb.

Once a band was confirmed as having existed, the band was classified as either a wind parallel band (Fig. 1) or shore parallel band (Fig. 2). Niziol et al. (1995) defined both of these band types. A wind parallel band is a band which forms parallel to the low-level wind direction and also one which forms parallel to the short axis of the lake. A shore parallel band is a band which forms parallel to the long axis of the lake. This band will form with a west wind over Lake Ontario and with a southwest wind over Lake Erie.

Using an EXCEL spreadsheet, a binary scale (0/1) was used placing a one where the band type occurred and over which lake also placing a zero if it did not. Then these values were added up obtaining a total for each band type over each month.

b. Results

Table 1 shows the results from analyzing every day from 1996 to 2001 of the months of October through March. The main months for lake effect bands are December and January for both lakes an average of about five events occurred per month per lake.

A difference between the two lakes is Lake Erie tends to have more events in October than Lake Ontario, but Ontario generally has more events toward the end of the lake-effect season in February and March when Lake Erie tends to freeze over. Another interesting note is the preponderance of shore parallel bands compared to wind parallel

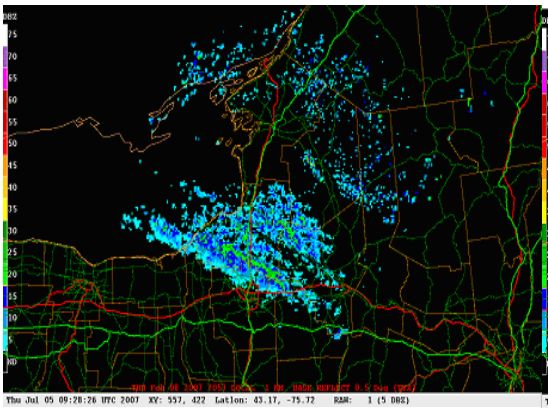


Fig. 1: Radar image of a wind parallel lake-effect

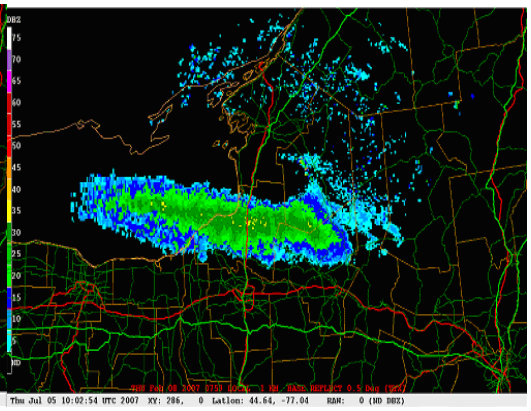


Fig. 2: Radar image of a shore parallel lake-effect band.

Table 1. Average of the number of lake effect bands per month.

Average	**Erie**		**Ontario**	
Month	*Shore-Parallel*	*Wind Parallel*	*Shore-Parallel*	*Wind Parallel*
October	1.6	0.2	0.8	0.4
November	1.8	1.6	1.8	1.8
December	4.4	1	4.6	0.8
January	4.6	0.4	4.4	1
February	1.2	0.4	1	1
March	1	1.2	2	1
Total	14.6	4.8	14.6	6

bands. Shore parallel bands double to triple wind parallel bands in occurrence. Both lakes have the same number of shore parallel events each season (14.6).

III. Lightning Study

The second facet of the lake-effect project deals with the electrical aspects of lake-effect storms. This research is focused on comparing lighting events between Lakes Erie and Ontario. This research is divided into three main sections. The first section compares the frequency of lightning events between Lakes Erie and Ontario. The second step in this research categorizes the lightning events by precipitation type, e.g., rain, snow or mix. The

third and final step in this research compares the intensity (number of flashes per storm) of lightning events by lake.

a. Data and methods

There is a particular methodology that was employed during all three parts of this research. All three parts centered on looking at lake-effect lightning events. Materials consisted of a list of lake-effect lightning events between 1996-2007 found by Hamilton et al. (2008), National Lightning Detection Network (NLDN) lightning density plots (e.g.,(Fig. 3) for those events, National Environmental Satellite, Data, and Information Services (NESDIS) radar data, National Severe Storms Laboratory (NSSL) archived surface observations, and upper air sounding data archived by the University of Wyoming.

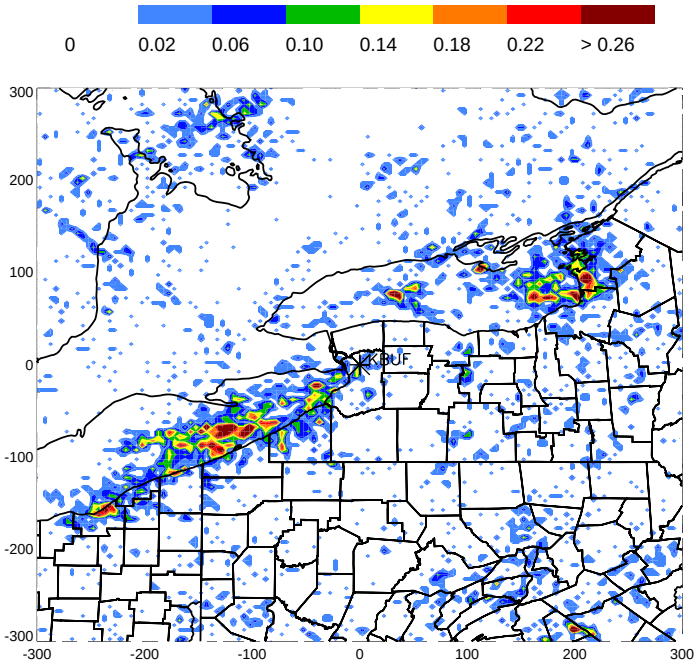


Fig. 3: Composite lightning density for all of the lake-effect events found by Hamilton et al. (2008).

The first section of research, comparing lightning event frequency by lake, centers on the NLDN lightning density plots. The first step in the diagnosing of lake-effect lightning is analyze a lightning density plot for a given lightning event. In some cases this is enough to determine which lake produced the lightning.

However there is often a lot of noise on some of the density plots, and it is hard to determine which lake produced lightning. If that was the case, radar data were examined for the given date, and looped in order to find lake-effect precipitation. However this was sometimes inadequate due to the low resolution of radar data, or lost data. In that case, sounding data (typically from measurements at the NWS at Buffalo) were examined; the

main parameter examined on a sounding was the wind direction in the lower levels of the atmosphere.

b. Results

It was found that both lakes, Erie and Ontario produced a very similar frequency of lightning (Fig. 4). In the 12 year data set of lake-effect lightning events provided by Hamilton et al. (2008) 60 of them were attributed to Lake Ontario and 58 were classified as Lake Erie events.

While the lakes had a similar frequency of lightning events, there were major differences between the two lakes with respect to time of year (Fig. 5). It was found that in the late fall and early winter (October & November) both lakes peaked in frequency of lightning events, which agrees with research done in the past (Niziol et al. 1995).

However, in the early lake-effect season (September through early December) Lake Erie dominated as the main producer of lake-effect lightning events, while in the core of winter in January and February, Lake Ontario produced a significantly higher amount of lightning than Lake Erie.

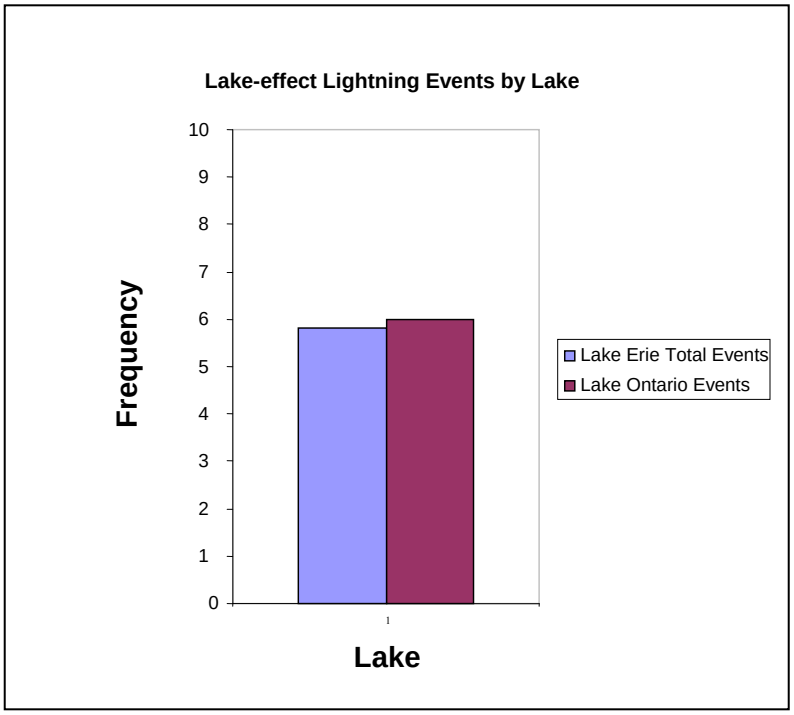


Fig. 4: Frequency of lake-effect lightning events per lake (1995-2007).

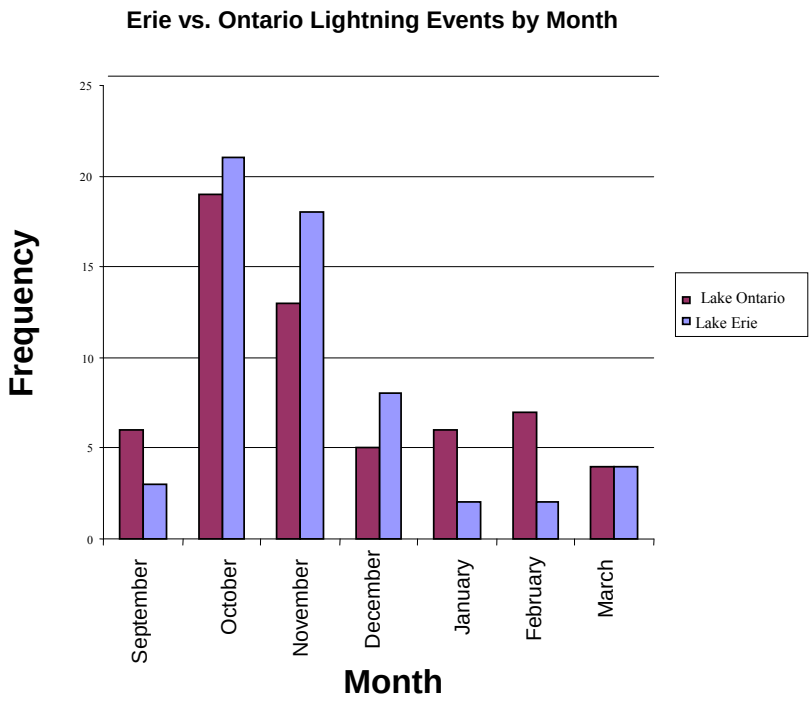


Fig. 5: Lake-effect lightning events compared by month (1995-2007).

c. Discussion

From the results several conclusions can be drawn about the climatology of lake-effect lightning for each individual lake. In the early part of the season lake-effect lightning is more apt to be found off of Lake Erie. This is most likely due to the fact that Lake Erie has a much smaller volume (484 km^3) than Lake Ontario (1640 km^3) (EPA, 2006). From this the inference can be made that Lake Erie's water temperature will fluctuate more through the seasons than Lake Ontario. Hence, in the early lake-effect season Lake Erie will be significantly warmer than Lake Ontario and therefore more likely to produce lake-effect thunderstorms. Subsequently Erie will cool down at a more rapid rate than Ontario throughout the winter and will be colder than Ontario in the latter part of the season.

This shows very good support for the accepted hypothesis that lake-induced instability is directly correlated with the strength of convective updrafts and lake-effect weather. This is because if there is a warmer lake, than there is a higher potential for greater amounts of lake-induced instability and more lightning.

IV. Future Research

The second phase of this project is to determine the precipitation type of lake-effect lightning events, and compare this characteristic for the two lower Great Lakes. This involved a little bit more methodology and is somewhat more subjective. Precipitation type of lake-effect can be difficult and relies on several different parameters.

To categorize precipitation type for this project, three categories (rain, snow, mix) were established and a funnel approach was utilized to make the data more manageable. First

the month of the event was investigated: if the month was January or later it was classified as snow. The reasoning for doing this is that the average lake temperatures for Ontario and Erie at this time of year are around 5°C or less. Since lake-effect storms require at least a 13°C difference between the surface lake temperature and the 850mb level, this incurs that the 850mb temperature must be at least -8°C. In the winter months this will almost always yield a below freezing temperature near the surface, and thus all lake-effect precipitation from January through until the end of lake-effect season will be classified as snow.

The next step in this funnel approach is to examine an upper air sounding from the date (typically taken at 12Z from Buffalo, NY). The temperature at 850 mb is examined and if this temperature is above 0°C then the event is classified as rain.

The next step is to examine surface data for every event. This is done in order to get an idea of the temperatures at the surface around the region, or in the lake-effect precipitation itself. However mixed precipitation can be found even if the temperatures are into the lower to mid 40s °F.

The final step in this methodology is to reexamine at the upper air sounding, and perform a "modified" top down method. A top down method used in diagnosing lake-effect precipitation is very similar to the top down method used to diagnose precipitation types in larger synoptic scale storms (Vasquez, 2002) except for one major difference. In larger scale storms the atmosphere is examined from high up in the atmosphere, where in lake-effect events only the lower section (700mb and lower) of the atmosphere will be analyzed as lake-effect clouds only extend about 3km AGL.

The top down method is fairly simple conceptually. It follows the path of a precipitation particle from start to finish (high to low altitude) and traces its track along the vertical axis of the atmosphere. Layers of above freezing air are typically examined the most thoroughly. If a precipitation particle falls though a significantly thick layer of above freezing air (> 600ft), it will at least partially melt. Rules of thumb have been developed by meteorologists (Vasquez, 2002) to determine how thick of a layer is needed to fully, or partially melt a frozen particle falling through the cloud. By using these accepted rules of thumb the top down method is generally accurate and efficient. The precipitation type of a lake-effect event can typically be accurately estimated by using this methodology.

For the lake effect occurrence study, data from 2002 to present still need to be analyzed in order to confirm the trend developed and explained earlier. After all these years are analyzed, there can be high confidence in the results. Also, satellite data and higher resolution radar data will be obtained for the questionable (unclassifiable) days and a consensus formed on whether or not lake-effect did occur.

The results for the precipitation type classification study are not listed because the data has not yet been fully analyzed. There are some indications that Lake Erie produces more rain and mix events where as Lake Ontario produces more snow events, however this cannot be concluded because there is still a large portion of data to go though.

In the future the precipitation type classification will be completed, and the third aspect of this research project can begin. Again the third and final step in the lake-effect lightning research will be to classify lightning intensity (number of flashes) of each lightning event,

and compare Lakes Erie and Ontario by lightning intensity in order to determine which lake produces the strongest storms.

V. REFERENCES

Ballentine, R. (2007). Improving the Understanding and Prediction of Lake-Effect Snowstorms in the Eastern Great Lakes Region. Final Report on COMET Partners Project, UCAR Award S06-58395.

Environmental Projection Agency. (2006, March 9). Great Lakes Fact Sheet. Retrieved April 28, 2008, from http://www.epa.gov/glnpo/factsheet.html

Hamilton, R., Keeler, J., Orville, R. E., & Steiger, S. M. (2008). Lake-effect Thunderstorms in the Lower Great Lakes. *Submitted to Journal of Applied Meteorology*.

Historical Weather Data Archives [Data file]. (n.d.). Retrieved Spring, 2008, from National Severe Storms Laboratory (NSSL) database: http://data.nssl.noaa.gov/dataselect/

NEXRAD National Mosaic Relflectivity Images [Data file]. (n.d.). Retrieved Spring, 2008, from National Environmental Satellite, Data, and Informational Service (NESDIS) database: http://www4.ncdc.noaa.gov/cgi-win/wwcgi.dll?wwnexrad~images2

Niziol, T. A., Snyder, W. R., & Waldstreicher, J. S. (1995). Winter Weather Forecasting throughout the Eastern United States. Part IV: Lake Effect Snow. *Weather and Forecasting, 10*, 61-75.

Upper Air Soundings [Data file]. (n.d.). Retrieved Fall/Winter, 2007, from University of Wyoming, Department of Atmospheric Science Web site: http://weather.uwyo.edu/upperair/sounding.html

Vasquez, T. (2002). Winter Forecasting. In *Weather Forecasting Handbook* (pp. 143-146). Garland Texas: Weather Graphics Technologies.

THE STRUCTURAL DEVELOPMENT OF THE KANMANTOO SCHIST BASED ON FIAS FROM SAMPLES AT PETREL COVE AND THE STRATHALBYN ANTICLINE

Laura Bianchi (Helen Daly Bohmer Quest best paper)
Department of Earth Sciences

Inclusions within porphyroblasts from Petrel Cove and the Strathalbyn Anticline provide evidence for metamorphic events not previously recognized in the Delamerian orogen because they contain more than five foliations defining at least two foliation intersection/inflection axes (FIAs). At Petrel Cove one FIA is preserved within cordierite. An identical FIA is preserved within staurolite in the Strathalbyn Anticline but another younger FIA is present in andalusite. The latter sequence appears to resolve the problem of apparently synchronous multiple phases of growth of staurolite and andalusite in these rocks (e.g., Adshead-Bell & Bell, 1999). Reactivation has destroyed these foliations or rotated them into parallelism with the bedding, which is why they were not distinguished until measurements of FIAs were made. The FIA succession distinguishes a progression of metamorphic events and further work of this type in the region will provide enough data for the shear senses suggested by this preliminary study to be confirmed.

1. Introduction

The start of orogenesis on the eastern margin of Gondwana in Australia occurred in the Precambrian (e.g., Preiss, 1987; Foden et al, 2002; Foden et al, 2006) with initiation of subduction and commencement of development of the Delamerian orogen (Talbot & Hobbs, 1968; Steinhardt, 1989; Sandiford & Alias, 2002). Although subduction is well documented, the processes associated with its initiation are poorly understood. This lack of understanding results from the problem of accessing information on what took place as deformation commenced. Multiples phases of deformation result in destruction of earlier phases through reactivation of compositional layering which both decrenulates and/or rotates developing and previously developed foliations into parallelism with the bedding (Bell, 1986; Ham & Bell, 2004). Fortunately, porphyroblasts and their inclusion trails preserve evidence for the interaction between heating and deformation during orogenesis for all deformation events after the first, even where this evidence has been entirely destroyed in the matrix because of reactivation of the bedding (Bell et al., 2003; 2004; 2005). Consequently, the measurement of FIAs provides access to both structural and metamorphic information that occurred at the commencement of subduction and which has been completely obliterated in the matrix. This paper uses this approach to access information on the early stages of orogenesis associated with the initiation of subduction on the Eastern margin of Gondwana.

I. FIA Measurement:

A foliation intersection or inflection axis in porphyroblasts (FIA) is measured for a sample by observing the orientation of the switch in inclusion trail asymmetry within porphyroblasts (clockwise or anticlockwise) from a series of vertically oriented thin sections observed consistently in the one direction around the compass (Bell et al., 1998). For example, Fig. 1A contains a simple spiral with a clockwise curvature from core to rim in the 040˚ section but an anticlockwise curvature in the 360˚ section. This switch in asymmetry takes place across the FIA trend. The principle is expanded in Figs. 1B and 1C.

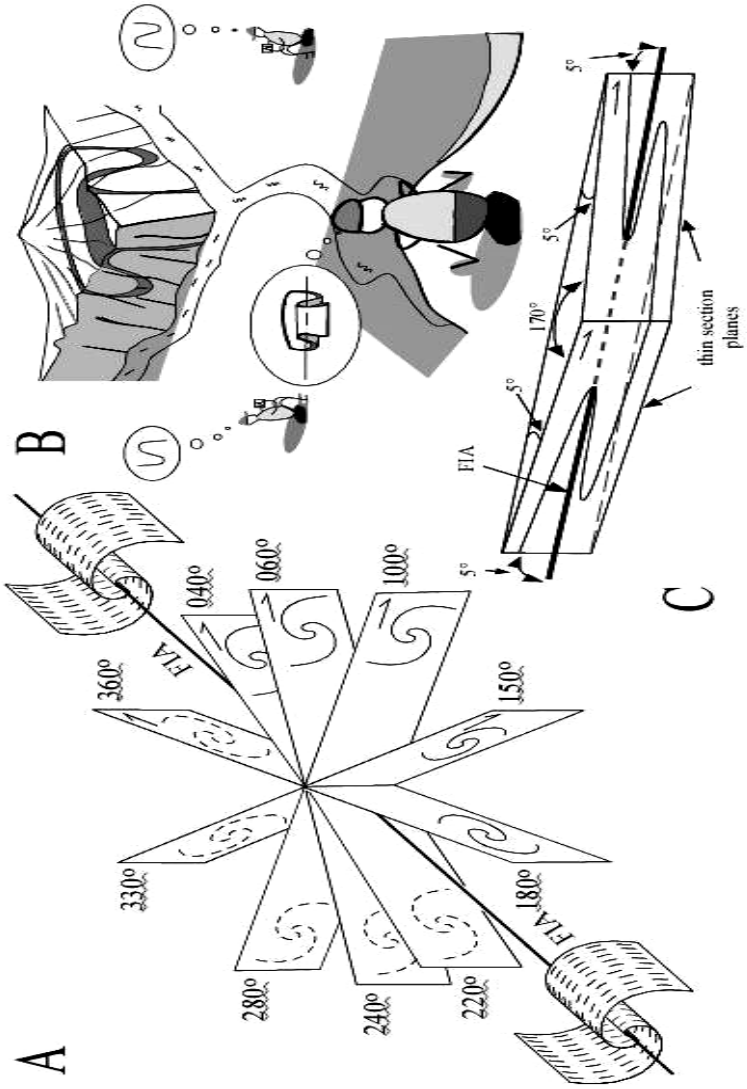


Fig. 1: A FIA trend is measured for a sample using all the porphyroblasts intersected by each vertical thin section, the principle of measurement is shown using a 3-D sketch drawn of a simple spiral. B: Sketch illustrating the principal with a fold preserved in an outcrop. The geologists to either side have no idea of the plunge direction. However the geologist in the centre does. C: Precise measurement of FIA made by cutting sections 10° apart and constraining the asymmetry switch within 10° (Modified from Bell & Newman, 2006).

II. Regional Geology

II. a. *Petrel cove:*

Petrel Cove is located 100 kms south of Adelaide (Fig. 2). The Rosetta Head area (Petrel Cove) in southern Australia form a part of the Kanmantoo Group of supposed Cambrian-age. There is a variety of metamorphic grades represented, from lower greenschist to amphibolite facies. The rocks of the region consist of schists with interbedded meta-sandstone which also contain well-bedded porphyroblastic biotite, andalusite, cordierite and chlorite (Steinhardt, 1989). The schists at Petrel Cove are not quite true pelites because they lack the normal percentage of potassium (Sandiford & Alias, 2002).

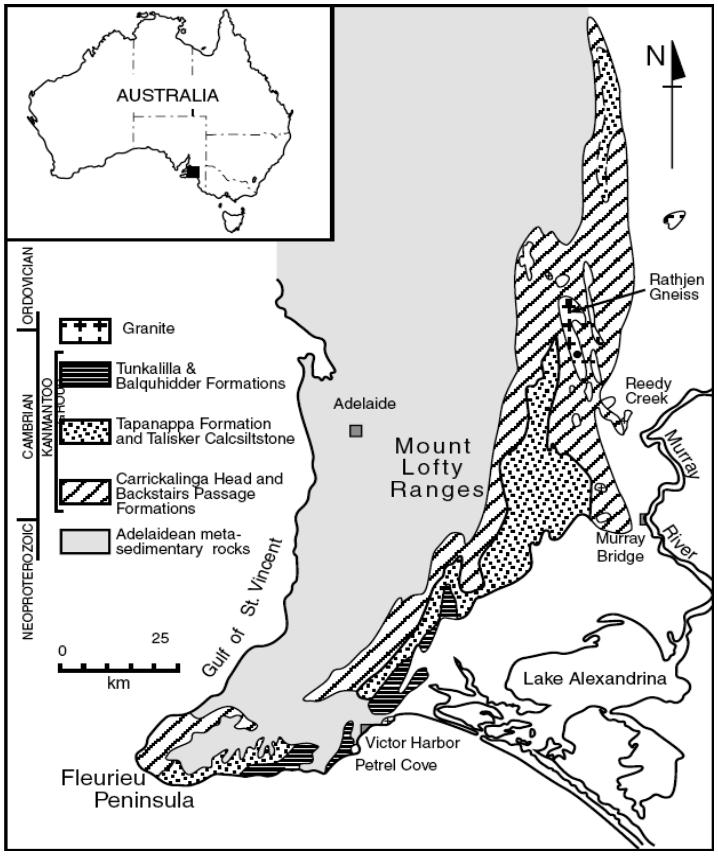


Fig. 2: Map showing location of Petrel Cove, approximately 100 km south-southwest of Adelaide (Modified from Adshead-Bell and Bell, 1999).

There has been much debate over the number of phases of deformation at Petrel Cove (e.g., Talbot & Hobbs, 1968; Steinhardt, 1989; Sandiford and Alias, 2002). The most recent workers (Sandiford & Alias, 2002) have suggested that there were three phases of deformation in these rocks. The significance of D_1 is uncertain because it is only preserved as a foliation parallel to compositional layering. Dolerite dikes and pegmatites were intruded prior to the development of the S_2 fabric. During the third deformation, megacrystic feldspar and biotite growth occurred and granite was emplaced.

The highest temperatures achieved within the Petrel Cove rocks are between 550-600°C. The following equations are chemical changes and mineralogical assemblages

that have been previously described as occurring throughout the Petrel Cove sequence (Sandiford & Alias, 2002),

(1) Chl + Msc → Cord + Bt
(2) Cord + Msc → And + Bt
(3) Cord → And + Chl
(4) And + Cord → St
(5) Msc + Chl → St + Bt + Qtz + H_20

but only the first of these was found in the rocks described herein.

II.b. *Strathalbyn anticline:*

The Strathalbyn Anticline is located approximately 45 km southeast of Adelaide (Fig. 3). These rocks were always somewhat enigmatic because their axial plane lay parallel to all the other regional folds of the Adelaide Geosyncline the north and west which had S_1 axial plane, yet these folds appear to have at least S_3 as their axial plane structure (Fleming & Offler, 1968; Offler & Fleming, 1968). This problem was addressed by Adshead-Bell & Bell (1999) and they showed that the regional folds all formed at the same time, and that the Strathalbyn Anticline was reused and modified several times during younger overprinting deformations.

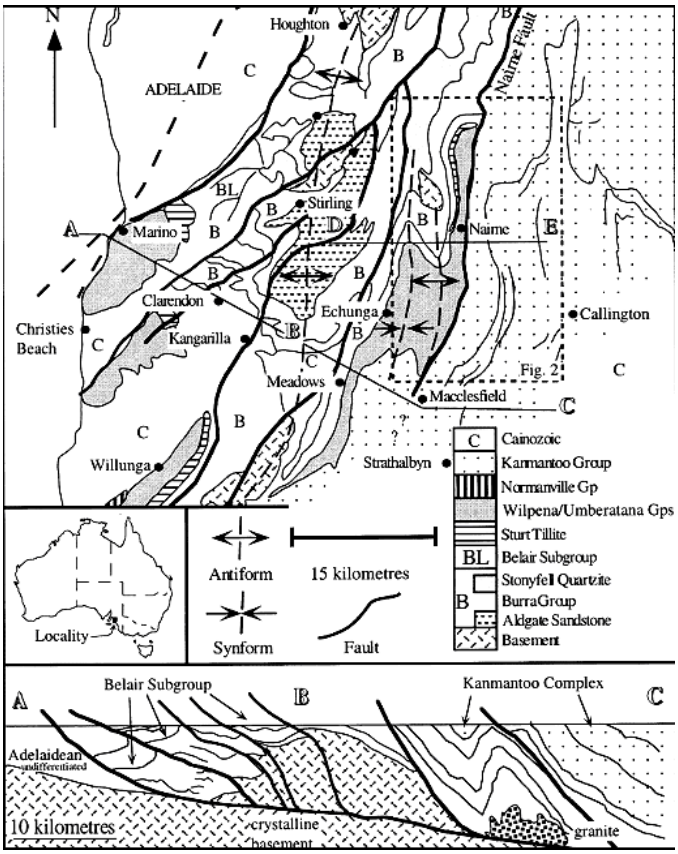


Fig. 3: Southern Australia, the Kanmantoo Group and the Strathalbyn Anticline. The Strathalbyn Anticline is approximately 40km east-southeast of Christies Beach. Two samples, K21 and K24, were examined for the purpose of this paper. The map above shows that K24 comes from just to the west of the anticlinal hinge whereas K21 comes from the east side but further away from the hinge. (Modified from Jenkins and Sandiford, 1992).

Around the Strathalbyn Anticline, the mineralogical compositions consist of namely staurolite and andalusite porphyroblasts in a matrix of quartz, biotite, and on occasion muscovite, ilmenite, sillimanite, and plagioclase. As one moves from west to east, the mineralogy changes from biotite, andalusite, and staurolite to sillimanite (fibrous and prismatic) and migmatites. The following reactions have been inferred based on previous works (e.g., Jenkins & Sandiford, 1992; Adshead-Bell & Bell, 1999).

(1) Msc + Str + Qtz → Al_2SiO_5 + Bt + H_2O
(2) Msc + Chl → Al_2SiO_5 + Bt + Qtz + H_2O
(3) Msc + Chl → Str + Bt + Qtz + H_2O
(4) Grt + Chl → Str + Bt + Qtz + H_2O
(5) Msc + Chl → Grt + Bt + H_2O

III. Sample Setting

III.a. *Petrel cove:*

Two samples, named PC2 and PC3, were collected from locations 2 and 3, respectively, at Petrel Cove (Figs. 4 & 5). Bedding, S_0, usually has a schistosity S_1 lying parallel to it. There appears to be at least two generations of cordierite porphyroblast growth within these schists but they contain the same FIA. A distinctive layering of metamorphic/deformational origin oblique to bedding occurs in phyllites, and cordierite/andalusite mica schists (e.g., Talbot & Hobbs, 1968). This layering is known as the "stripy layering" because of its distinctive appearance. The layers alternate from light to dark-gray bands, and the lighter bands are generally much thinner on the scale of mm-cm in size. It lies parallel to S_2 which dips shallowly SE. Bedding can be very difficult to observe within an outcrop because of the dominance of the stripy layering and many have previously mistaken this structure for bedding (T. Bell pers. comm., 2007). Bedding, S_0, is shown up close in Fig. 5c. The stripy layering in samples PC2 and PC3 consists dominantly of cordierite porphyroblasts and quartz. These porphyroblasts contain the same FIA as that preserved in porphyroblasts in the adjacent matrix. However, the inclusion trails in porphyroblasts within the stripy layering are commonly truncated by the dominant matrix foliation S_2 suggesting that they both predate this foliation and that it may have been rotated into parallelism with them by reactivation (Fig. 6, Bell et al., 2004). Fig. 7 shows detail of some folds at Petrel Cove.

III.b. *Strathalbyn anticline:*

Two samples, K21 and K24, were examined for the purpose of this paper. Fig. 8 shows that K24 comes from just to the west of the anticlinal hinge whereas K21 comes from the east side but further away from the hinge. Sample K24 contains smoothly-curving sigmoidal-shaped inclusion trials in andalusite and staurolite porphyroblasts which are continuous into the matrix. The matrix, defined by slightly elongate quartz grains and aligned biotite grains, shows a flat foliation.

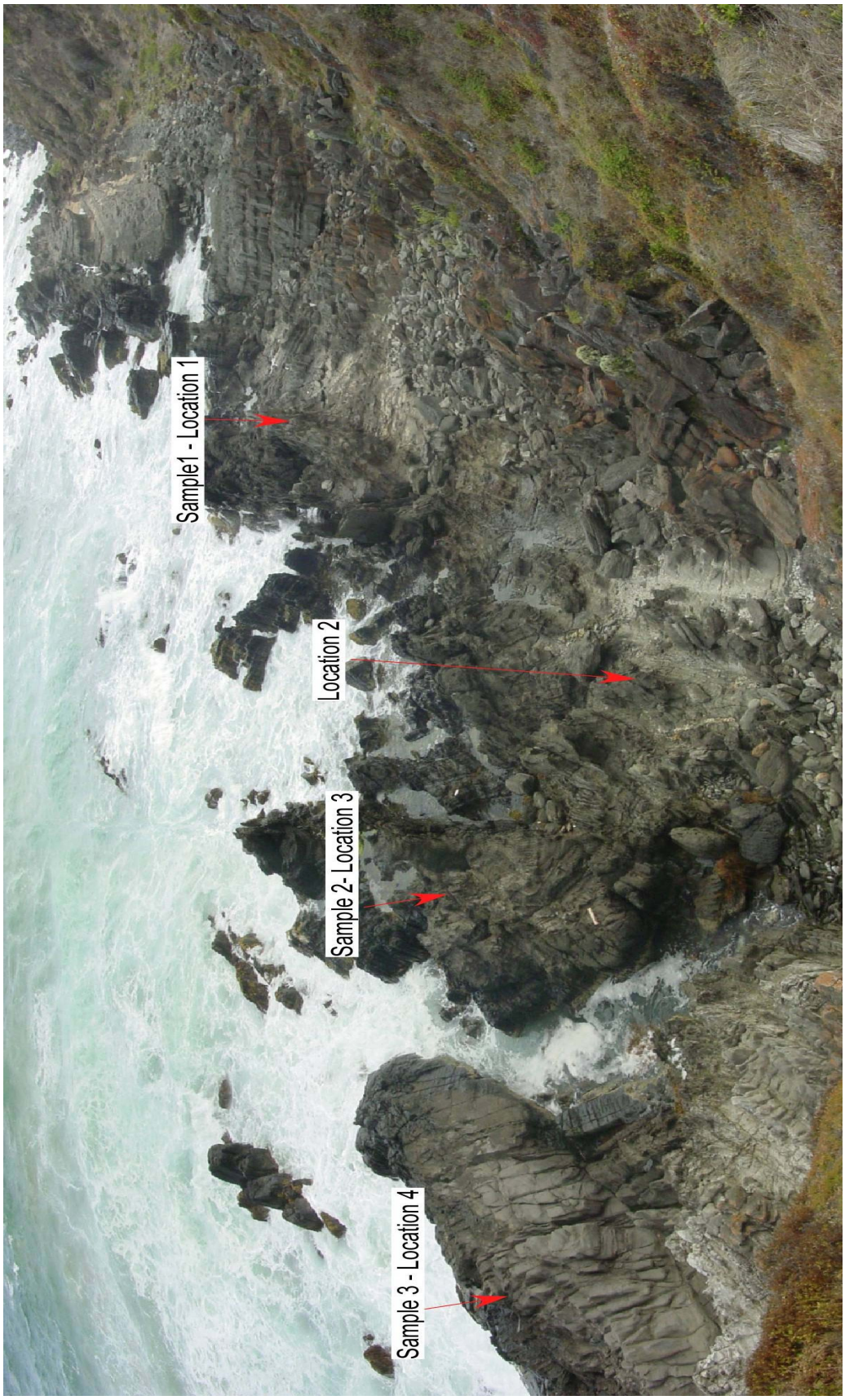


Fig. 4: Petrel Cove, sample setting looking southwest.

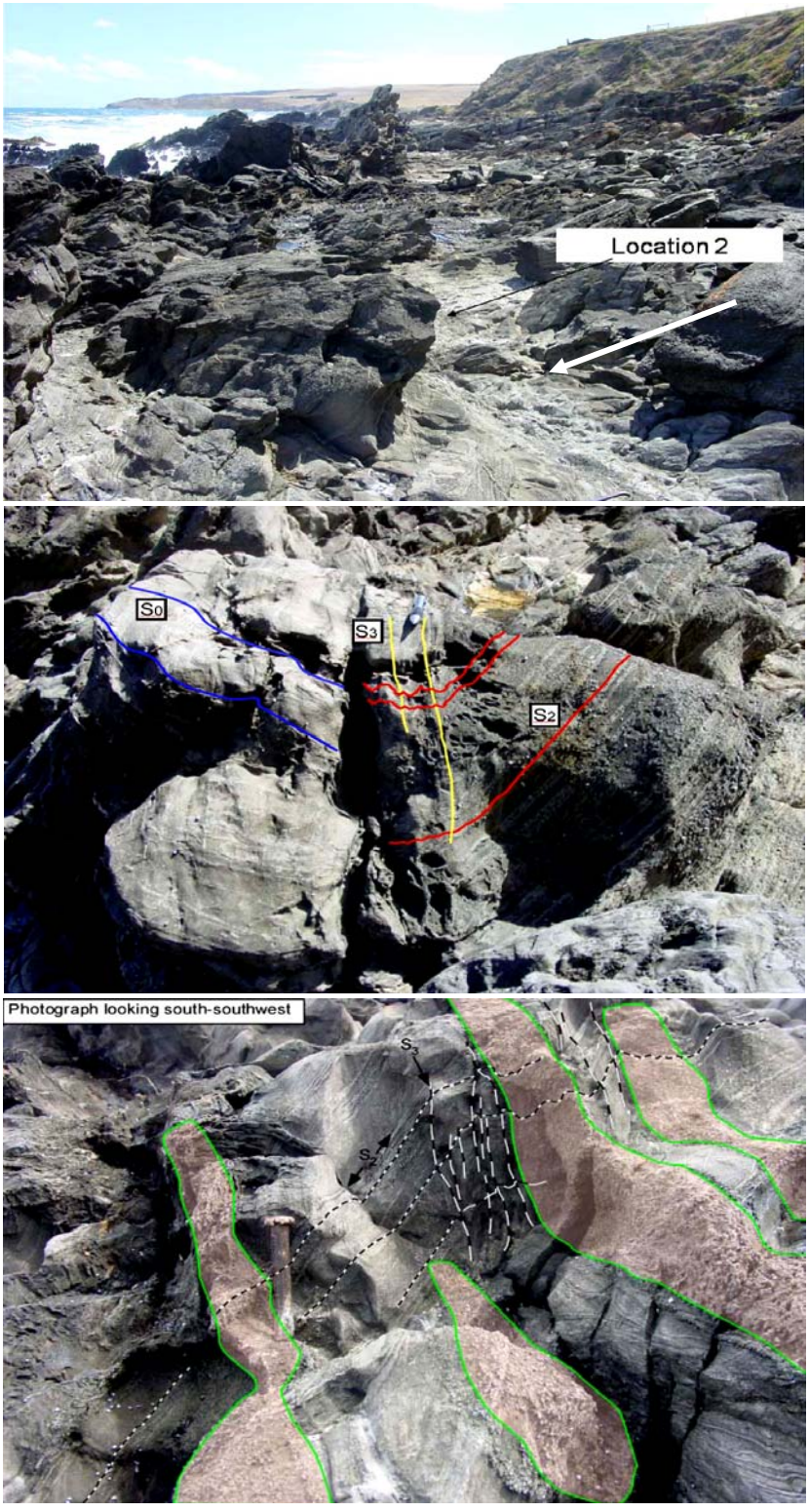


Fig. 5: a) Petrel Cove location 2, facing south. S_0 dips west across the outcrop and is locally outlined in blue. S_2 is parallel to the stripy layering (red), and S_3 has formed sub-vertically (yellow); b) Close up of location 2; c) Petrel Cove station PC 3, photo looking south-southwest. S_0 is outlined in green. S_2 lies perpendicular to bedding and in this outcrop is parallel to the stripy layering. S_3 has formed sub-parallel to bedding.

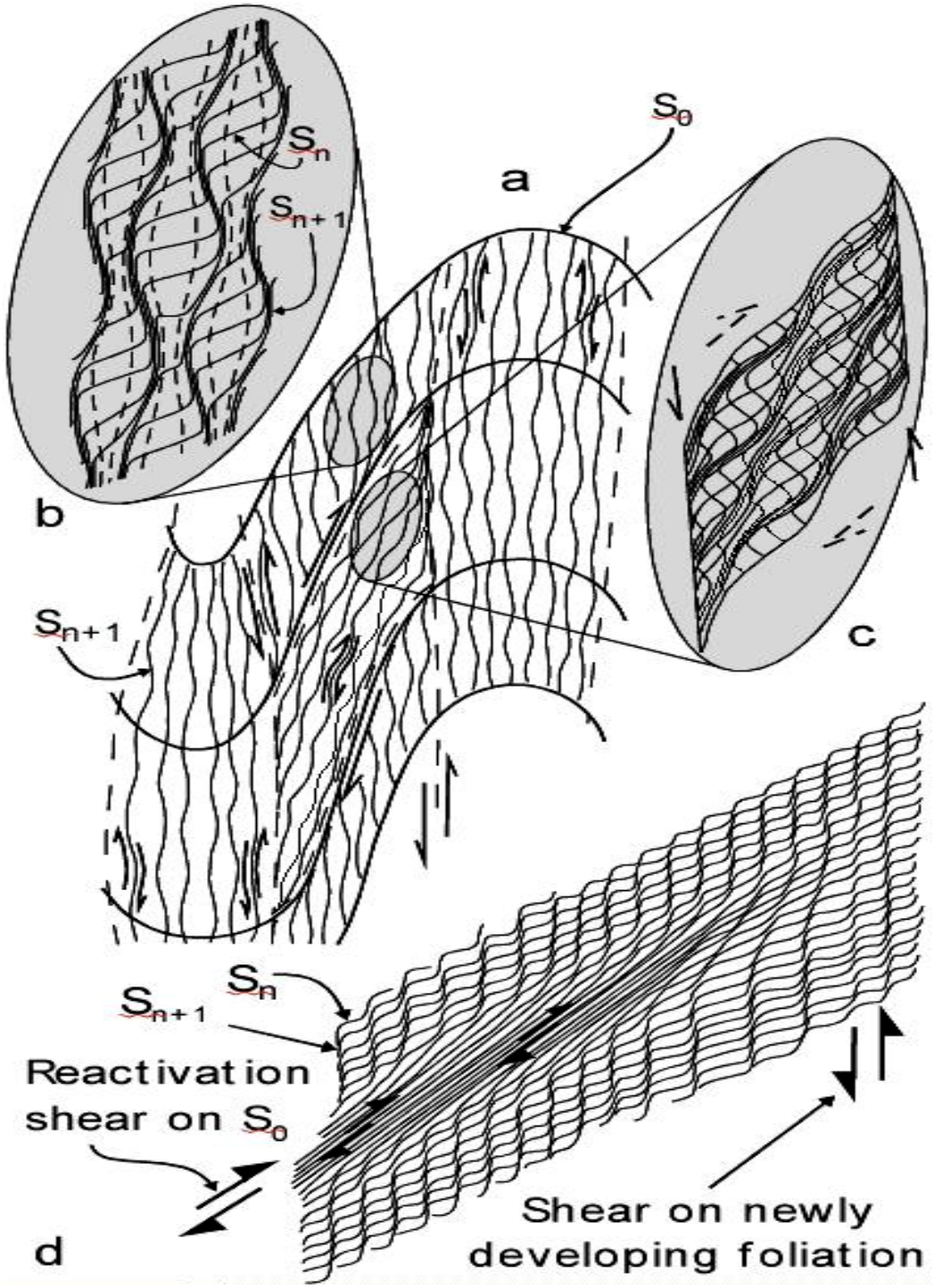


Fig. 6: Model of crenulation cleavage (S_{n+1}) development during folding by buckling. In this example fine-scale crenulations of S_n (sub-parallel to S_0) form by buckling due to bulk shortening (a, b, c), cleavage development (S_{n+1}) occurs through "pressure solution" of long limbs (f, h), and the geometry of the cleavage orientation is a function of the competition between buckling rotation of the fold limbs S_0 (c, e) versus shear on the folded foliation by flexural flow (f, h). Any shear on the cleavage S_{n+1} (f, h) only occurs late in the development of the fold. (Modified from Ham & Bell, 2004).

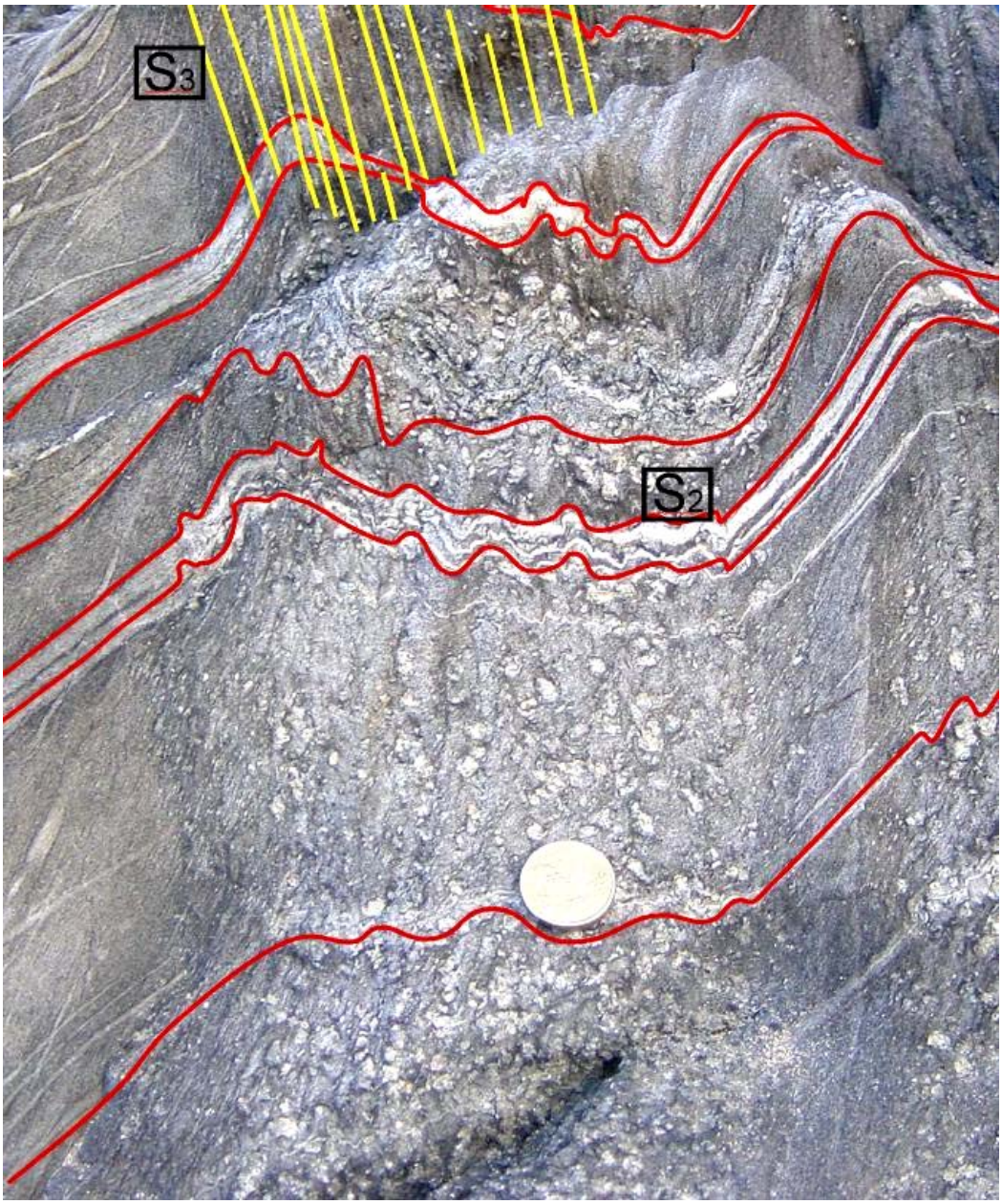


Fig. 7: Australian $2 coin for scale. Foliations seen within a Petrel Cove outcrop looking SW**.** Folded stripy layering, which lies parallel to S_2 (red) and andalusite layering, lies sub-parallel to sub-vertical S_3 (yellow). The shear sense on S_3 is right side up (NW side up).

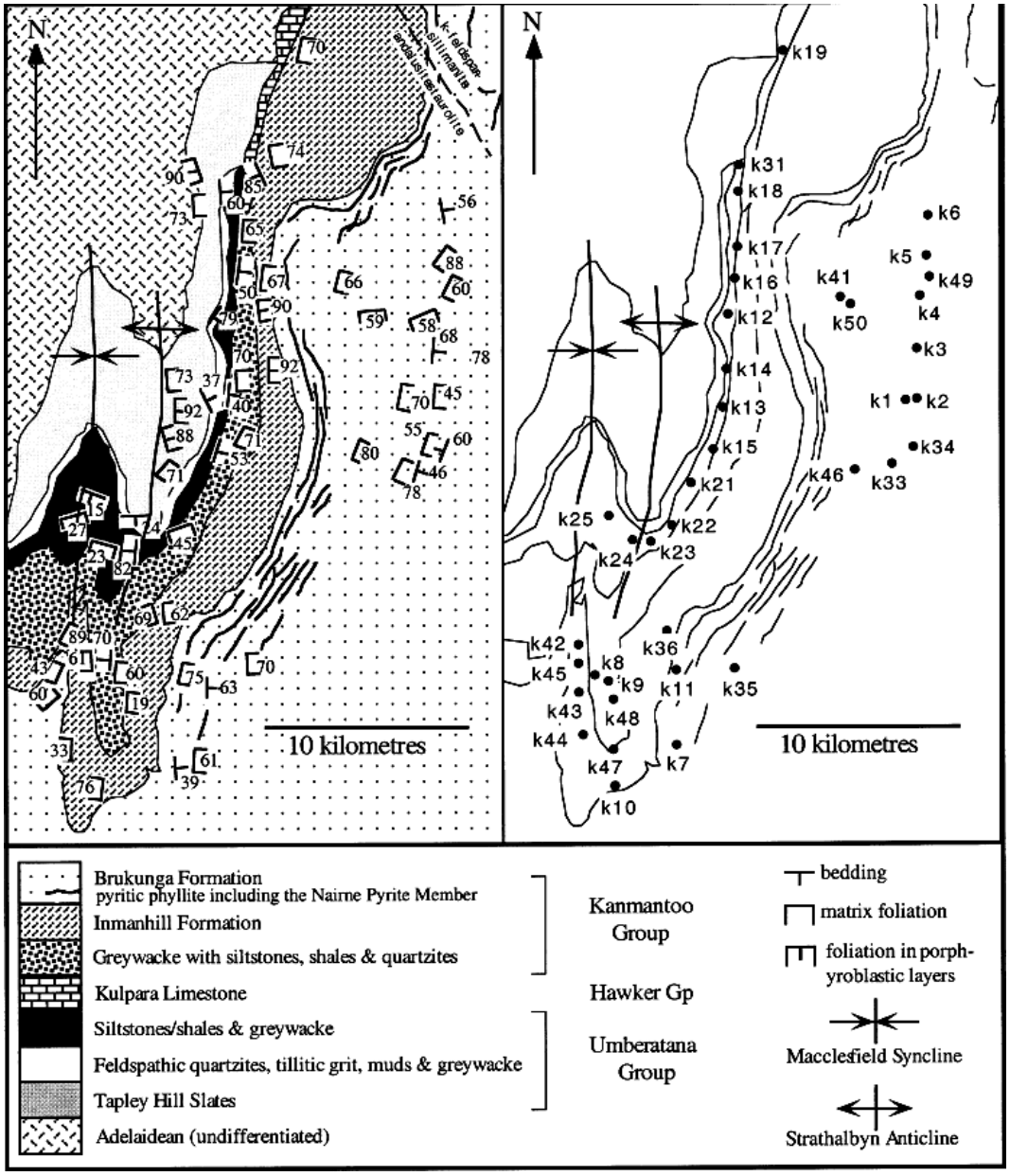


Fig. 8: Location of the Kanmantoo Group and Strathalbyn Anticline (Modified from Adshead-Bell and Bell, 1999).

IV. Structural Setting

IV.a. *Petrel cove:*

Petrel Cove is a region of fine-grained meta-psammites and meta-siltstone that has experienced lower greenschist to almandine amphibolite facies deformation during the Cambrian (Sandiford & Alias, 2002; Talbot & Hobbs, 1968). The S_1 schistosity lies parallel to S_0. Cordierite porphyroblast growth is mainly controlled by different composition beds (Figs. 5b, 5c, & 7), but thin section work has revealed that cordierite

growth defines the stripy layers in many locations and that S_2 lies parallel to them as well. Biotite grains are oriented randomly inside the matrix. However, where they are more aligned, cordierite porphyroblasts can be easily found. The folded stripy layers (Fig. 7) were suggested by Talbot & Hobbs (1968) to contain andalusite/cordierite assemblages associated with intrusion of the Rosetta Head Granite**.**

At the PC2 sample site, there is an increase in the intensity of S_2 (Fig. 9). The number of porphyroblasts also increases within the S_2 fabric, to create the white striped layering. At the PC3 sample site, the S_2 fabric is perpendicular to bedding while S_3 remains parallel to bedding (Figs. 5 & 7).

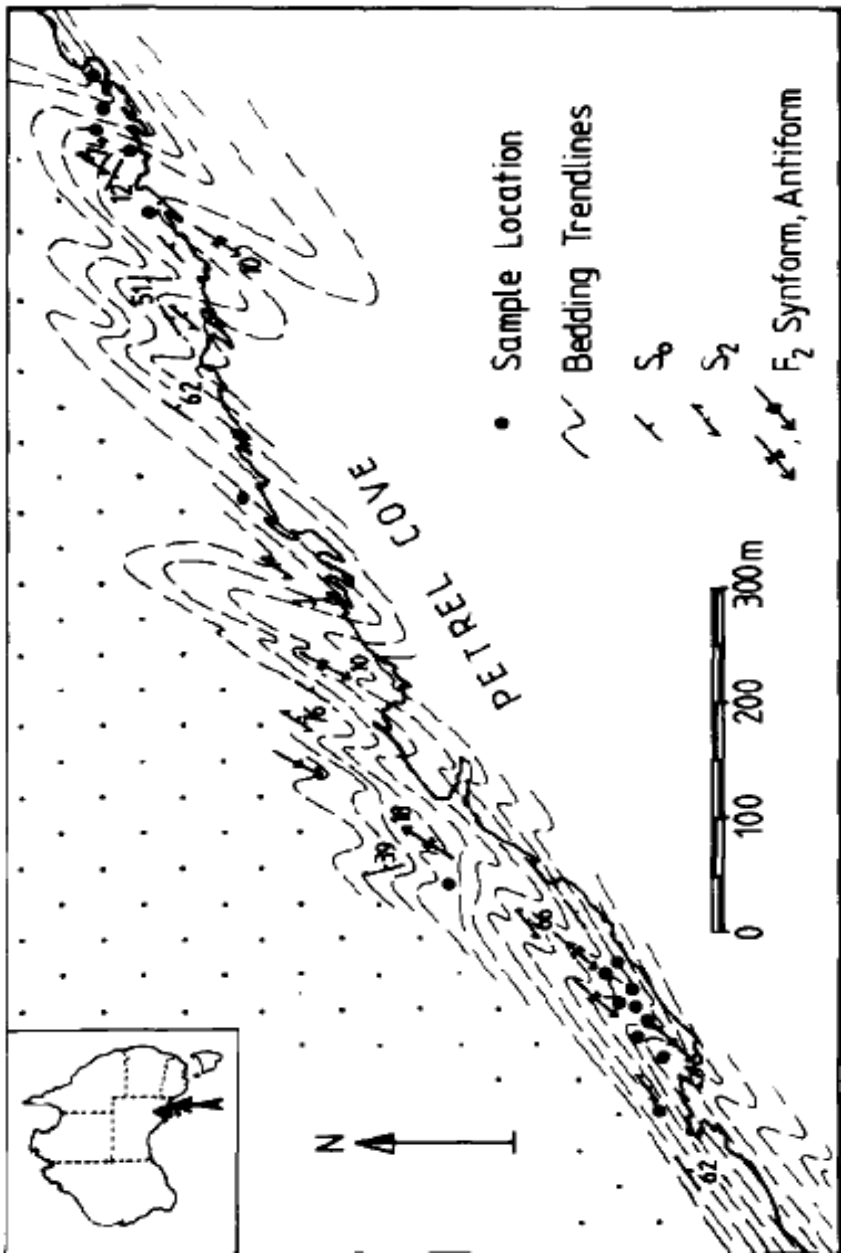


Fig. 9: Closer location of Petrel Cove, showing second deformation event at Petrel Cove (Modified from Steinhardt, 1989).

IV.b. *Strathalbyn anticline:*

The Precambrian rocks of the Adelaide Fold Belt lie below the Cambrian Kanmantoo Group. The Strathalbyn Anticline can be described as a dextral and asymmetrical macroscopic fold which lies within the Kanmantoo group and the Adelaide Fold Belt (Adshead-Bell & Bell, 1999; Bell, 1994). This fold, formed by a thrusting tectonic setting, left the stratigraphy near the surface relatively stable and untouched. The fold is presumed by some to have formed at the same time as the spiral inclusion trails, assuming that the porphyroblasts didn't rotate (Adshead-Bell & Bell, 1999; Bell, 1994). The structure of the anticline has a near-vertical axial plane and a shallowly-plunging hinge (Adshead-Bell & Bell, 1999). The first deformation event left folds preserved which show axial-plane slaty cleavage within the Adelaide Fold Belt; the young crenulation cleavage lies parallel to the axial-plane (Adshead-Bell & Bell, 1999). At least five foliations can be seen which alternate from steep to shallow orientations, thus possibly suggesting that these folds formed as late structures because a differentiation crenulation cleavage and bedding that is parallel to schistosity exist. However, the folds formed in the early part of orogenesis (Adshead-Bell & Bell, 1999).

V. Microstructures/FIAs

V.a. *Petrel cove:*

Cordierite porphyroblasts generally have symmetrical strain shadows (Steinhardt, 1989). According to Talbot & Hobbs (1968), the massive Rosetta Head porphyritic granite intruded well-bedded schists, which contain well preserved bedding, ripple marks, and slump structures. The rocks are broadly folded and contain a well-developed schistosity overprinted by more than one crenulation cleavage with associated asymmetric micro-folds.

Sample locations PC2 and PC3 contain only cordierite porphyroblasts (Figs. 10, 11 & 12). The cordierite porphyroblasts are relatively fresh in PC2. The FIA in PC2 is located at 30°. The stripy layering contains cordierite porphyroblasts and quartz and lies sub-parallel to S_2. The compositional heterogeneity provided by the stripy layering causes it locally to behave like bedding and reactivate when suitably oriented relative to the deforming forces and shear sense (e.g. Ham & Bell, 2004). The stripy layering/S_2 probably formed sub-horizontal and has been rotated NW side up by the S_3 shear which is NW side up in both PC2 and PC3.

PC3 contains cordierite grains that are altered into muscovite and chlorite (Fig. 12). The FIA lies between 30° and 60°. Some cordierite porphyroblasts show NW side up shear indicating a similar relationship to that of location PC2. Most grew in a steep to flat event with shear sense top to SE.

V.b. *Strathalbyn anticline:*

Although many of the matrix foliations that have developed have been subsequently destroyed by reactivation, porphyroblasts and minerals in their strain shadows are preserved. Most of inclusion trails can be traced into the matrix, especially in andalusite, but also can be seen as continuous with staurolite (Fig. 13). Those that are truncated in staurolite commonly show younger andalusite or biotite with muscovite

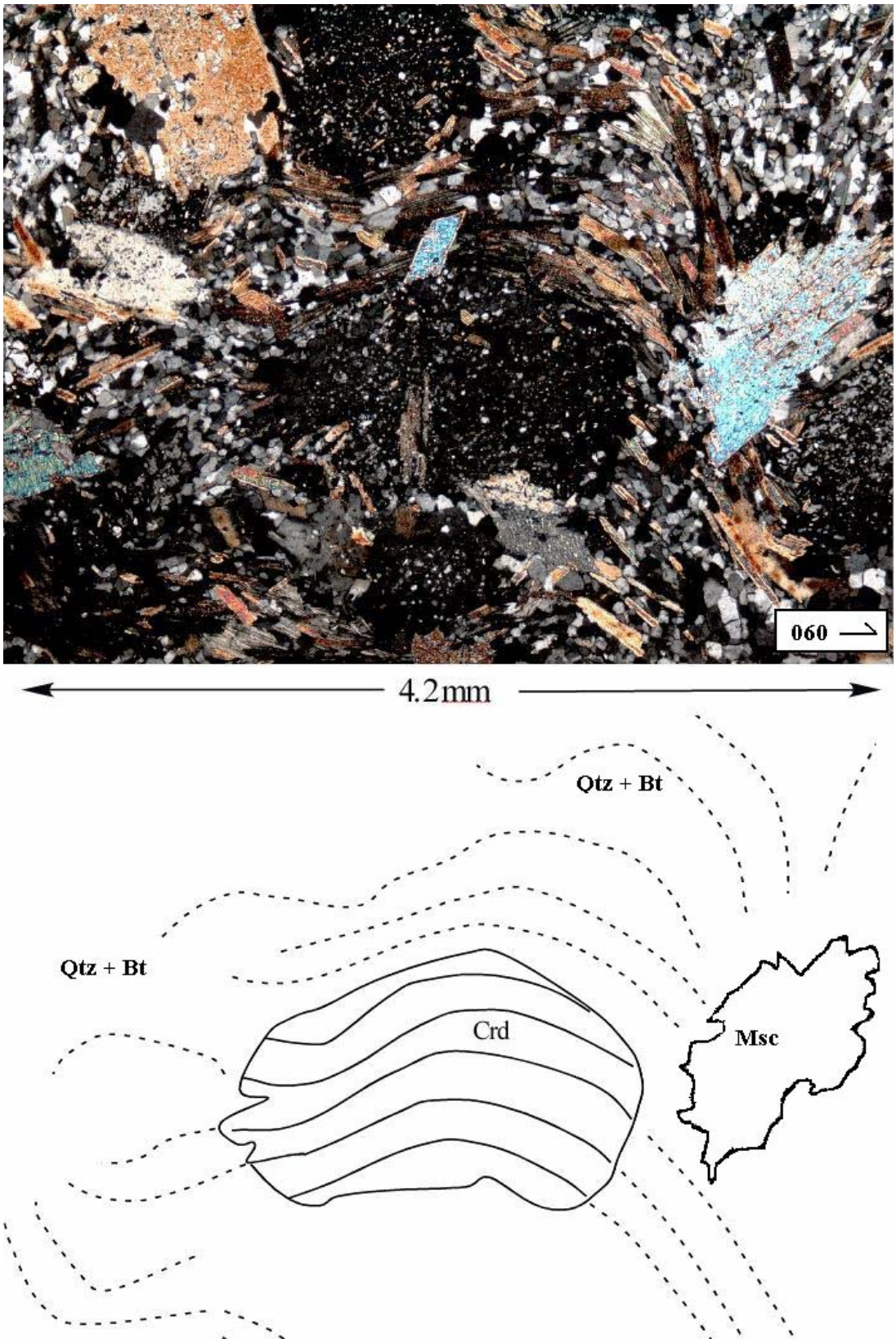


Fig. 10: Sample taken from PC2 at 30º. Sample shows a general clockwise trend in the inclusion trails which are continuous into the matrix of quartz, biotite, and muscovite**.**

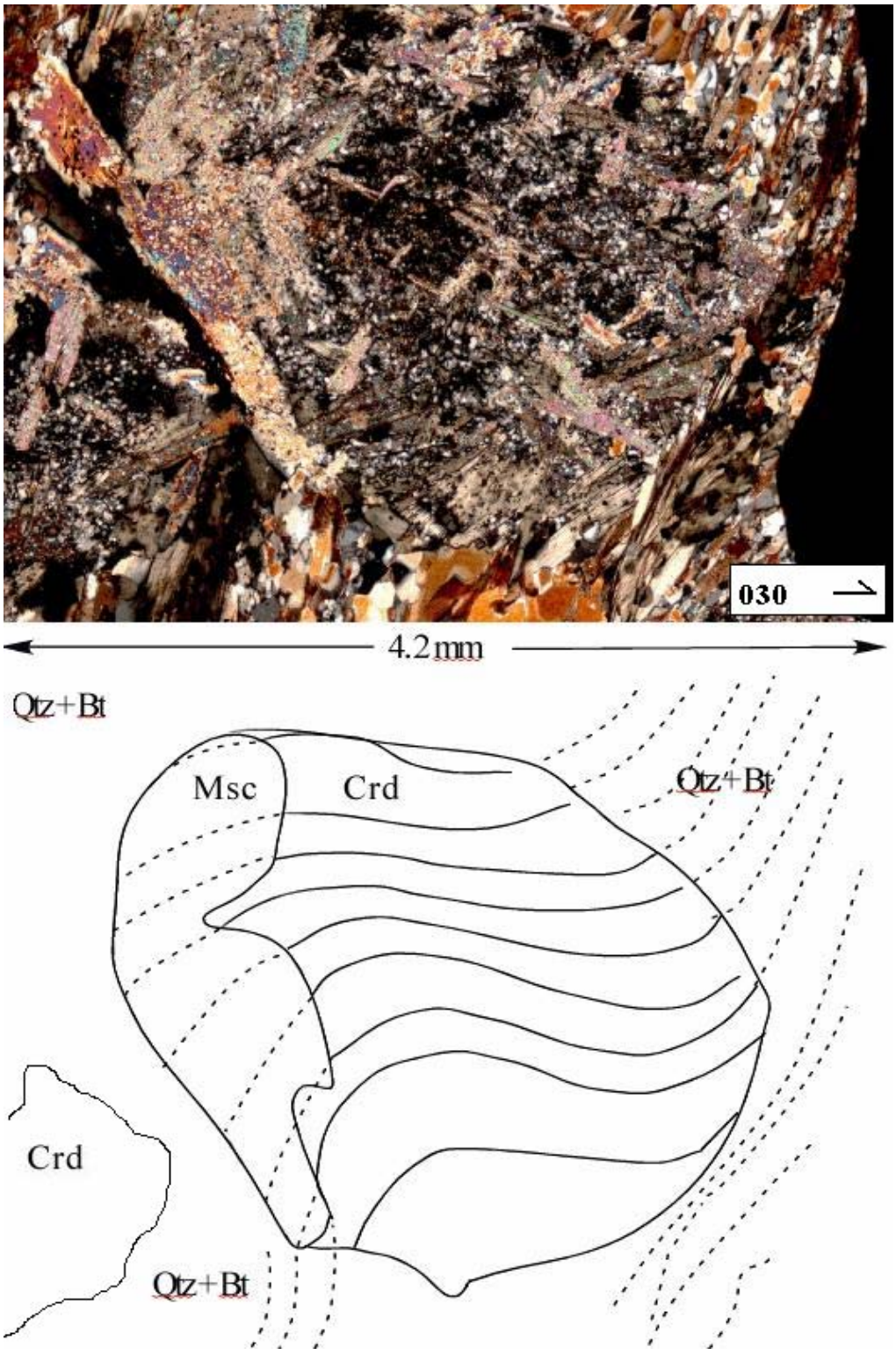


Fig. 11: Sample PC3 trending at 30º. Cordierite has been overgrown by muscovite, and the inclusion trails are continuous into the matrix. Quartz and biotite grains are aligned preferentially.

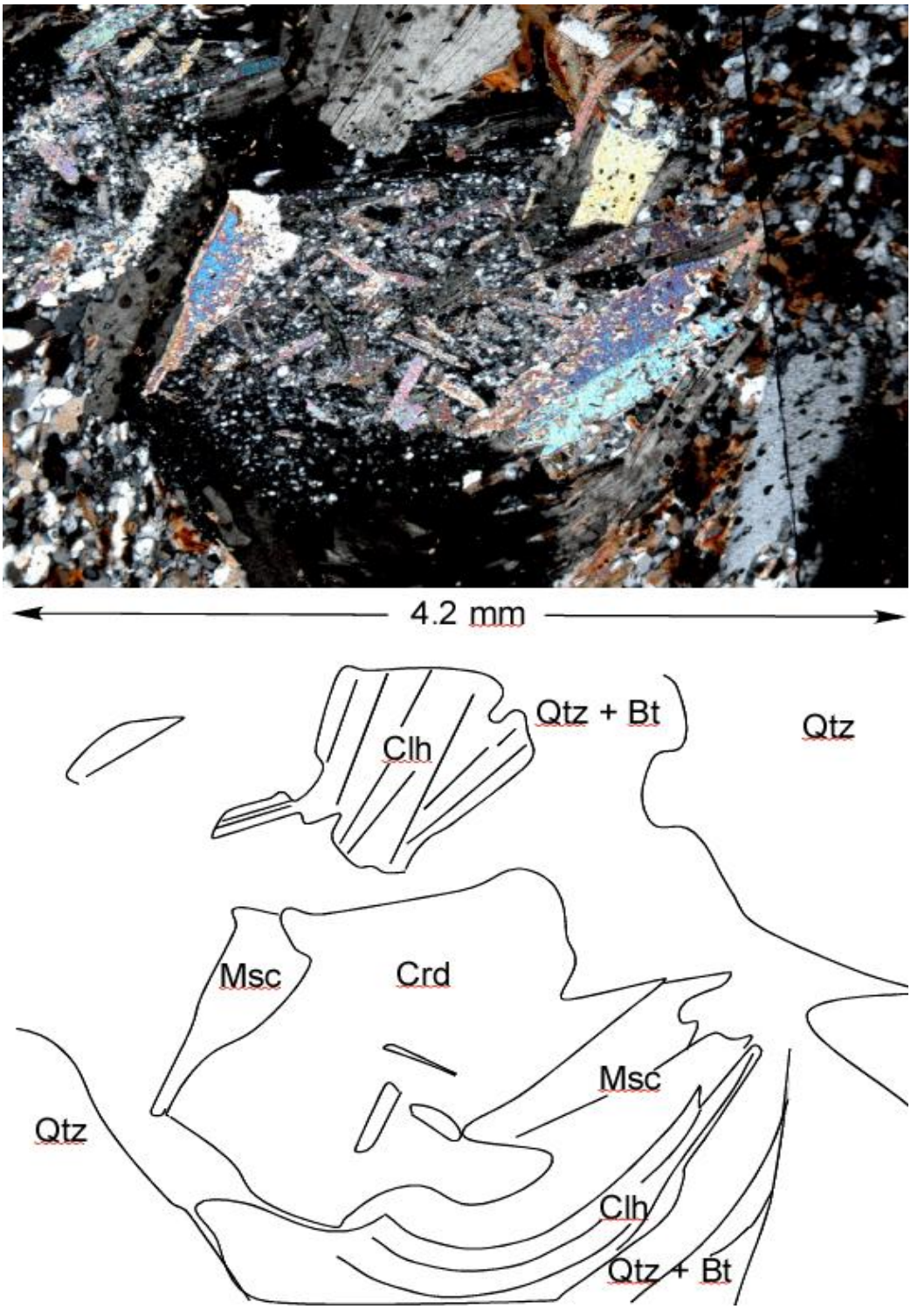


Fig. 12: This figure from sample PC3 with a trend of 60º shows that the cordierite grain in the center has been overgrown by muscovite, chlorite, and biotite. Quartz grains are preferentially aligned. Muscovite grains are randomly aligned within the cordierite suggesting that these grains are relatively older than the matrix.

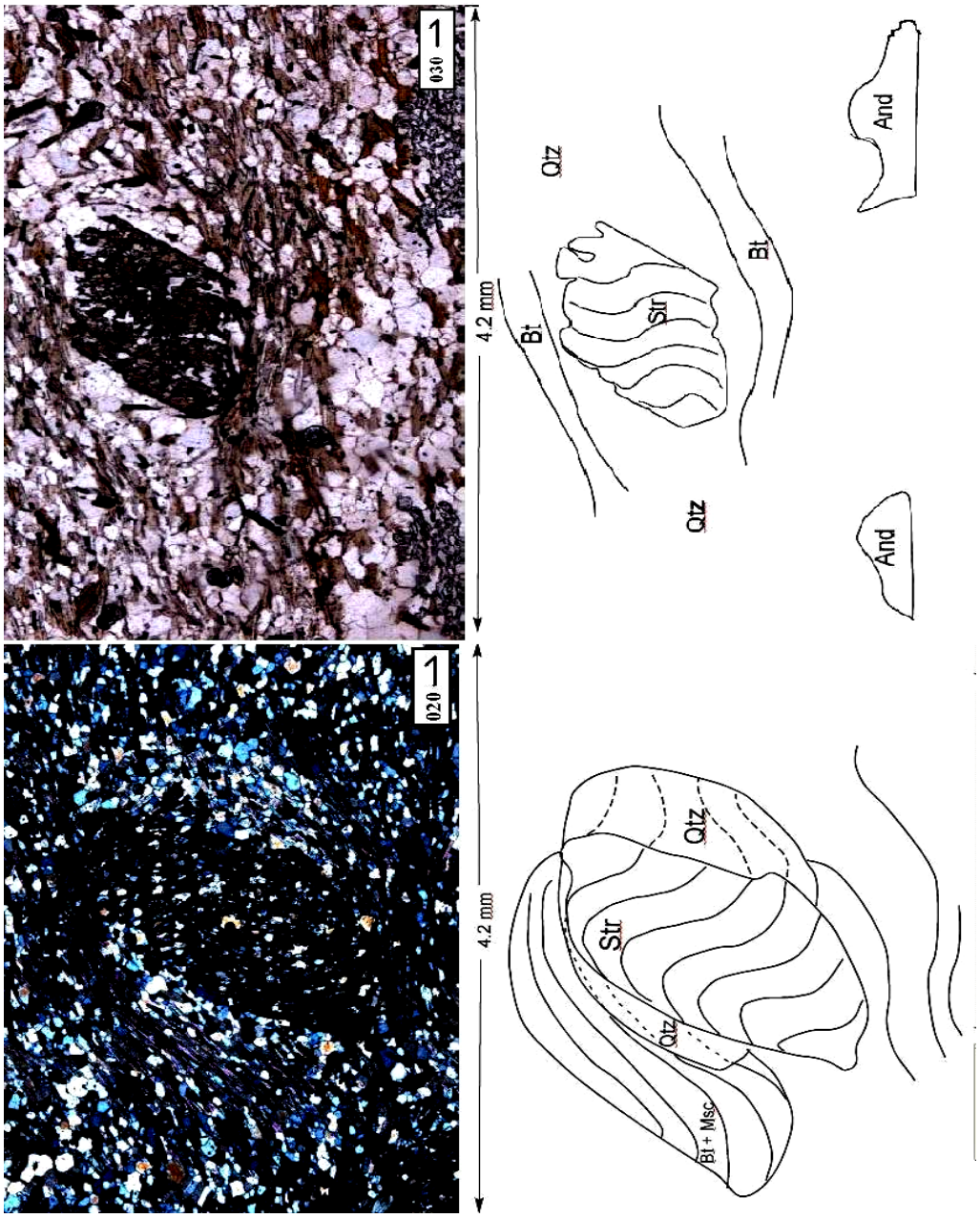


Fig 13: Sample K24 from the Strathalbyn Anticline showing staurolite grains switching from an anticlockwise pattern to a clockwise direction from the 20° to the 30° section. The FIA lies between these two sections. The inclusion trails in the LH figure could be continuous with the matrix foliation as inferred by the dashed lines. The inclusion trails in the RH figure are definitely truncated by the matrix foliation.

that has overgrown the truncational foliation and, of course, where the trails are continuous with those in the matrix. The andalusite grains are highly poikiloblastic with over 50% of their mass made of inclusions. Their trails consist of biotite, quartz, and a little illite and muscovite. The trails are mostly straight across the porphyroblast with slight curvature near the rims, but some are sigmoidal (Figs. 14 & 15).

Sample K21 contains two different FIAs (Table 1). The FIA is located at 30° $\pm$ 10° for staurolite and 55° $\pm$ 5° for andalusite. Within the 0° and 30° thin section samples, the inclusion trails within staurolite are generally straight and gently pitching with a slight curve on their edges. Inclusion trails in staurolite in sections to either side

Table 1: The data collected shows the changes between ACW and CW FIA patterns as well as similarities between the Petrel Cove samples and Stratha lbyn Anticline samples.

Petrel Cove Samples:				CONSIDER THE FOLLOWING:
				*Items in **red** indicate FIA location.
Sample	Cordierite	Andalusite		
PC 2 →				**Blank boxes underneath mineral name
0	ACW		**Presumed FIA for Cordierite is between**	indicates minerals not found in that sample
30	CW &ACW		**0° and 30° (closer to 30°)**	
60	CW			***Underlined minerals indicate that photos
90	CW			have been taken of that particular sample.
120				
150	CW			
PC3 →				
0				
30	fia			
60	fia		**Coridierite: BETWEEN 30° and 60°.**	
90				
120				
Strathalbyn Anticline Samples:				
Sample	Andalusite	Staurolite		
K21 →				
20	ACW	ACW		
30	ACW		**Andalusite FIA is 55° +/- 5°**	
40		CW	**Staurolite FIA is 30° +/- 10°**	
50	ACW	CW		
60	CW			
70	CW			
80	CW	CW		
K24 →				
0	ACW	ACW		
10	ACW	ACW		
20	ACW	ACW	**Andalusite FIA is 45°.**	
30	ACW	CW	**Staurolite FIA is 25°.**	
35	ACW			
40	ACW	CW		
50	CW	CW		
90	CW			

are steeply pitching and tend to be sigmoidal in shape. Inclusion trails in andalusite tend to be sigmoidal and steeply pitching close to the FIA and straight and gently pitching further away.

Sample K24 also contains differently trending FIAs within staurolite versus andalusite. The staurolite FIA trends around 25° whereas the andalusite FIA trends at 45°. Both andalusite and staurolite grains contain steeply pitching sigmoidal trails. Inclusion trials in most staurolite grains are truncated while those in andalusite are

continuous with the matrix. The inclusion trails show NW side up for flat to steep changes in pitch and top to the SE for steep to flat changes in pitch identical to those observed and remarked on by Adshead-Bell & Bell (1999).

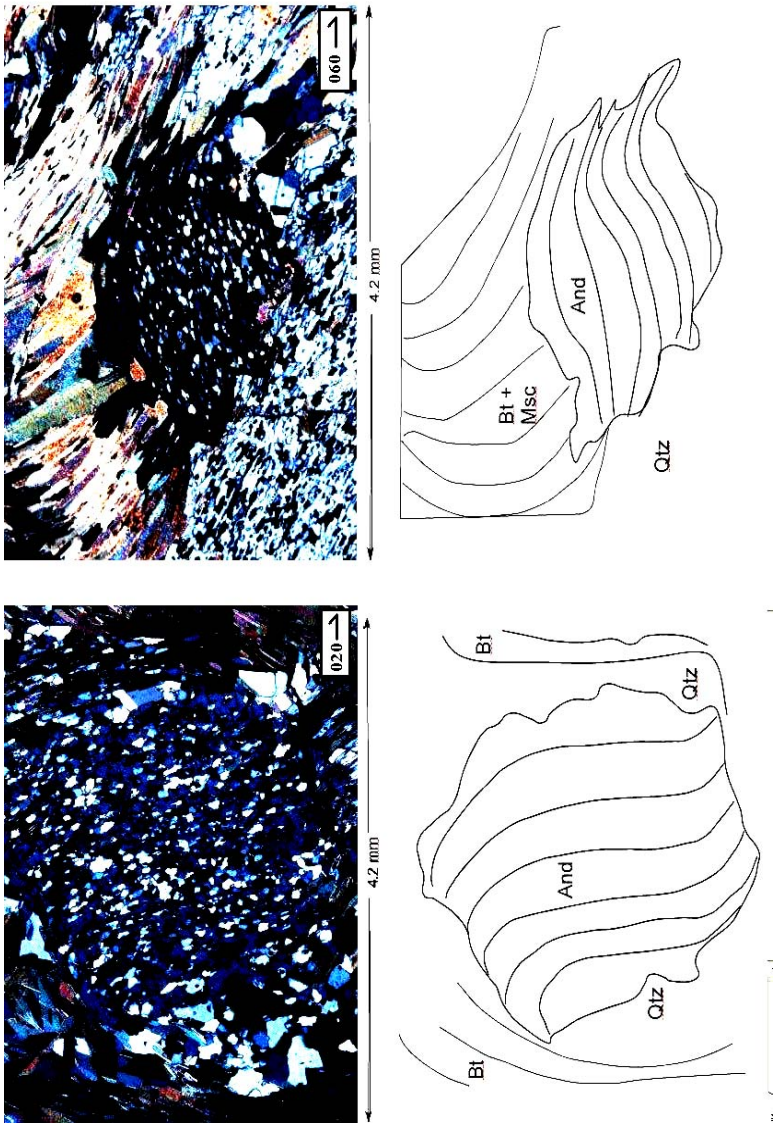


Fig. 14: ACW inclusion trails in Andalusite in the 20° section from Strathalbyn for sample K21, switch to CW in the 60° section across the FIA. The FIA trend is controlled by the sub-vertical foliation-forming event (Bell & Bruce, 2006). The strike of this sub-vertical foliation therefore lies on the FIA. Steeply pitching inclusion trails swing through the horizontal on the FIA as one crosses the strike of the vertical foliation creating them in 3D. Therefore the FIA lies close to 60°. The inclusion trails are truncated by the matrix foliation.

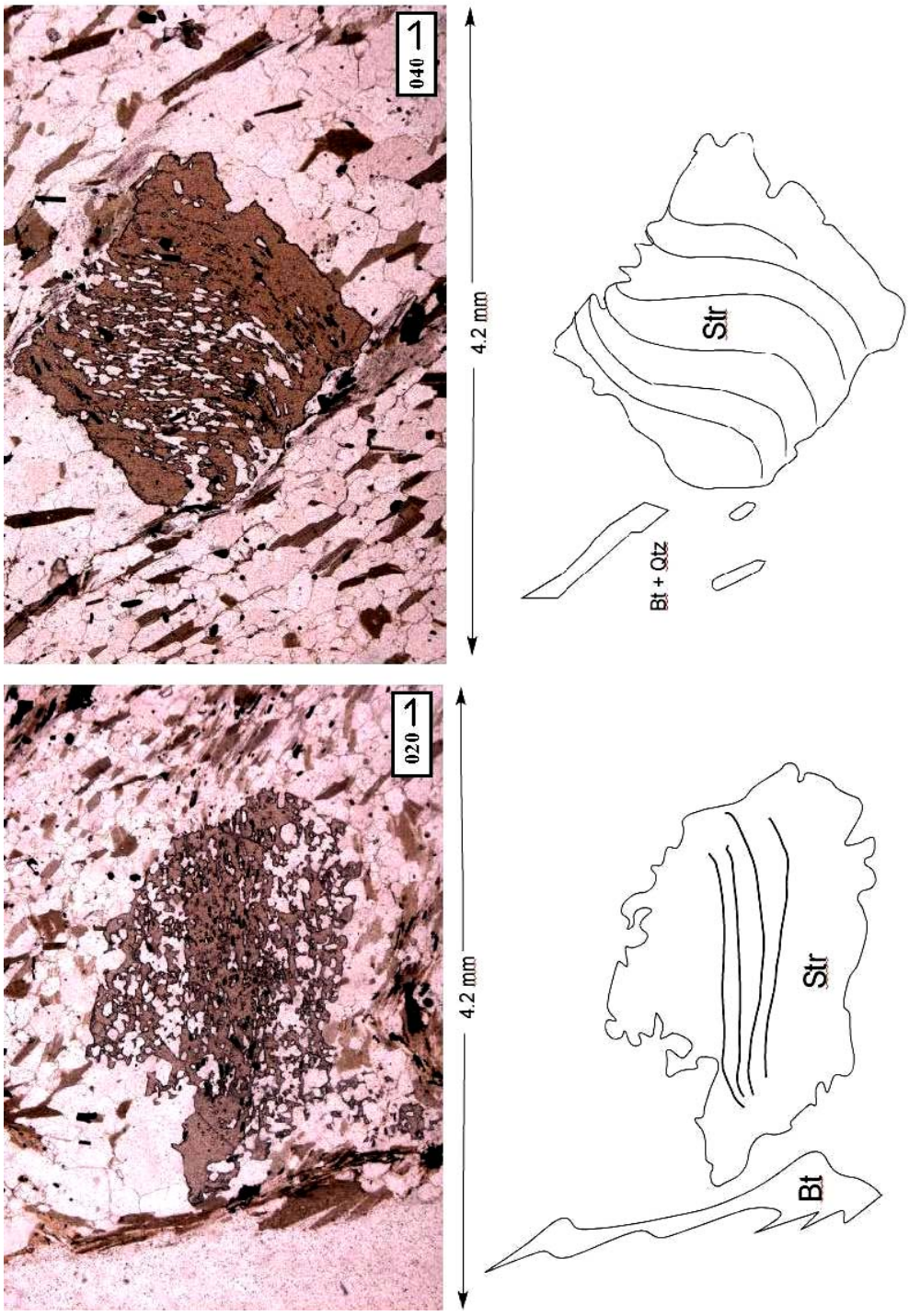


Fig. 15: Shows the switch in asymmetry in staurolite porphyroblasts from ACW in the 20° section to CW in the 40° section across the FIA trend (30°) in sample K21 from the Strathalbyn Anticline. The FIA trend is controlled by the sub-vertical foliation-forming event (Bell & Bruce, 2006). The strike of this sub-vertical foliation therefore lies on the FIA. Therefore, steeply pitching inclusion trails swing through the horizontal on the FIA as one crosses the strike of the vertical foliation crating them in 3D. This suggests the FIA trend is close to 20° than 40°. The inclusion trails in the RH staurolite are truncated. Inclusion trails commonly appear continuos in thin sections that lie sub-perpendicular to the matrix foliation because they exit he porphyroblast into strain shadows relative to that foliation (Cihan, 2004).

VI. Interpretation

VI.a. *Petrel cove:*

The stripy layering that dominates these rocks in outcrop, and which is commonly quite oblique to bedding, controlled where cordierite porphyroblasts grew in samples PC2 and PC3. Therefore, it is likely that in these rocks this layering formed as a result of fluids emanating from a granite that altered the margins of the fractures along which they escaped (c.f., Alias & Sandiford, 2002; Talbot & Hobbs 1968). The steeply pitching inclusion trails preserved in these cordierite porphyroblasts curve anticlockwise looking SW at the rims suggesting that the cordierites grew during top to the SE shearing. S_2 my have formed sub-horizontally at this time as a sub-horizontally pitching foliation in porphyroblast strain shadows merges with S_2 in the matrix. The sub-parallel stripy layering to S_2 indicates that the stripy layering developed relatively early. S_2 therefore may have been associated with W to E thrusting (Jenkins & Sandiford, 1992). D3 resulted in NW side up shear and the local development of a differentiated S_3 in both samples.

VI.b. *Strathalbyn anticline:*

The inclusion trails preserved here show NW side up shear on the steep foliations and top the SW shear on sub-horizontal foliations suggesting uplift to the NW and thrusting to the SE.

VI.c. *Petrel cove versus strathalbyn anticline:*

FIA 1 in the Strathalbyn Anticline is similarly oriented to with FIA 1 at Petrel Cove suggesting that they formed at the same time. FIA 2 in the Strathalbyn Anticline formed subsequently but no porphyroblast growth at this time was seen in the sample from Petrel Cove.

Furthermore, the steep to flat changes in inclusion trail geometries seen in FIA 1 at both Petrel Cove and the Strathalbyn Anticline are identical and strongly suggest top to the SW thrusting during orogenesis. Even though the FIA trends changed by some 30º during the new development of FIA 2 in the Strathalbyn Anticline, the shear sense did not change suggesting that thrusting continued to the SE. This is supported by the fact that flat to steep changes in inclusion tail geometry during the development of FIA 1 in both regions were NW side up indicating uplift to the NW and thrust the potential fold gravitational collapsed and thrusting to the SW at this time (Bell & Johnson, 1989; Bell & Newman, 2006).

VII. Discussion

VII.a. *Porphyroblast rotation?*

A debate over whether porphyroblasts rotate or not has been taking place ever since Bell (1985) proposed that in general they do not. All quantitative data that has been presented on FIA trends indicates that porphyroblasts do not rotate (e.g., Bell & Newman, 2006). The argument as to whether they do rotate has been entirely theoretical and model driven (Fay et al., 2008, in review). Indeed, modellers argued that porphyroblast non-rotation was impossible in a continuous medium. Fay et al (2008)

have recently demonstrated that this is not the case. Indeed, they have discovered the phenomenon of gyrostasis whereby once an anastomosing or millipede geometry (a general strain developed within heterogeneous rocks) has been established by porphyroblast growth early during essentially coaxial deformation, all rotation ceases even in progressive bulk inhomogeneous simple shear (Fig. 16)! Although pressure shadows have been described in the past as being the results from rotation, they can also form as a result of gyrostasis, a process modelled in the form of a mesh-type structure which demonstrates its irregular shape as different magnitudes of strain are forced upon it (Fay et al, 2008; Cihan, 2004). Steinhardt (1989) argued that no porphyroblast rotation had occurred at Petrel Cove because the inclusion trails in all porphyroblasts that he measured were essentially sub-horizontal. Jiang & Williams (2004) inferred that non-rotation of porphyroblasts during non-coaxial deformation is mechanically impossible. However, the results presented herein disagree with his data because they suggest that many porphyroblasts overgrew a steep foliation rather than a horizontal one. One way in which his data could be reconciled with this data is if he cut a preponderance of thin sections striking within 30° of the FIA which trends at 30° in one sample and between 30° and 60° in the other.

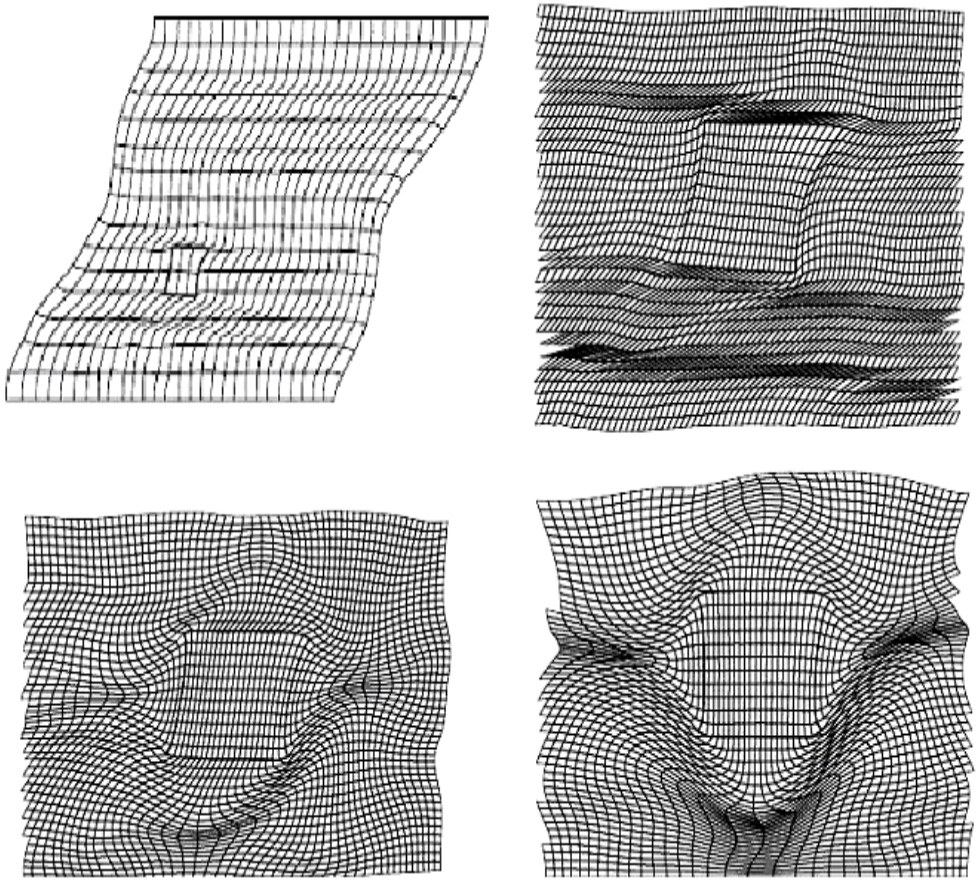


Fig 16: Figures showing the effects of shearing on porphyroblasts. The millipede geometry demonstrates that there are various levels of strain being exerted on the mesh squares, yet none of the above illustrations demonstrate rotation. Bulk shortening is the result of strain on porphyroblasts (Modified from Fay et al, 2008 in review).

The similar FIA trends determined at Petrel Cove versus FIA 1 in the Strathalbyn Anticline, which lie approximately 50 kms apart (Fig. 2), suggest that there has been no porphyroblast rotation (Bell, 1981; Fay et al, 2008) particularly knowing that at least 2 locally penetrative deformations have affected the rocks in the Strathalbyn Anticline post the development FIA 1.

VII.b. *Stripy layering:*

The origin of the white stripes within the rock formations at Petrel Cove has been controversial. Talbot & Hobbs (1968) suggested that they formed in a non-dilatational event and represent an in situ differentiation process with some degree of chemical exchange between the host rock and a solution that may have emanated from the adjacent granite. Sandiford & Alias (2002) have suggested that this layering occurs at a shallow angle relative to schistosity and that the porphyroblasts form augen that are elongate parallel to this foliation. The results presented here provide a solution to both these observations. The stripy layers presumably formed due to fluid emanating from a granite which caused alteration to either side of the fractures in the country rock through which they passed. During subsequent deformation these altered zones were of a bulk composition such that cordierite preferentially grew within them at the prevailing PT. The deformation that accompanied cordierite growth would have caused rotation of the stripy layers towards the developing foliation. Subsequent deformations tended to cause the stripy layering to behave like bedding and thus reactivate (e.g., Fig. 6) and this further rotated any previously formed foliations and decrenulated any newly developing foliation leaving S_2 sub parallel to them (e.g., Ham & Bell, 2004).

VII.c. Thrusting*:*

Previous workers (e.g., Jenkins & Sandiford, 1992) have argued that thrusting occurred from the west to east on the western side of the Mount Lofty Ranges (Fig. 2) during the Delamerian orogen. If this was the Precambrian margin of the southern Australian craton, subduction would have occurred of an oceanic plate to the east under these rocks to the west, and this accords with the direction of thrusting observed by Jenkins & Sandiford (1992). The inclusion trail asymmetry data from the FIAs accords with this both at Petrel Cove and in the Strathalbyn Anticline.

VIII. References

Alias, G., Sandiford, M. (2002). The P-T record of synchronous magmatism, metamorphism and deformation at Petrel Cove, southern Adelaide fold belt. *Journal of Metamorphic Geology* 20, 351-363.

Ashead-Bell, N. S., and Bell, T. H. (1999). Progressive Development of a macroscopic. *Tectonophysics* v 306, 121-147.

Bell, T. H. (1981). Foliation development: the controls, geometry and significance of prograde bulk inhomogeneous shortening. *Tectonophysics* 75, 273-296.

Bell, T. H. (1985). Deformation partitioning and porphyroblast rotation in metamorphic rocks: A radical reinterpretation. *Journal of Metamorphic Geology* 3, 109-118.

Bell, T. H. (1986). Foliation development and refraction in metamorphic rocks; reactivation of earlier foliations and decrenulation due to shifting patterns of deformation partitioning. *Journal of Metamorphic Geology* 4, 421-444.

Bell, N. S (1994). The Structural and metamorphic reinterpretation of the Kanmantoo Foldbelt, South Australia. *Honours Thesis from James Cook University.*

Bell, T. H., Ham, A. P., and Hickey, K. A. (2003). Early formed regional antiforms and synforms that fold younger matrix schistosities; their effect on sites of mineral growth. *Tectonophysics* 367, 253-278.

Bell, T. H., Ham, A. P., and Kim, H. S. (2004). Partitioning of deformation along an orogen and its effects on porphyroblast growth during orogenesis, *Journal of Structural Geology* 26, 825–845.

Bell, T. H., Ham, A. P., Hayward, N., and Hickey, K. A. (2005). On the development of gneiss domes. *Australian Journal of Earth Sciences* 52, 183-204.

Bell, T. H., and Johnson, S. E. (1989). Porphyroblast inclusion trails; the key to orogenesis. *Journal of Metamorphic Geology* 7, 279-310.

Bell, T. H., and Newman, R. (2006). Appalachian orogenesis: the role of repeated gravitational collapse In:-Styles of Continental Compression, Eds R. *Butler and S. Mazzoli, Special Papers f the Geological Society of America* 414, 95-118.

Bell, T. H., Hickey, K. A.., and Upton, G. J. G. (1998). Distinguishing and correlating multiple phases of metamorphism across a multiply deformed region using the axes of spiral, staircase and sigmoidal inclusion trails in garnet. *Journal of Metamorphic Geology* 16, 767-794.

Cihan, M. (2004). The drawbacks of sectioning rocks relative to fabric orientations in the matrix: A case study from the Robertson River Metamorphics (Northern Queensland, Australia). *Journal of Structural Geology* 26, 2157-2174.

Fay, C., Bell, T. H., and Hobbs, B. E. (2008) In Review. Porphyroblast rotation versus non-rotation: Conflict resolution!

Fleming, P. D., Offler, R. 1968. Pre-tectonic crystallization in the Mt. Lofty Ranges, *South Australia. Geol. Mag.* 105, 356–359.

Foden, J., Elburn, M. A., Turner, S.P., Sandiford, M., O'Callaghan, J., and Mitchell, S. (2002). *Journal of Geological Sciences London* 159, 601-621.

Foden, J., Elburg, M. A., Dougherty-Page, J., and Burtt, A. (2006). *Journal of Geology* 114, 189-210.

Ham, A. P, and Bell, T. H. (2004). Recycling of foliations during folding. *Journal of Structural Geology* 26, 1989-2009.

Jenkins, R. J. F. & Sandiford, M., (1992). Observations on the tectonic evolution of the southern Adelaide fold belt. *Tectonophysics*, 214, 27-36.

Jiang, D., and Williams, P. F. (2004). Reference frame, angular momentum, and porphyroblast. *Journal of Structural Geology* 26, 2211-2224.

Offler, R., Fleming, P. D., (1968). A synthesis of folding and metamorphism in the Mt. Lofty Ranges, *South Australia. J. Geol. Soc. Aust*. 15, 245–266.

Preiss, W. V. (1987). *Geological Survey of Australia Bulletin* 53, 438.

Sandiford and Alias, (2002). P-T Record. *Journal of Metamorphic Geology* 20, 351-363.

Steinhardt, C. (1989). Lack of porphyroblast rotation in non-coaxially deformed schists from Petrel Cove, South Australia, and its implications. *Tectonophysics* 158, 127-140.

Talbot, J. L, and Hobbs, B. E. (1968). The Relationship of Metamorphic Differentiation to other structural features at three localities. *Journal of Geology* 76, 581-587.

SYSTEMATIC FRACTURE ANALYSIS USING HIGH-RESOLUTION IMAGERY: EXAMPLES FROM THE HUDSON HIGHLANDS AND THE LAKE ONTARIO SHORE, NEW YORK

[1]Joshua D. Valentino, [1]Regina L. Cuevas, [1]David W. Valentino & [2]Alexander E. Gates
(Bill Bosch Quest 2008 Best Poster)
[1]SUNY Oswego, Department of Earth Sciences
[2]Rutgers University, Department of Earth and Environmental Sciences

Detailed bedrock joint patterns were studied in two parts of NY, integrating outcrop data and analysis of high-resolution images obtained through the NYS GIS Clearinghouse and supported by The Institute for the Application of Geospatial Technology, and the NY Geological Survey. The areas examined include the Hudson Highlands, Orange Co. and the shore line of Lake Ontario, Oswego Co. Low water level on Lake Ontario provided access to outcrops of the Ordovician Oswego Fm that would normally be covered. Most outcrops include two subvertical fracture sets. The oldest strikes NW, and fresh surfaces contain plumose structure typical of Mode I joints. A second set, strikes ENE and has a variety of morphologies including sinistral en-echelon zones and fracture zones with microgouge. The first set is often displaced across the second with a sinistral shear sense. High-resolution images of Lake Ontario reveal a detailed pattern of fracture density that was compared with the outcrop data. The digital images were spatially analyzed and lineament maps were produced. Master joints that appear in the images were then confirmed at places where they come ashore, and many were mapped beyond the extent of the image using a small motor boat with GPS. The image data was sorted by length and strike, and consistently shows that the second set dominates this region. The western region of the Hudson Highlands is underlain by Proterozoic crystalline rocks and was mapped over the past decade including fracture analysis to compile maps at the scale of 1:24,000. Within Harriman and Sterling Forest St Parks, there are extensive bedrock outcrops with minimal cover, therefore, high-resolution images were used to augment the field analysis. The Kanawauke fault traces ENE across the Hudson Highlands, crosscuts the Proterozoic structures, and has dextrally offset various units. There are two dominant subvertical fracture sets. One set occurs at almost every outcrop and strikes NW, but the second appears to be more regionally restricted to ENE striking faults. Again, digital images of outcrops near the KF were spatially analyzed to produce high-resolution fracture trace and fracture density maps. As anticipated, the fracture density of the second set increases in rocks adjacent to the fault. With this study, it was demonstrated that field data could be integrated with available image data with overall positive results.

I. Introduction

An understanding of bedrock jointing (fracturing) is fundamental any construction project that requires a solid foundation, such as large buildings, industrial sites and bridges. As well, an understanding of bedrock joint patterns, including attitude, distribution, aperture and density (number of joints per volume of rock), is critical in fields that involve subsurface fluid flow, such as in the petroleum and mineral resource industries. Finally, joints are often the primary bedrock porosity and serve as the source for potable water, or as conduits for pollution dispersion and migration in the shallow subsurface.

The quantitative and qualitative study of joints typically requires intensive field work over long periods of time. This field work may involve measurement of the attributes of joints, such those listed above. It may also involve the documentation of mineralized domains (veins), kinematic evidence (movements) and surface morphology to piece together the history of bedrock failure (Twiss & Moores, 2006).

A lineament is a linear feature that appears on a map, air photo or remote sensing image. Lineaments reflect the geomorphology of the land surface at the map scale, but with the acquisition of high-resolution imagery lineaments may represent zones of joints, discrete faults, or in some cases individual joints. Therefore, the analysis of high-resolution air photographs and remote sensing data can aid in the understanding of bedrock joints and fault patterns. However, it is important to "ground-truth" image analysis in different areas with the integration of image analysis with more traditional bedrock study.

Detailed bedrock joint patterns were studied in two parts of New York State (Fig. 1), integrating outcrop data and analysis of high-resolution images obtained through the NYS GIS Clearinghouse and supported by The Institute for the Application of Geospatial Technology, and the New York Geological Survey. The image analysis was completed in conjunction with detailed field work for validation. Two image analysis projects were completed during the summer 2007, and field work was done during the fall and winter 2007. This paper is a preliminary report of the study results.

II. Bedrock Geology and Joint Patterns of the Eastern Lake Ontario Basin

The central Appalachian basin is host to a number of regional joint sets associated with the late Paleozoic Alleghanian orogeny (Engleder, 1979; Engelder & Geiser, 1980; Evans et al., 1989; Zhao & Jacobi, 1997; Younes & Engelder, 1999; Engelder et al., 2001). A number of researchers demonstrated the orthogonal relationship between fracture sets and the trend of broad-open folds in the Paleozoic strata, and concluded that a major set are cross fold joints (Engelder & Geiser, 1980; Engelder, 1985). Cross fold joints are fractures that develop perpendicular to fold axes, and occur as non-parallel sets (Zhao & Jacobi, 1997) in the Appalachian basin of Pennsylvania and New York due to the arcuate shape of the orogen. Generally, the cross fold joints radiate perpendicular to the acuate trend of the orogen. In central New York, the northern extent of the Appalachian basin, steeply dipping fractures that strike generally northwest are interpreted to be cross fold joints (Zhao & Jacobi, 1997). Development of these joints requires a component of orogen parallel tension.

In a systematic study of regional joints, Engelder (1982) proposed a tie between specific joint sets and the contemporary stress within the lithosphere. This set of steeply dipping joints strikes generally east-northeast and were documented throughout the Appalachian basin from Ohio to central New York and it was proposed that they are related to the contemporary stress field (Engelder, 1982).

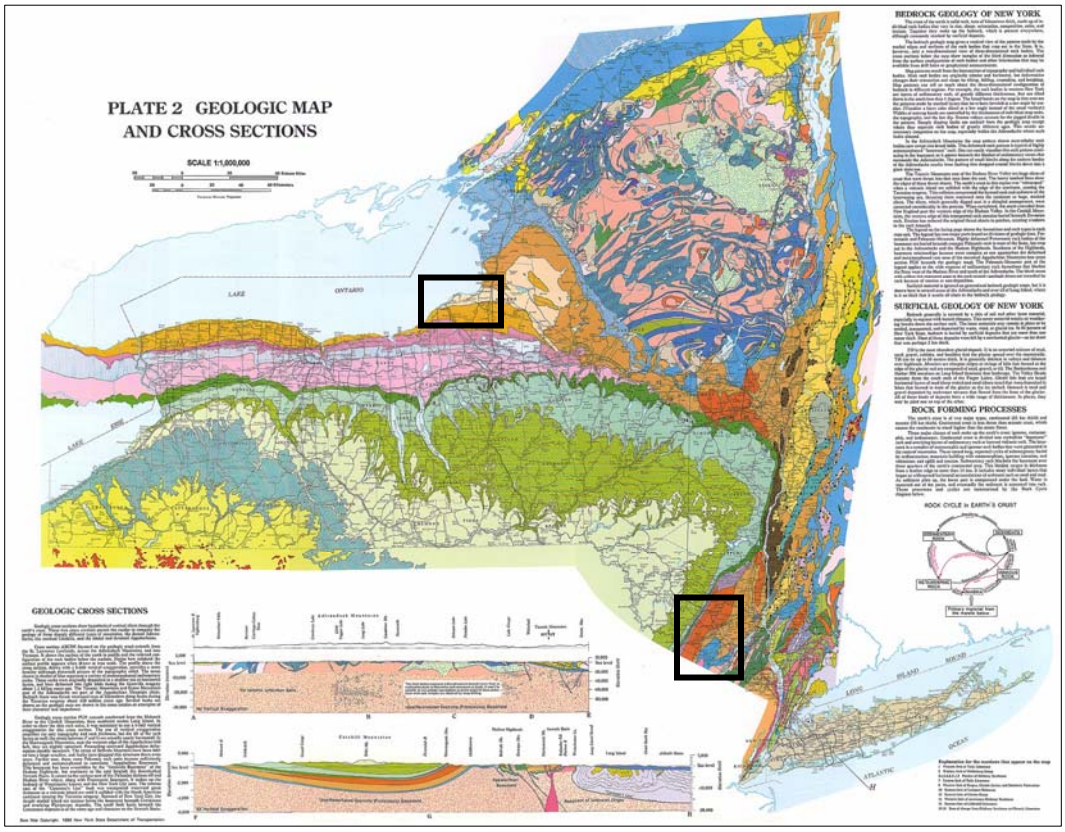


Fig. 1: Geologic map of New York showing the areas that were studied during this investigation (NY State Museum).

Along the southeastern shoreline of Lake Ontario there are abundant outcrops of Ordovician Oswego Formation. As well, the eastern rivers within the basin are down-cut into the underlying Pulaski Formation forming excellent bedrock exposures in the Salmon River gorge. Both of these rock formations are dominated by clastic sedimentary rocks, are generally flat lying and contain a number of systematic joint sets. A recent study of these joints documented sets with orientations similar to the cross fold joints described above. Additionally, a set of east-northeast striking joints have evidence for sinistral shear including displacement of earlier joints and parallel zones of en-echelon fractures. During this field trip six locations will be visited to show the regional and local fracture sets in the Ordovician Oswego and Pulaski Formations in the eastern Lake Ontario basin, with special emphasis on evidence for sinistral shear.

The central region of Oswego County, New York is underlain by the Ordovician sedimentary rocks of the Pulaski and Oswego Formations. The Pulaski Formation consists

of interlayered gray-red sandstone, siltstone and shale with the amount of shale decreasing upward in the section. The best outcrops occur in the Salmon River gorge in the eastern part of the county. The Pulaski Formation is overlain by the Oswego Formation, named for the excellent bedding-plane exposures in the Oswego River and along the shoreline of Lake Ontario in the city of Oswego, NY. The Oswego Formation is characterized by interlayered thin to thick beds of gray-green sandstone and siltstone containing large cross beds, abundant ripple marks, channel structures, and rare trace fossils (Patchen, 1978). The contact between the Pulaski and Oswego Formations is gradational with the Pulaski Formation containing substantial shale and the Oswego Formation containing only minor shale partings (Fisher & Laird, 1978). Taken together, these formations represent a regressive package of sediments associated with the transition from marine to terrestrial depositional environments and the onset of the Taconic orogeny (Patchen, 1978). The Pulaski and Oswego Formations are generally flat-lying with only minor eastward dip of a few degrees locally. No major faults are shown to break these formations in Oswego County, however, both formations contain extensive joints and minor faults.

II.a. Bedrock joint data

Joints were studied at outcrops of the Oswego Formation along the Lake Ontario shoreline, in the riverbed of the Oswego River, and in the Pulaski Formation in the Salmon River gorge. Outcrop maps were completed for all the exposures along the lake-shore and in the Oswego River. These maps include the attitude of every major joint (joints that transect the entire exposure) and a systematic sampling of typical discontinuous joints. As well, these maps document joint density for specific sets, the distribution of en-echelon fracture zones, and domains of high-density joints. The joint density was quantified at most outcrops by counting the number of joints that occur over the outcrop distance, measured perpendicular to the joint set, and the joint density was determined for each joint set separately. Where available, joint patterns were interpreted from high-resolution air photographs (Figs. 2 & 3). The length and strike of individual joint traces were collected from the photographs. Finally, any kinematic information, such as slicken-sides, en echelon zones, or off-set markers were noted to assess potential displacements associated with shear fractures.

II. b. Regional joint patterns

In general, two joint dominant joint sets occur in the Oswego and Pulaski Formations in Oswego County. At any given outcrop, just one may be present, both present with either one as the dominant (higher density) set. Fig. 4 shows the typical distribution and geometry of joint sets in the Oswego Formation at SUNY-Oswego. The Oswego Formation contains thick beds of sandstone with large cross beds common (Fig. 5A), and the Pulaski Formation consists of interbedded shale, siltstone and shale (Figs. 2 & 6). Generally, the two joint sets are subvertical and strike NW and ENE. The regularity of the spacing and

Fig. 2: High-resolution air photograph of the shoreline of Lake Ontario at the campus of SUNY Oswego. Air photograph data was obtained from the New York State GIS Clearinghouse internet database. The locations of bedrock k exposures discussed in the text are shown. The image also shows major underwater fracture sets.

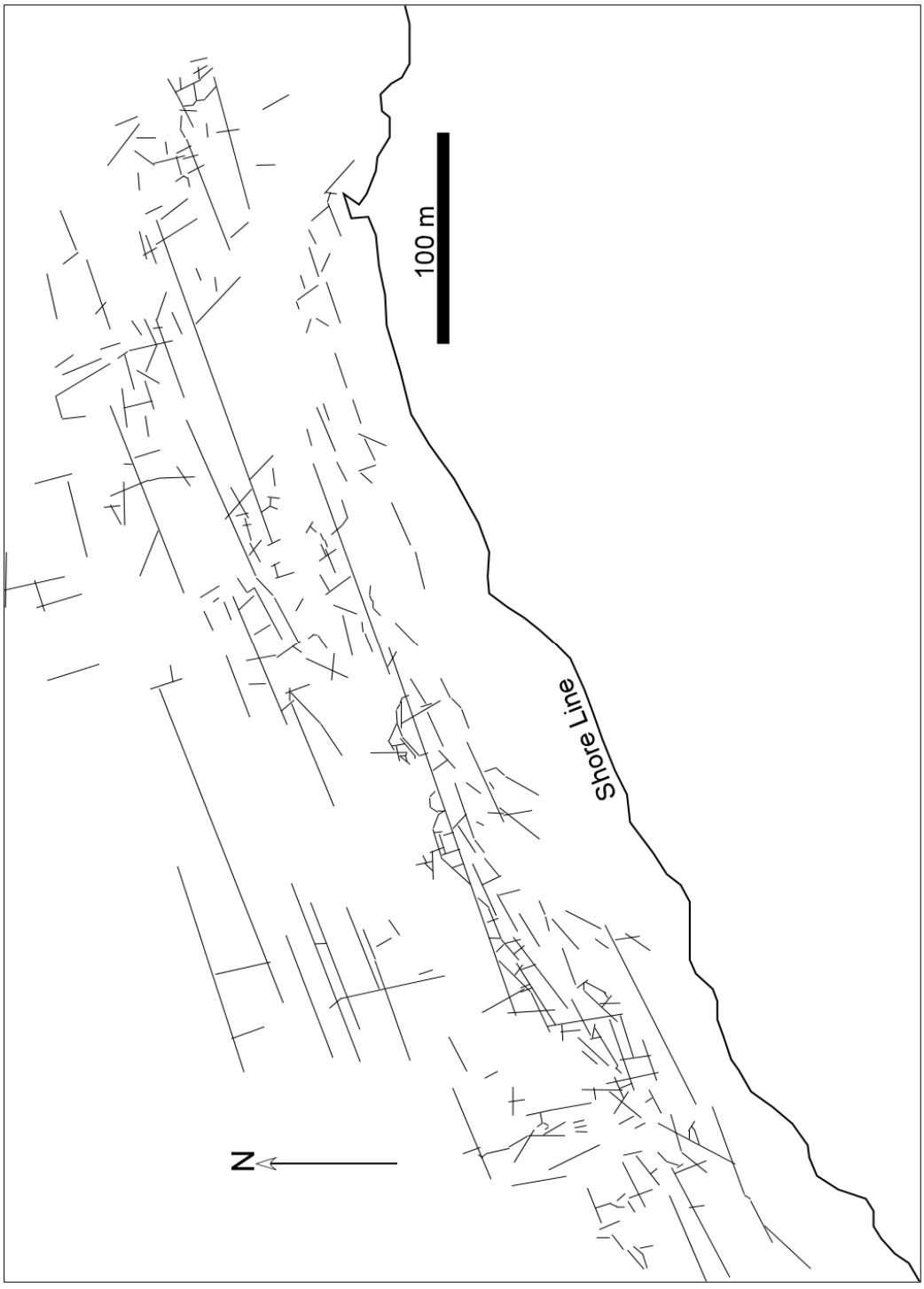


Fig. 3: Fracture trace map interpreted from the air photograph image of Fig. 2. Note the continuous fractures that strike east-northeast. They range in length from 10 meters upward to more than 200 meters. These fractures are believed to be the same set that shows sinistral shear in the lake side bedrock exposures.

planar morphology of two joint sets produced a "diamond" shaped pattern in the bedding planes (Figs. 5B, 5C). Few plumose structures were observed on joint surface in the Oswego Formation, however, there are few favorable exposures to view joint surfaces, and many are severely weathered. On the contrary, the exposures of the Pulaski Formation exhibit excellent plumose structures. Some well exposed joint surfaces exhibit irregular jagged, or "stair-shaped" asperities (Fig. 5D). Systematic sampling of hundreds of joint orientations from the Lake Ontario shoreline reveals the consistency of joint attitude in the Oswego area. Rose diagrams were produced to portray the strike of the joints, and the obvious joint populations that are viewed at the outcrop dominate the plots, however, the attitude of the joints varies about 10^{o} to 20^{o} from the Oswego shoreline and in the Oswego River bed (Fig. 4). A third population of joints with a generally N-S strike occurs in the rose diagrams, but are less apparent at most outcrops. The trend of joint traces interpreted from high-resolution air photographs (NYS GIS Clearinghouse) taken immediately offshore at SUNY-Oswego show nearly the same pattern that is observed at the outcrop, with the exception that the individual joints are much larger than the average shoreline exposures. These data was collected using a scale-calibrated image in a computer-mapping program (Canvas GIS), where interpreted joint traces were quantified for attitude and length. A map of the interpreted joint traces appears in Fig. 3. The interpreted joint traces were sorted according to length and attitude, and the rose diagrams of Fig. 6 show the systematic variation. Most very long joints are parallel to the east-northeast striking set.

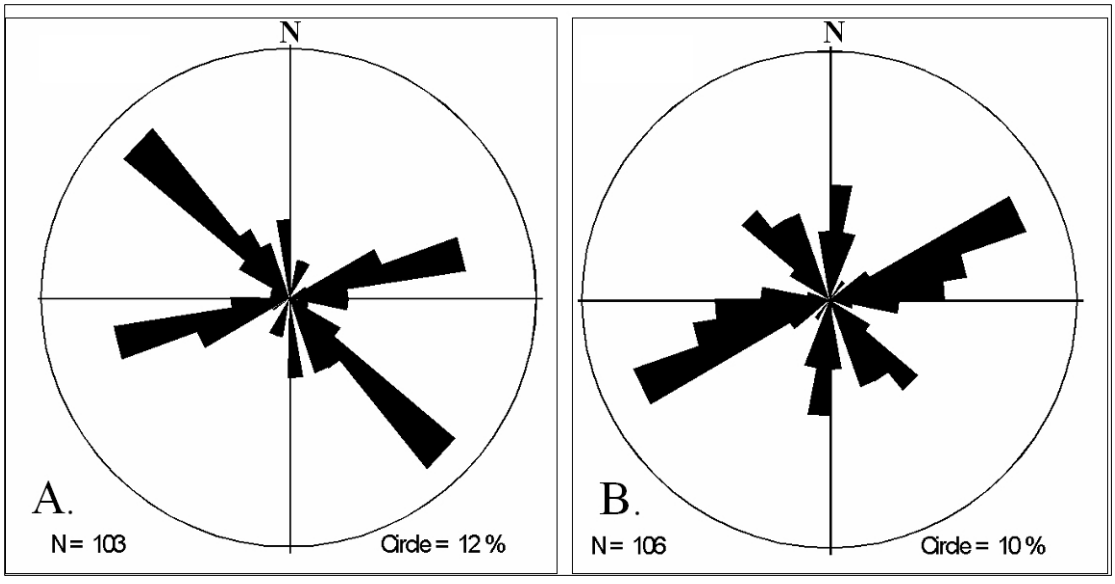


Fig. 4: Rose diagrams for joints that occur in the Oswego Formation in the area of Oswego, NY. A. A composit of shore line outcrops along Lake Ontario at SUNY-Oswego; B. Pavement outcrops that occur in the bed of the Oswego River south of the Utica Street Bridge (these exposures are only available during the lowest water times and will not be viewed during this trip).

Fig. 5: Outcrop photographs of the Oswego Formation on Lake Ontario. A. Meter-scale crossbeds in sandstone. B. Bedding plane exposure with multiple joint sets. The view is looking southeast. C. Two intersecting joint sets. View is looking easterly. D. Close up view of the NW striking joint showing the irregular joint surface.

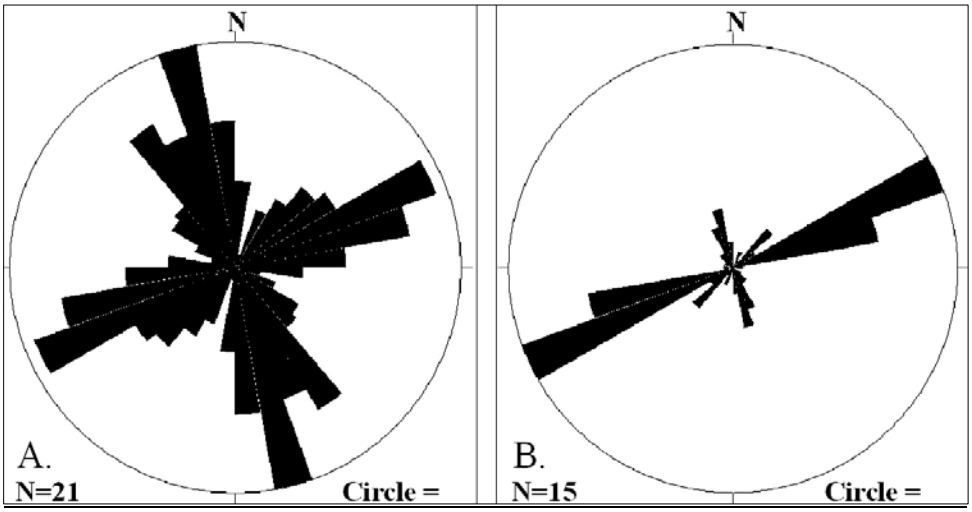


Fig. 6: Rose diagrams for joint traces interpreted from the high-resolution air photograph of Fig. 3. The scale of the image was calibrated in a computer mapping program, and the length of joint traces were measured in addition to the attitude. A. Joint traces less than 10 meters in length; B. Joint traces greater than 10 meters in length.

II.c. Evidence for sinistral shear

En-echelon fracture zones have been found in the Oswego and Pulaski Formations. The individual fractures that make up the zones strike about 25 ^{o} to 30 ^{o} anticlockwise to the general east-northeast strike of the fracture zone. The boundaries of these en-echelon zones are roughly parallel to the joints that strike east-northeast. In some cases, the trace of an individual fractures terminates by curving into the leading en-echelon fracture in a zone. The orientation of these en-echelon domains, and the distribution of fractures relative to each other is consistent with a component of left-lateral shear. Often these en-echelon domains are confined to discrete beds in the Oswego Formation, and terminate where the intersection with the bedding plane. Although the en-echelon fracture zones have been seen throughout the study region, they are best developed in the Oswego Formation along the shoreline of Lake Ontario.

Small faults have been documented in the Oswego Formation along the Lake Ontario shoreline. These faults strike about 070 ^{o} and are subvertical (Fig. 5B). They are approximately parallel to the dominant east-northeast striking joints and parallel the local en-echelon fracture zones. As well, they are parallel to the longest fractures that were interpreted from the high-resolution air photograph of Fig. 3. Due to limited exposure, the length of these faults has not been accurately determined, however, the longest one observed was traced more than 70 meters parallel to strike. Sometimes these faults occur as discrete breaks in the bedrock, but more often they occur as fracture zones 10 to 20 centimeters wide. These faults appear to cross cut all other joints in the Oswego Formation and the northwest striking set have been used as offset markers that document a left lateral slip history (Fig. 5B). The most amount of displacement inferred for any one of these faults was about 1 meter, and the average displacement is about 20 centimeters. The dominant joints appear to be mutually intersecting at most outcrops with no apparent displacement, however, in the vicinity of these minor faults, both sets of fractures show evidence for left lateral displacement of the other joint set. Fig. 7 shows two examples of northwest striking joints with apparent left lateral displacement of the east-northeast striking joints.

II.d. Conclusions about Oswego joints

From out observations of regional and local joint patterns along the southeastern Lake Ontario shoreline and in the Salmon River gorge, there are joints related to those described by earlier researchers in central New York. Specifically, the northwest striking set of subvertical joints is probably related to the late Paleozoic Alleghanian orogeny which produced regional cross fold joints in the Paleozoic strata (Engelder, 1979; Engelder & Geiser, 1980; Zhao & Jacobi, 1997). The other dominant joint set, east-northeast striking, are parallel to joints described by Engelder (1982) and proposed to be associated with the modern stress field in the crust. At most outcrops of the Oswego and Pulaski Formations, it is difficult to determine the relative timing between these two joint sets because they appear to be mutually intersecting. However, the east-northeast striking left lateral en echelon fracture zones and minor left lateral faults appear to displace the apparent Alleghanian joints (NW striking set). Engelder et al. (2001) demonstrated minor displacement on joints

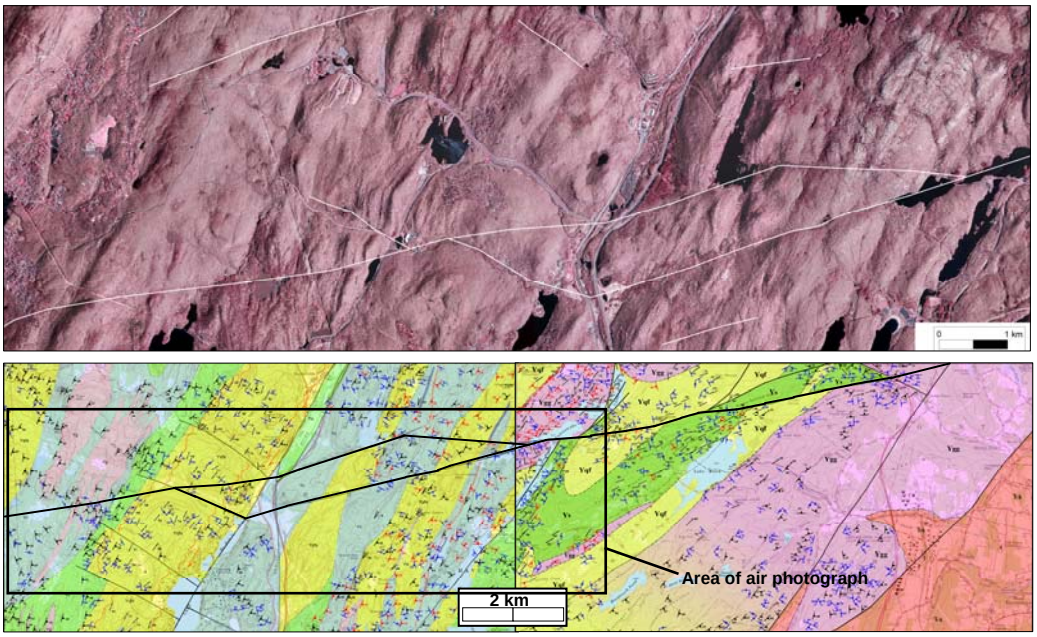


Fig. 7: Preliminary bedrock geologic map for the area of the Hudson Highlands (Gates & Valentino, 2000; 2003) including the Kanawake Fault. The inset on the geology map is the area of the high-resolution air photograph (NY GIS Clearinghouse). The E-W trace of the Kanawake fault is represented by the series of continuous white lines.

in the Finder Lakes region, but attributed it to layer parallel deformation during the Alleghanian orogney. Only in the vicinity of these minor faults do the northwest striking fractures show evidence of shear. We interpret this local shear near the left lateral faults as sympathetic reactivation on older joints, possibly those with a locally favorable orientation relative to the stress field that produced the overall left lateral slip. Engelder (1982) proposed that the origin of the east-northeast striking joints could be the result of the contemporary stress field in the lithosphere. If these joints are directly associated with the en-echelon fracture zones and minor left lateral faults, then the local stress field should be consistent with this type of displacement. Overcoring data were collected for Ninemile Pont during the construction of the nuclear power stations (Dames & Moore, 1978). Engelder & Geiser (1980) compiled these data and show that the maximum horizontal compressive stress trends northeast-southwest. The orientation of te inferred stress would be consistent with left lateral slip on east-northeast striking failure surfaces.

III. Joints and Faults in the Hudson Highlands, New York

In conjunction with bedrock mapping for the StateMap program in NY State, joint maps, including joint attitude and joint density for multiple sets were produced for the Monroe, Sloatsburg, Thiells, Popolopen Lake and Greenwood Lake 7.5' quadrangles, located in the Hudson Highlands, southern New York State. This area of the Hudson Highlands is mostly underlain by high-grade metamorphic and igneous rocks, and experienced a long history of ductile deformation associated with the Grenville orogenic cycle (Gates et al., 2004). Joint

orientations are highly consistent across the entire Grenville massif of the western Hudson Highlands and some joint sets correlate with mapped faults and topographic lineaments. Minor joints that are parallel to foliation occur in all bedrock units. The general strike of these joints is northeast-southwest, and the dip is variable depending on the attitude of the local foliation. Although most topographic lineaments are parallel to, and can be correlated with bedrock units, there are many, that range in length from 1 to 5 km, that cross the regional bedrock structure. These lineaments are most pronounced west of the Ramapo Fault in the Sterling Forest, but a few transect the valley between the western and eastern structural blocks that form the Hudson Highlands. These lineaments extend for 10's of kilometers, and correlate with faults that offset the bedrock Proterozoic structure. This section of the paper summarizes the relationship between outcrop joints and faults, joint intensity, and patterns observed using high-resolution imagery.

III.a. Bedrock geology: a summary

Previous mapping in the Hudson Highalnds area (Figs. 8 & 9), divided the bedrock units by discrete rock types. Gundersen (1986) suggested that lithologic and stratigraphic associations and rock sequences should be grouped as units in a kind of sequence stratigraphy or lithofacies for metamorphic rocks. This system of grouping rock types was adopted for the latest round of geologic maps by Gates et al. (1999), hence, the lithofacies described below constitute the bedrock of the central Hudson Highlands structural blocks.

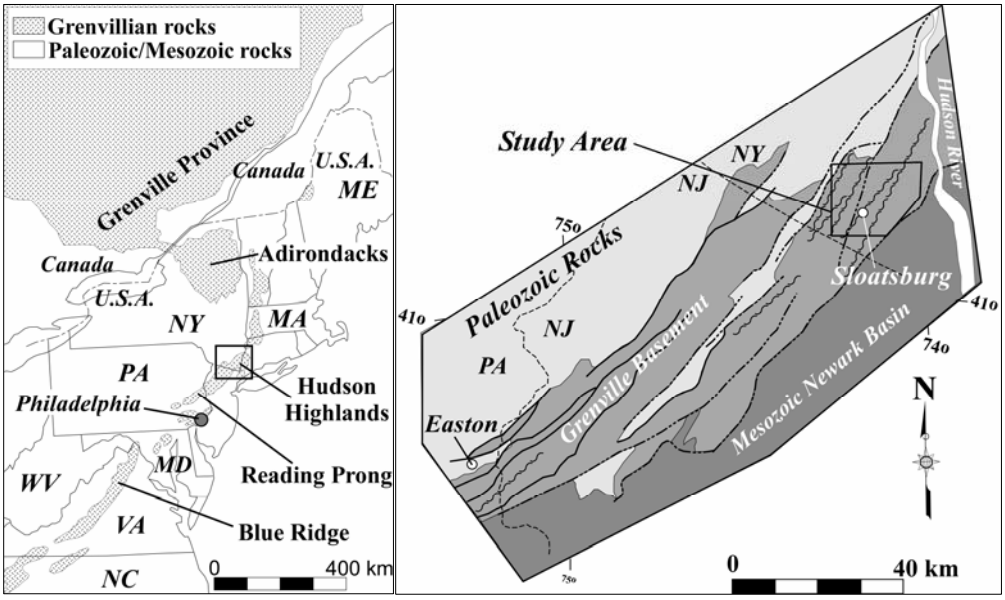


Fig 8: General geologic map showing the location of Proterozoic basement massifs in the Appalachians and a general geologic map of the Hudson-New Jersey Highlands.

The Highlands Province of the north central Appalachians is also known as the Reading Prong and extends from Pennsylvania to Connecticut (Fig. 8). It is composed of crystalline Mesooproterozoic Grenville-age rocks and as such forms a link between the

larger Blue Ridge Province to the south and the Berkshires-Green Mountains province to the north. The Highlands are subdivided into western, central and eastern blocks by the Reservoir Fault to the west and the Ramapo Fault to the east. The western Highlands are variably overlain by and disappear beneath the Paleozoic sedimentary rocks of the Valley and Ridge to the west. The eastern Highlands are variably overlain by and tectonically interleaved with Paleozoic sedimentary and metasedimentary rocks of the Manhattan Prong and Taconics to the east and variably overlain by Mesozoic sedimentary and volcanic rocks to the southeast. The central Highlands overthrust and are overlain by early Paleozoic sedimentary rocks to the north and northwest. The Hudson Highlands form the central Highlands within southeastern New York State, primarily within the Hudson River valley north of New York City.

There are numerous geologic maps of the Hudson Highlands that attempted to map every lithology separately. These maps are very complex and do not reveal the structural and stratigraphic relations clearly. The main problem is that it is difficult to even determine if a highly deformed granulite is a metasediment or a metaigneous rock much less detailed relations. In an attempt to resolve these problems, a new mapping project has been underway since 1997, that considers the area in a sequence stratigraphy approach. Three primary lithofacies have been identified with several intrusive units.

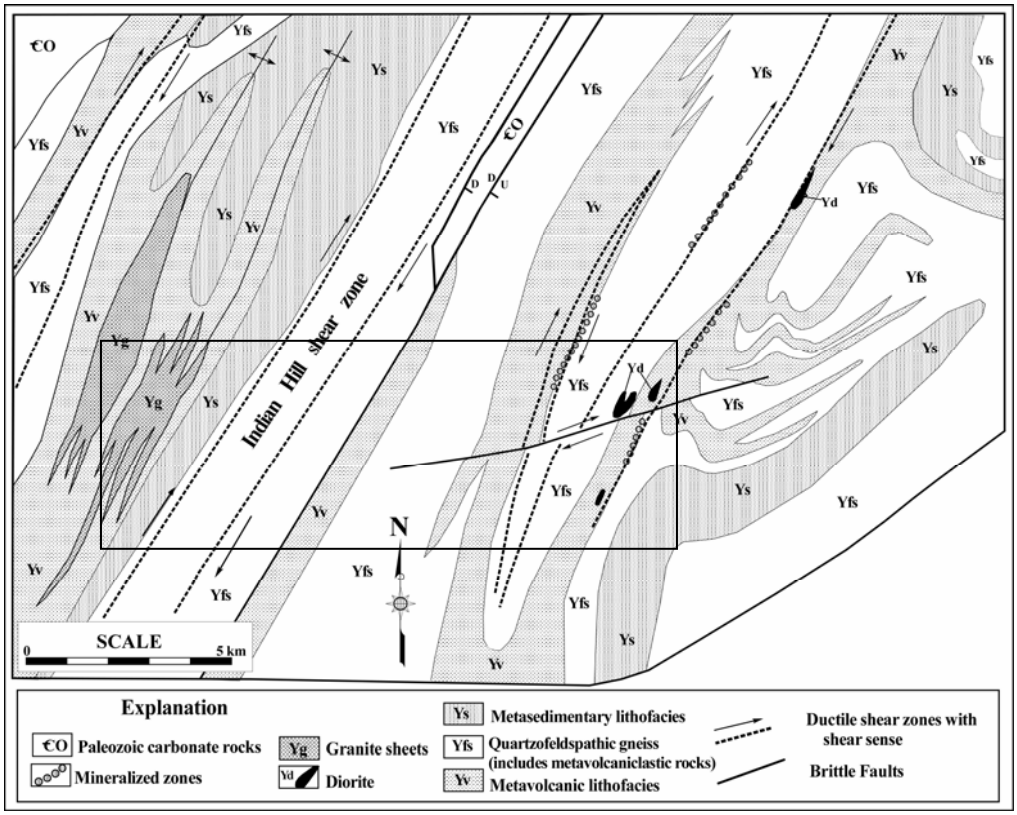


Fig. 9: Generalized bedrock geology map of the Hudson Highlands, New York (modified from Gates et al., 2004). The geographic distribution of the three major lithofacies and general structural trends are represented. The rectangle shows the approximate location of Fig. 11.

III.a.i. Metavolcaniclastic lithofacies.

A discontinuous sequence of medium to coarse grained massive to layered (10 cm-2 m) quartzofeldspathic gray gneiss is interepreted as the metavolcaniclastic lithofacies. These rocks contain quartz, plagioclase and varying amounts of hornblende, biotite, garnet and K-feldspar as a function of the layering. They are locally interlayered with, and in gradational contact with metavolcanic rocks. Thinly interlayered K-feldspar-pyroxene calc-silicates are also present in this lithofacies.

III.a.ii. Metavolcanic lithofacies.

Numerous blocks of thin to thickly banded black and white gneiss are interpreted as the metavolcanic lithofacies. Mafic, intermediate and felsic compostion layering occurs at the sub- to meters scale. Mafic layers contain pyroxene, hornblende, magnetite and local biotite. Quartz and plagioclase comprise the felsic layers with minor mafic phases previously listed. These rocks are locally interlayered with calc-silicate gneiss containing diopside, K-feldspar and quartz, and gradational contacts with the metavolcanic lithofacies.

III.a.iii Metasedimentary lithofacies.

The metasedimentary lithofacies is inferred from an interlayered group of gneisses including biotite-sillimanite-garnet metapelitic gneiss with cordierite locally, quartzite with minor biotite and/or garnet, K-feldspar-pyroxene calc-silicate, graphite-sulfide gneiss with garnet, sillimanite, biotite, quartz and feldspar and rare thin lenses of calcite marble. Typically, the contact with the metavolcaniclastic and metavolcanic lithofacies is sharp with no obvious gradation as in the other lithofacies.

The three rock units previously described were involved in two main tectonic episodes. The first event is the most pervasive of the two, and reflects the main Grenville orogeny in the Highlands terrane. The rocks were deformed into tight recumbent west-directed fold nappes with subhorizontal, highly attenuated lower limbs and shallowly dipping, more intact upper limbs. This nappe propagation was accompanied by the granulite facies metamorphic conditions, and the age of this event appears to be about 1060 to 1020 Ma (Gates et al., 2004). The second event was a large scale dextral strike-slip episode. It began after the intrusion of the 1008 Ma Tiorati diorite and continued until sometime prior to 981 Ma (Gates et al., 2004; Price, 2005).

There were several periods of igneous intrusion. They include pre- to synkinematic granites and granite sheets, a 1008 Ma diorite-granite suite that intruded after the first tectonic event but at the initiation of the second tectonic event, and coarse pegmatites that intruded at 981 Ma (Gates et al., in review).

III.b. Joints and joint patterns

There are two dominant joint sets in the Hudson Highlands with other localized sets that appear to be associated with faults. The most dominant joint set is characterized as steeply dipping to subvertical, strikes northeast-southwest, and can be found at almost every bedrock exposure. The general joint density ranges from 3-6 per meter counted

perpendicular to the joint strike. The strike of these joints generally parallels the foliation in the bedrock. A second set of very common subvertical joints strikes east-northeast (Fig. 10). This joint set appears to be more localized, but occurs in bedrock that is near long

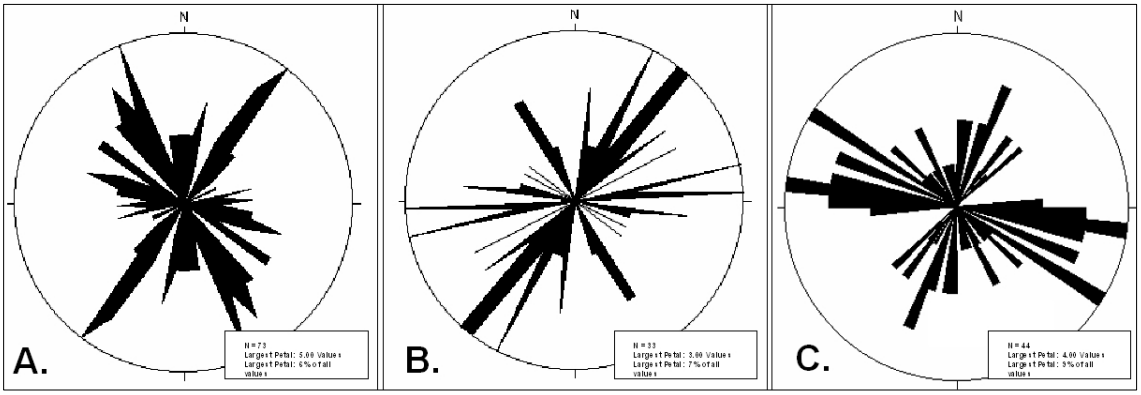


Fig. 10: Rose diagrams of joints from three locations in the Hudson Highlands. A. Sterling-Tuxedo Lake area; B. Indian Hill area; C. Torne valley area.

topographic lineaments that also transect the regional structural grain of the Hudson Highlands. A third set of subvertical joints strikes west-northwest, and appear to be linked to faults with the same attitude that show oblique-dextral displacement (north-side-up) (Figs. 10 & 11). Joint surface morphology is highly varied. The joint set that is generally parallel to the regional foliation appear to be Mode I, with no apparent displacement (Fig. 12). However, the second and third joint set commonly exhibits minor displacement and mineralized slick surfaces (Figs. 12D).

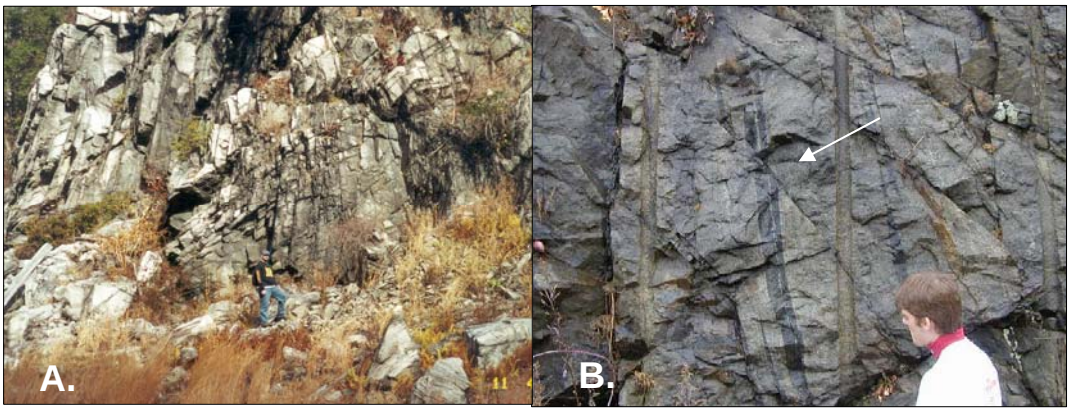


Fig. 11: Photographs showing examples of densely spaced joints. A. Metavolcanic lithofacies along the Kanawake fault. Two joints sets are apparent. The one inclinded to the left is roughly parallel to the location foliation. The set that is subvertical is subparallel to the Kanawake fault. B. Joints with oblique-reverse displacement in quartzofeldspathic gneiss with minor amphiblite layers. The white arrow points to amphibolite layers offset across a joint.

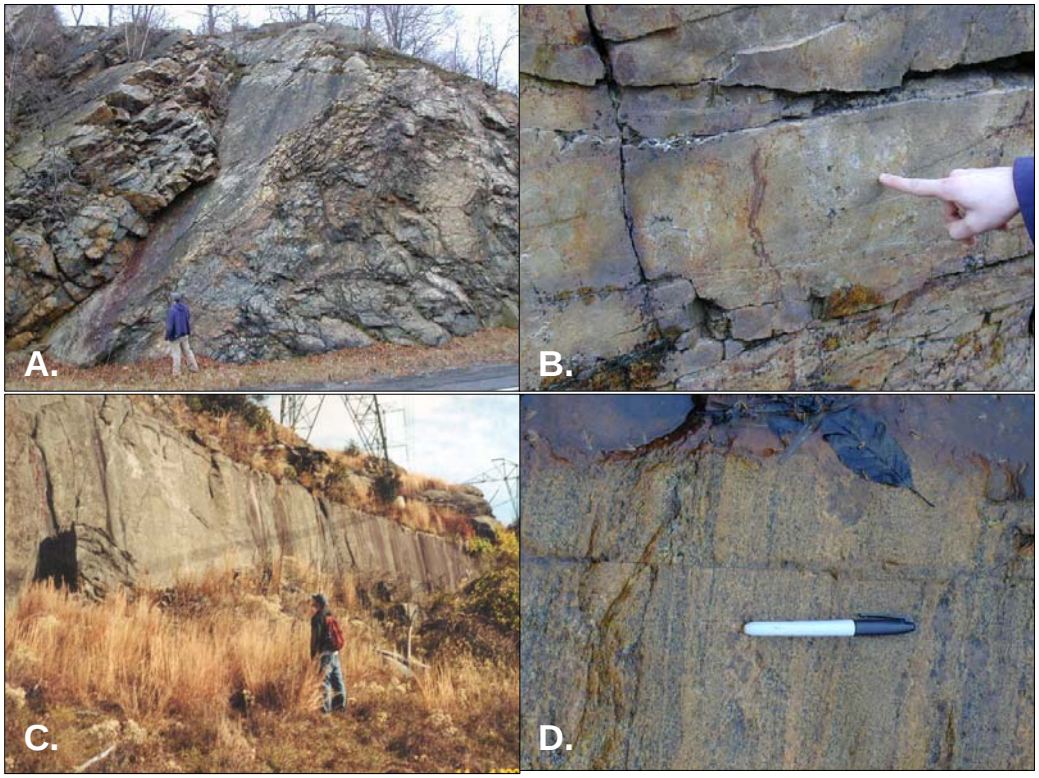


Fig. 12. Photographs of joints feature and minor faults in the Hudson Highlands. A. Fault surface in the gneiss of the metasedimentary lithofacies, Seven Lakes Drive, Harriman State Park; B. Close up view of a highly polished, mineralized joint surface at the same location as A.; C. Northeast striking, apparent fault scarp in quartzofeldspathic gneiss near the New York Thruway; D. Joint with minor dextral displacement in pelitic gneiss in Torne Valley.

III.c. Active (?) faults in the Hudson Highlands

Using the bedrock geology as a marker, some of the faults in the central Hudson Highlands have up to 2 km offset. These faults also correlate with strong topographic expression (Fig. 6). One of the more pronounced faults traces eastward across the Sterling Forest into Harriman State Park, and through Lake Kanawake (Kanawake Fault). This fault displaces major geologic units. In the area of north of Tuxedo, NY, the fault jogs northwestward and merges with another east-northeast striking fault that cuts across the Sterling Forest area. Although this region is highly glaciated, some of the Kanawake Fault segments trace through fresh talus slopes. As well, this fault traces into a region of modern seismicity suggesting the fault has minor activity (Fig. 8). Seismicity data for the region was obtained from the Lamont-Doherty Cooperative Seismic Network database (Kim, W. Y.-personal communication). Of the numerous earthquake events compiled for the New York City metro area, on that occurred on September 3, 1951 lies approximately along the eastern extension of the Kanawake Fault in Rockland County. This earthquake had a Richter (ML) magnitude of 3.6 and a maximum intensity (MM) of V, and is cited as the eleventh largest

earthquake in this region. Some of the more notable earthquakes in the area are apparently association with Mesozoic faults such as the Ramapo Fault. However, the joints and the faults associated with the trend of the Kanawake fault appear to be the youngest structural features and may reflect current stress fields.

IV. References

Dames & Moore (1978). Geologic investigation: Ninemile Point nuclear station unit 2, Doc. 04707-125-19 to Niagara Mohowk Power Corporation, Syracuse, New York.

Engelder, T. (1979). The nature of deformation within the outer limits of the central Appalachian forland fold and thrust belt in New York state, *Tectonophysics*, 55, 289-310.

Engelder, T. (1982). Is there a genetic relationship between selected regional joints and contemporary stress within the lithosphere of North America?, *Tectonics*, 1, 161-177.

Engelder, T. (1985). Loading paths of joint propagation during a tectonic cycle: an example from the Appalachian Plateau, U.S.A., *Journal of Structural Geology*, 7, 459-476.

Engelder, T. & Geiser, P. A. (1980). On the use of regional joint sets as trajectories of paleostress fields during the development of Appalachian Plateau, New York, *Journal of Geophysical Research*, 85, 6319-6341.

Engelder, T., Haith, B. F. & Yones, A. (2001). Horizontal slip along Alleghanian joints of the Appalachian Plateau; evidence showing that mild penetrative strain does little to change the pristine appearance of early joints, *Tectonophysics*, 336, 31-41.

Evans, K. F., Engelder, T. & Plumb, R. A. (1989). Appalachian stress study 1. A detailed description of insitu stress variations in Devonian shales of the Appalachian Plateau, *Journal of Geophysical Research*, 94, 1729-1754.

Fisher, J. A. & Laird, H. S. (1978). Engineering geology at Ninemine Point, New York, New York State Geological Association Guidebook, 50th Annual Meeting, Syracuse, NY, p. 257-272.

Gates, A. E., Valentino, D. W. & Krol, M. A. (1999). The Grenville Province in the western Hudson Highlands, New York, Friends of the Grenville Field Conference Guidebook, Montreal, Canada.

Gates, A. E. & Valentino, D. W. (2000). Bedrock geologic map of the Sloatsburg 7.5 Minute Quadrangle, New York, New York Geological Survey, Open-file Report.

Gates, A. E. & Valentino, D. W. (2003). Bedrock geologic map of the Thiels 7.5 Minute Quadrangle, New York, New York Geological Survey, Open-file Report.

Gates, A. E., Valentino, D. W., Chiarenzelli, J. R., Solar, G. S. & Hamilton, M. A. (2004). Exhumed Himalayan-type syntaxis in the Grenville Orogen, northeastern Laurentia, *Journal of Geodynamics*, 37, 337-359.

Gundersen, L. C. (1986). Geology and geochemistry of the Precambrian rocks of the Reading Prong, New York and New Jersey; implications for the genesis of iron-uranium-rare earth deposits, U.S. Geological Survey Circular 19.

Patchen, D. G. (1978). Depositional environments of the Oswego sandstone (Upper Ordovician), Oswego County, New York, New York State Geological Association Guidebook, 50th Annual Meeting, Syracuse, NY, p. 368-385.

Price, R. P. (2005). Age constraints for granitoids in the Hudson Highlands using SHRIMP U-Pb on zircons, M.S. Thesis, Rutgers University, Newark, NJ.

Twiss, R. J. & Moores, E. M. (2006). Structural Geology, Freeman.

Younes, A. I. & Engelder, T. (1999). Fringe cracks; key structures for the interpretation of the progressive Alleghanian deformation of the Appalachian Plateau, *Geological Society of America Bulletin*, 111, 219-239.

Zhao, M. & Jacobi, R. D. (1997). Formation of regional cross-fold joints in the northern Appalachian Plateau, *Journal of Structural Geology*, 19, 817-834.

GREEN OSWEGO

J. Dorsey, A. Hazar, L. Lee, & M. K. Morris
Marketing and Management

The purpose of this paper is to discuss environmental issues concerning Oswego State University. These issues include current commitments we are undertaking and changes that have already been made. It will also include remedies other schools are undertaking that we could consider in the future.

One of the things that we found during the course of our research was the American College and University Presidents Climate Commitment, or the ACUPCC (http://www.presidentsclimatecommitment.org/). Deborah Stanley, our own President here at Oswego, was one of the charter signatories. We asked ourselves, what does this commitment mean? One of the things we committed to by signing it was initiating the development of a comprehensive plan to achieve climate neutrality as soon as possible. Within two months of signing this document in September 2007, we agreed to create institutional structures to guide the development and implementation of the plan. Within one year of signing this document, all signatories were to complete a comprehensive inventory of all greenhouse gas emissions (including emissions from electricity, heating, commuting, and air travel) and update the inventory every other year thereafter. Within two years of signing this document, we would develop an institutional action plan for becoming climate neutral. We were also to initiate two or more of the following tangible actions to reduce greenhouse gases while the more comprehensive plan is being developed: establish a policy that all new campus construction will be built to at least the U.S. Green Building Council's LEED Silver standard or equivalent; adopt an energy-efficient appliance purchasing policy requiring purchase of ENERGY STAR certified products in all areas for which such ratings exist; establish a policy of offsetting all greenhouse gas emissions generated by air travel paid for by our institution; encourage use of and provide access to public transportation for all faculty, staff, students and visitors at our institution; within one year of signing this document, begin purchasing or producing at least 15% of our institution's electricity consumption from renewable sources; establish a policy or a committee that supports climate and sustainability shareholder proposals at companies where our institution's endowment is invested; and participate in the Waste Minimization component of the national RecycleMania competition, and adopt three or more associated measures to reduce waste. Lastly, we would make the action plan, inventory, and periodic progress reports publicly available by providing them to the Association for the Advancement of Sustainability in Higher Education (AASHE) for posting and dissemination.

SUNY Oswego also participated in Focus the Nation, joining with more than 1,550 other colleges and universities on January 31st, 2008. This was a country-wide day of teach-ins about environmental issues. Among the speakers was Patrick Moore, co-founder and former director of Greenpeace International, who spoke about Nuclear power and how it is the only way to provide the energy the world needs and simultaneously reduce humankind's carbon footprint.

We are already doing a lot here at Oswego State to deal with the environmental issues that we face. For example, we are in the process of switching to an electronic campus, decreasing the impact made by the overwhelming amounts of paper used on college campuses. We are also active in recycling, and we use energy efficient lighting in many of the buildings. We offer Distance learning courses which studies show can cut carbon emissions by two thirds even for part time on campus students, and we have implemented greener building plans and environmentally friendly construction techniques.

In the course of our research we also found some very interesting things other schools are doing, which we could implement here at Oswego to make an even bigger difference. Middlebury College in Vermont touts an environmentally friendly dining hall with a vegetated roof that opened in 2005. UC-Santa Cruz is an institution recognized for reversing the outsource trend by moving from a food service company contract to an in-house operation, and now, instead of using national mega-suppliers, the dining halls buy organic produce grown on local farms, which supports the area economy and reduces transportation time (and thus gas emissions). UCSC also recycles used cooking oil and use paper plates made from recycled materials. The dining operation is also reducing landfill waste by donating leftover food to the "Second Harvest Food Bank." When UCSC started buying fresh, locally grown organic produce the local farmer's organization ALBA's output increased by 300 percent in a single month. At Case Western Reserve University, a new housing complex will feature a kiosk that displays statistics on the buildings energy use. The University of South Carolina uses recycled materials (right down to the carpets) and keeps energy use to a minimum through good engineering and usage changes. Think lots of natural light, low-flow faucets, and motion-detected lighting. This is the way we need to go in all buildings.

There is also the impressive Northland College Environmental living and learning center (ELLC). Living arrangements feature three styles: regular double rooms, suites, and apartments. It also includes nine lounge or study areas, and four different energy and waste management systems. Among the special environmental features is a 120-foot 20 kilowatt wind tower. Three photovoltaic arrays will provide efficient active solar energy collection and help study the efficiency -- one array is stationary, a second one tracks the sun's path horizontally, and the third tracks both horizontally and vertically to maximize solar gain. Fourteen solar panels placed on the roof will preheat hot water for use by residents. Composting waterless toilets in two of the apartments will provide a demonstration of their function and efficiency. The apartments have passive solar design and share two greenhouses open to student use. The school also utilizes Heat recovery units in the building's ventilation system. They have high efficiency natural gas w/ solar preheating and use green building materials in the construction. The environmental

features of the ELLC result in a 50% greater efficiency than a similar building designed to code. Students living in the dorm said it made them more ecology-conscious. This is an excellent example of where we could go with our school in the future.

There are also some simple things that students themselves can do. First of all, we can recycle everything, especially paper! Think about how many pieces of paper a student can go through in a semester. There are class notes, scrap copies, term papers, daily school newspapers, and assorted stuff that has been printed out from the internet. It all adds up. Look for a paper recycle bin on campus -- they are often near a dormitory or large classroom building. And we can recycle other items as well, including cans, bottles, and cardboard boxes. We can also use our printers more carefully. We can save paper by printing on both sides of the page. Many professors don't mind if a student turns in a paper like this -- we just need to ask first. Students can save pages that have been printed and use the backs to print out drafts and other things that don't need to be turned in. In addition, many printers have multiple settings for print quality. Use the high quality print setting for things that have to look nice, but use the low quality setting for things that don't. This will save ink. While we're at it, we can consider cutting down on the things we print out. Do we really need to print out that web page, or can we just bookmark it? Another thing students can do is limit the use of disposable cups and plates, and the use of paper napkins. If a student moves into his or her first off-campus apartment, it can be tempting to buy disposable cups and plates to save time. This adds up to a lot of waste *and* money. He or she can buy some inexpensive plates and wash them. This also works in a dorm setting. Many dorms have a kitchen, and if it doesn't, students can wash dishes in the bathroom sink. Also, since college students eat a good deal of fast food, napkin use can add up. One napkin will usually do the trick, instead of using a whole pile of them.

There are a lot of other things students themselves can do to reduce their impact on our environment. You can use compact fluorescent light bulbs. These bulbs cost more, but they last longer and ultimately save you money. If you live in a dorm, get yourself a lamp and screw in one of these bulbs. Lamp light is much more pleasant and environmentally efficient than overhead dorm lighting. Another thing you can do is walk, bike, and limit your use of a car. Most campuses are very pedestrian friendly, and many college towns offer good public transportation and bike paths. Ask yourself if you really need a car as a college student, because if you can get by without one, you can save a good deal of money on gas, repairs, and overpriced student car insurance. If you do own a car, try to use it as little as you can. Buy recycled products whenever you can, especially paper. Buy environmentally safe cleaning products as well. Some of these products cost more-- but many don't, or the price difference is negligible. Students can also carry a water bottle instead of drinking from disposable plastic bottles. Think of how many bottles of water get consumed on a campus every day. Save waste and money and carry a refillable bottle. If the tap water on your campus is questionable, buy large containers of waters to refill your bottle. Use refillable binders, not disposable notebooks. This is a simple way to save waste. If you want to save your notes after the semester is over, take them out of the binder and staple them. Or you can go electronic and take all of your notes on a laptop. Lastly, students can buy, sell, and donate at second hand stores. Lots of students do this to save

money, but it's also a great thing to do for the environment. Reusing clothes decreases the use of resources to make clothing and puts a dent in the problem of worldwide sweatshops. All these small sacrifices can add up, especially at Oswego State.

Also, something you don't think about much, but there are Green mutual funds out there, ones that not only avoid companies that harm the environment but actively pursue investment with companies that make a positive impact on the world. Not only should our University look into this type of investing but all individuals can too.

In conclusion, we at Oswego are doing a great job in our opinion, and we appear to be at the forefront of schools trying to make a difference. We need to make more advances in the areas of Environmental education by implementing course work in Photovoltaic Systems and the National Electrical Code, Solar Hot Water Education, Wind Energy Science, Criterion-Referenced Testing and Nuts and Bolts of Ethanol and Biofuels-Hydrogen and Fuel Cells. I got this directly from Renewable energy online's article "Educating the Green Workforce." But that is what we need. More challenging course work involving these fields would be instrumental in making advancements in the green movement, bringing more students who are actively involved in the changes that must be made to our campus and our world. Many would wonder why here? According to Bartleby of Associated Content it is because "campuses are diverse in their development: they combine residential components like dorms and dining halls with offices, classrooms, laboratories, sports venues, theatres, and all kinds of other spaces." They are like models of a city or even of a country. We can and do learn so much from what we have here. Universities are also where many people learn how to live on their own, forming habits that will last throughout their lives. That is if they are good habits and being green can be part of that.

We would like to offer a quote in closing.

"Sustainability is the art of ever after, assuring that people in the future will have what they need to lead fulfilling lives."

Impact of New Product Development of Store Brands on Branded Manufacturers in the U.S. Grocery Industry.

N. G. Leader & I. D. H. Cuthill
Marketing Department, School of Business

Store brands of U.S. packaged food products have made considerable strides in the last 20 years. Readily available technology has enabled manufacturers and retailers to understand consumer trends and buying habits and quickly offer private label substitutes. U.S. consumers recognizing value are more willing to buy private label, and the level of penetration is approaching once lofty European share levels. This has branded manufacturers looking harder at NPD as a weapon to fend off private label and grow category share. This paper will explore the process of new product development in private label and the role private label plays for manufacturers, retailers, consumers and brands. Further a research proposal will be outlined to determine the impact of private label growth on branded NPD strategies.

I. Introduction

The proportion of private label (PL) food products on U. S. supermarket shelves has been growing steadily for the last twenty years at the expense of national brands (NB) until it is now on average across categories about 16% of total supermarket sales (PLMA 2006). In the past year, more than 41% of consumers bought PL (PLMA 2007). There is only a finite amount of shelf space in supermarkets, so as retailers have allocated more shelf space to their own brands, it has become more difficult for manufacturers to get distribution for both existing and new products (Amrouche & Zaccour, 2007). This is expected to continue for the foreseeable future, presenting manufacturers with a dilemma about the best way to compete with their own customers (Amrouche & Zaccour, 2007). Our study will evaluate the moderating factors that govern the extent to which this trend will continue and provide strategic direction for NB marketers on how to deal with it.

Several additional factors have further intensified competitive pressure on manufacturers. The U.S. food market has become increasingly open to foreign firms due to lowering of trade barriers, for example the North American Free Trade Agreement (NAFTA) with Canada and Mexico. The trend towards elimination of trade barriers globally will continue as virtually every country has an interest in the World Trade Organization (WTO). At the same time the U.S. market is attractive to foreign firms because of its size and diversity. Technology has speeded up and automated communications and fostered development of websites that offer 24 hour availability of every conceivable form of merchandise on a worldwide basis. All of these influences have

intensified competition for food marketers in a mature market where total demand is increasing only in proportion to population growth, no more than 3% in dollar volume annually.

Additionally, structural changes in the industry due to improved methods of communication, gathering and analysis of check-out scanner data, and automation of inventory control have made it possible for supermarkets to measure the sales and profit contribution of every stock keeping unit (SKU) in the store on a daily basis. Due to this new-found capability, during the last twenty years the balance of power in food industry distribution channels has shifted from the manufacturer towards the retailer. This is a powerful illustration of the old adage, "Knowledge is Power." Whereas previously only large manufacturers like Procter & Gamble (P&G) or market research firms like A.C. Nielsen had the capability to collate and analyze checkout scanner data, now retailers can do it. They have gone further and turned access to checkout data into a profit center. Wal-Mart used to give scanner data to Nielsen who charged them for analyzing it. Now, they sell it to them. Also, since they have better data sooner than P&G, they have been able to negotiate deals which take into account the savings that P&G can obtain in manufacturing efficiency because of access to Wal-Mart inventory information. The result of this shift in power is that, in order to get distribution, manufacturers now have to convince retailers that a new product introduction will generate at least as much profit as the product it will displace.

Since it is impossible to forecast exactly the potential demand for a new product, retailers also demand a "slotting allowance" which may run as high as $40,000 per store, before agreeing to carry a new product (Desiraju, 2001). Some manufacturers have claimed that this is a restraint of trade and particularly a barrier to entry for small manufacturers. Several Congressional committees have pondered whether or not to outlaw the practice, so far without any result. It is evident, however, that retailers' own new brands have a competitive pricing advantage in not having to carry the added cost of a slotting allowance

Branded packaged food products have been marketed in the U.S. for more than 100 years by manufacturers who were much quicker than grocery stores and supermarkets to see the benefits of branding to marketing strategy. However, with their new-found sophistication in marketing, retailers have also come to realize that branding their own food products is a good way to differentiate themselves from other retailers. (Tarnowski, 2007) In this respect, European food retailers such as Sainsbury in the U.K., and Ahold in the Netherlands have been much quicker to adopt a branding strategy for their own label packaged goods. It is estimated that on average, the proportion of private label in U.S. supermarkets is around 16%, while in Europe it averages around 40% and in the UK chain, Sainsbury's, it is as high as 65% (Harcar et al., 2006; PLMA 2006). Loblaw's President's Choice PL line is so popular that it is licensed to non-competing retailers and is distributed internationally. Wal-Mart, the largest food retailer in the U.S. has its Great Value brand, and Target has its Archer Farms label.

Perhaps one of the most disappointing revelations of recent studies is that the heavy users of PL products are middle and upper middle income consumers (PLMA 2006).

The proportion of relatively affluent consumers who shop at discount stores but who also demand quality has increased (Nunes, Johnson & Breene, 2004). Astute retailers like Target have based their marketing strategy on catering to this segment and Wal-Mart is trying to follow suit. The most brand-loyal consumers tend to be lower income and ethnic minorities.

Few food retailers possess their own manufacturing operations. Exceptions include Kroger which supplies around 4,300 SKUs from 41 manufacturing plants for distribution to its own and other stores. Retailer and wholesaler-sponsored co-ops also have limited manufacturing capabilities. Most depend upon contract manufacturers who produce and package PL products to their specifications. In some cases, manufacturers of branded products also produce private label products. (Table 1) This may seem paradoxical behavior when private label products are directly in competition with them. However, they do it when it allows them to reduce their cost of goods although the profit margins for manufacturers on private label products are typically about half that of branded products. There are also competitive advantages to supplying PL products that will be discussed later.

Table 1 - Wal-Mart Share of Consumer Package Goods Companies' 2004 Sales

Company	Global Sales ($ millions)	% Sales to Wal-Mart
Procter & Gamble	56,741	16
General Mills	11,244	16
Kellogg	9,614	14
Kraft	32, 168	14

Source: Kumar & Steenkamp, Brand Versus Brand (2007)

The conventional wisdom is that failure rates of new food products are typically around 90%. Scholarly studies have shown that it actually varies from 33 – 50%, depending upon the uniqueness of the product and the nature of the product category (Crawford & Di Benedetto, 2008). One of the factors that increases the likelihood of failure is the similarity of the new product to others already on the market.(Crawford & Di Benedetto, 2008). In this respect, PLs are mostly "me-too" products. However, they will always be guaranteed distribution, an advantage not enjoyed by national brands. As the proportion of private label grows, manufacturers will be competing for less and less shelf-space in order to get distribution and will be pressed to introduce new products that are well-differentiated from PL products.

Store brands are priced lower than equivalent national brands by about 20 – 30% (Mendez 2008, Tortola 2007). Until recently, they relied mostly on price to differentiate themselves and did not apply the same marketing methods as those used by national brands. However, that has changed (Bell 2008). Consolidation in manufacturing has led some marketing managers to cross over into the retail side of the business. The result is that PL products have improved in product and packaging quality and now compare favorably with NBs. It follows that as private label grows, manufacturers will be forced to develop more innovative products in order to get distribution and maintain their market share.

It should also be mentioned that the cross –category average proportion of store brands does not reveal the complete competitive picture since it varies widely by product category. For example, in 2006 the average PL market share of milk was 58.7 %. For packaged frozen entrees it was 2.4% and for chewing gum only 0.7% (Table 2).

Table 2 – Retail Grocery Market Share of U.S. Private Label Food Products

Category	PL % Share of Market	PL % Annual Growth
Top Ten High Share		
Milk	58.7	+ 5.8
Butter	45.5	+20.1
Frozen Vegetables	45.5	- 1.0
Frozen Seafood	38.2	+6.2
Natural Cheese	36.1	+ 7.1
Snack Nuts	29.6	+ 25.6
Pickles/Relish	28.3	-1.8
Frozen poultry	24.3	+6.1
Bottled Water	21.1	+10.1
Pasta	20.1	-6.2
Top Ten Low Share		
Baby Food	0.4	+47.1
Gum	0.7	-5.5
Chocolate	1.2	+ 9.1
Frozen Dinners	2.4	-5.2
Granola Bars	3.8	+0.3
Salty Snacks	4.7	- 3.5
Prozen Appetizers	4.9	+13.8
Candy	5.7	+7.2
Carbonated Beverages	6.4	+ 6.0
Dressings/Deli	8.1	+ 8.0

Source: Stagnito (2005)

The food industry is often referred to as a “low-tech” industry. However, this is not an accurate description. For example, today’s food industry is actually based upon extremely “high-tech” developments in animal and crop husbandry, food preservation systems, and ingenious high speed packaging systems which are based upon basic scientific discoveries in biology, chemistry, genetic engineering, mathematical modeling, microbiology, nanotechnology, nutrition, physics, physiology, and process engineering. All of this is not apparent to the consumer of the finished product, but without it our current global food distribution system would not exist. National brand manufacturers possess more of the technological know-how required to produce truly innovative new products than do retailers or the contract manufacturers who supply the bulk of the private label trade and who tend to be technology followers. This, along with their marketing expertise and

financial strengths, provide the national brand manufacturers with their most significant competitive edge in developing new products.

II. Literature Review

An extensive literature has developed during the past twenty years, much of it from European scholars, about the relationship between private label and national brands. It covers analyses of the economic basis and nature of consumer buying preferences for PL products and the strategic alternatives available to retailers and manufacturers in dealing with them (Amrouche & Zaccour, 2007; Ashley, 1998; Batra & Sinha, 2000, Bronnenberg et al, 2007; Kumar, 2007). This paper is concerned mostly with the impact of this relationship on new product development practices, and so we review the literature from that standpoint.

As PL market share increases, most researchers conclude that manufacturers should develop more innovative products that are well-differentiated from PLs (Ashley, 1998), Kumar & Steenkamp, 2007). Prior research has indicated that failure rates of new products that are similar to existing products are higher than those that are well differentiated (Crawford & Di Benedetto, 2008). Also, various researchers have suggested that the leading brand in a category can use contract packing to reduce unit manufacturing cost and assist a PL product to fulfill the role of a lower quality flanker NB (Wu and Wang, 2005). As manufacturers lose market share to PLs, it is reasonable to suppose that they will increase their rate of new product launches in order to compensate for the loss. In fact new food and beverage product introductions have continued to increase for the last twenty years. This is undoubtedly due to the stage of the U.S. food industry life cycle which is mature and intensely competitive. However, the number of new product introductions seems greatly disproportionate to the market need. For example, the total number of new packaged food products introduced annually has been more than 10,000 for the past ten years (Table 3). A recent report in Industry Week stated that consumer packaged goods companies increased new product introductions by 25% over the last three years and a similar increase is forecast for 2008 (Industry Week, Feb 28, 2008). The average U.S supermarket carries 30–40,000 stock keeping units (SKUs), while a Wal-Mart Supercenter may stock 100,000, of which around 40,000 are of food products. An SKU can be a different package size of the same brand. Also, many of the products taking up shelf space such as bread, milk, meat, and vegetables are staples that have been in distribution for centuries. According to Marketing Intelligence, Ltd., retail chains are offered around 200 new SKUs each week of which at least 70% are rejected by buying committees. Of those that are accepted, it is often at the store manager's discretion if they will be stocked in his store. On average, as we have said, 16% of the existing stock is private label (Adkawski & Bari, 2004). It is extremely unlikely that a new NB food product being introduced will obtain distribution, let alone become a success. Many of these products may only be seasonal or intended for regional rather than national or international distribution. Nevertheless, the difficulties in new product development faced by national food marketers are apparent and likely to get worse.

Table 3 : New Packaged Food Product Introductions

Year	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**
Total	8420	10641	9866	9770	10320	13451	13397	14596	13840
Top 20 Firms	1549	1713	1718	1566	1562	1562	1935	1377	1430

Source: Marketing Intelligence Service, Ltd., 2000

Most scholars who have studied the dilemma of marketers faced with the prospect of intensifying competition with their retailer and wholesaler customers have advised them to develop value-added products that carry benefits for the target market that cannot be matched by PL. (Harcar et al, 2006; Herstein & Gamliel, 2006; and Kumar & Steenkamp, 2007) They have also suggested that the technology base of NB products should be protected by patent applications or maintenance of trade secrets. (Kumar & Steenkamp, 2007) They have also advocated greater use of advertising to build brand equity and customer loyalty rather than trade promotions which favor a sales or short-term strategy. (Kumar, 2007) Retailers favor promotions because they defray the cost of local advertising and often include allowances that go to their bottom line (Abraham & Lodish, 1990). At the same time research has shown that only 16% of trade promotions actually provide incremental profit to manufacturers (Abraham & Lodish, 1990). All of these suggestions play to the strengths of manufacturers which are in marketing and manufacturing expertise and research and development capability (Lewis, 2006). However, in order to accomplish this they will also have to wean their customers off of their current reliance on trade promotions, which will not be easy (Abraham & Lodish, 1990). If manufacturers follow these strategic recommendations, one would expect the ratio of unique new products to incremental new products to increase.

If, as seems likely, the increased marketing competence of retailers leads them to develop PL line extensions within each category, this will bring them even more directly into competition with NBs (Anonymous, 2007). The large manufacturers, such as P&G, General Mills, and Kraft have recently extended their new product development activities to worldwide external R&D networks including consultants, universities, and contract suppliers (Huston & Sakkab, 2006). Providing they have sufficient resources, the large chains such as Wal-Mart, Safeway, Target and Costco could do the same to compensate for their lack of NPD infrastructure. The British retailers, Tesco and Sainsbury have their own PL premium brands, as have Loblaws in Canada (Anonymous, 2007). In the U.S., Food Lion has launched a food product line differentiated by quality (Daniel, 2007). While this trend may grow in the short term, and despite the fact that Kroger has been able to do so for many years, it is unlikely that retailers would be able to or even want to sustain a heavy NPD program or make the significant investment in plant and equipment required. Some scholars have estimated an optimum ratio of PL as about 50% overall of market share (Kumar & Steenkamp). However, this issue is best looked at on a category basis rather than cross-category as most researchers have done. For example, the PL market share of

commodity type products such as bread, milk, butter and eggs already exceeds 50% (Table 2, Wellman, 2003). It seems likely that there is an inverse relationship between PL market share and product degree of innovativeness (Lewis, 2006). However, if the U.S. market develops in the same way as has European markets, this picture may be changing. For example, Tesco in the U.K., Loblaw in Canada, and Food Lion are already marketing premium PL brands. In fact, Food Lion is using a three tier approach of marketing generic, flanker and premium products

The conventional wisdom in the food industry is that it is difficult to compete as the third or lower market share brand. If market share grows as expected to European or Canadian levels, PL will become at least the third brand in most categories, making it extremely difficult for firms with lower market share to maintain distribution (Harcar et al, 2006; Herrstein & Gamliel, 2006; Ward et al., 2002; Wu & Wang, 2003). Also, consumers like variety in food products but supermarkets are continually pruning their assortment of brands with respect to volume and profit per square foot of shelf space . Consequently it is to be expected that the average number of NBs per store will decline as the marginal market share competitors drop out.

III. Research Questions

1. As PL market share increases, how will this affect the number of NB new product launches?
2. As PL market share increases, how will this affect the success rates of new NB products?
3. As PL market share increases, how will this affect the ratio of radical to incrementally new NB products?
4. As PL market share increases, will retailers introduce less incrementally new and more unique products?
5. As PL market share increases, will this squeeze out tertiary NBs?

IV. Hypotheses

H1: As PL market share increases, the number of new NB market launches will increase.
H2: As PL market share increases, the ratio of NB unique product market launches will increase
H3: As PL market share increases, the ratio of PL unique new product introductions will increase
H4: As PL market share increases, the success rate of NB new product introductions will increase
H5: As PL market share increases, the number of NB brands will decrease.

V. Methodology

It is proposed to conduct a survey of retailers, private label contract manufacturers and national brand manufacturers in order to determine the current role that PL plays in their NPD strategy and future plans with respect to new product development (NPD)

Data will be collected via a self-administered questionnaire using a random sample collected from regional, national, and large NB manufacturers, PLMA members, and retailers with respect to five high market share and five low market share categories of PL products.

Measures will be developed regarding:

- current and planned role of PL in retailer, contract manufacturers and NB manufacturers strategy, e.g. retailer increase in profit contribution per square foot of shelf space; contract manufacturers into proprietary technology; NB manufacturers to reduce unit manufacturing cost and increase pressure on competition
- current and planned NPD practices of Retailers, Dedicated PL Mfrs, and NB Mfrs.
- current and planned brand strategy of retailers, e.g. me-too in direct competition with NB, or differentiated from them.

It is proposed to carry out analysis of the data in order to determine relationships between variables using Correlation Analysis and Regression Analysis. We will also conduct Factor Analysis in order to determine if there are combinations of variables that influence relationships. Regression analysis will be used to model the relationships between variables.

VI. References

Abraham, M. M., & Lodish , L. M. (1990). Getting the Most Out of Advertising & Promotion. *Harvard Business Review,* May-June, Vol 68, Iss 3, 50-57.

Adlawski, K., & Bari, H. (2004) An Empirical Analysis of the Determinants of Retail Margins: The Role of Store Brand Share. Journal of Marketing, 68 (January) 147-165.

Amrouche, N. & Zaccour G. (2007). Shelf-space Allocation of National and Private Brands. *European Journal of Operational Research,* 180 648 – 663.

Anonymous. (2007). Analysis: Anger at Competition Commission's Supermarket Whitewash. *Marketing Week*, Nov. 8, 2007.

Anonymous (2007). $55 billion in Sales Could Shift to Store Brands, McKinsey Study Predicts at Private Label Trade Show in Chicago. *PR Newswire*, New York, Nov. 13, 2007.

Anonymous. (2007). Retailers Across the Country Win Awards for Outstanding Store Brand Products. *PR Newswire*, New York, Oct 2, 2007.

Ashley, S. R. (1998). How to Effectively Compete Against Private-Label Brands. *Journal of Advertising Research,* January*February, 75-82.

Batra R. & Sinha I. (2000). Consumer-level factors moderating the success of private label brands. *Journal of Retailing*, Vol 76(2), 175-191.

Bell, N. M. (2008). Which Can Would You Buy? Store Brands Have Come A Long Way from Old White Label Days. *McClatchy- Tribune Business News*. Washington: Feb 3, 2008.

Bronnenberg. B. J., Dhar S. J. & Dube J-P. (2007). National Brands, Local Branding: Conclusions and Future Research Opportunities. *Journal of Marketing Research*, Vol. XLIV, (February), 26-28.

Burton S., Lichtenstein D. R., Netemeyer R. G. & Garretson J. A. (1998). A Scale for Measuring Attitude toward Private Label Products and an Examination of its Psychological and Behavioral Correlates *Journal of the Academy of Marketing Science*, Volume 26, No. 4, 293-306.

Choi, S. Chan & Coughlan A. T. (2006). Private label positioning: Quality versus feature differentiation from the national brand" *Journal of Retailing*, 82(2) 79-93.

Crawford, C. M., & Di Benedetto, A. (2008). New Products Management, (9th Ed.), New York, McGraw-Hill/Irwin.

Daniel, Fran. Retail Notes: Done Right, Private Label Spells Profit. *Knight Ridder Tribune Business News*, August 19, 2007.

Desiraju, R. (2001). New Product Introductions, Slotting Allowances, and Retailer Discretion. *Journal of Retailing*, 77, 335- 358.

Harcar, T., Kara, A., & Kucukemiroglu, O. (2006). Consumers' Perceived Value and Buying Behavior of Store Brands: An Empirical Investigation. *The Business Review*, Vol. 5, Num. 8, Summer, 55–62.

Herstein R. & Gamliel E. (2006). Striking a Balance with Private Branding. *Business Strategy Review*, Autumn, 39-43.

Huston, L., & Sakkab, N. (2006). Connect & Develop: Inside Procter & Gamble's New Model for Innovation, *Harvard Business Review*, Mar, Vol 84, Iss. 3, 58-66.

Industry Week (2008). CPG Manufacturers Didn't Reduce Inventory Levels in 2007, *Industry Week*, Feb 28, 2008. http://www.industryweek.com/PrintArticle.aspx?ArticleID=15864

Kumar, N. (2007). The Right Way to Fight for Shelf Domination," *Advertising Age*, Jan 22, 2007, Iss. 4, 14-16.

Kumar N. & Steenkamp J-B.(2007) Brand Versus Brand. *International Commerce Review*, Volume 7, No. 1, Spring 47-53.

Kumar, N., & Steenkamp, J-B E. M. (2007) Are Brands Dead? *Chief Executive*, July/August, 40 – 43.

Lewis, H. (2006). What is Happening with Private label in Food? Is the Market Growing at the Expense of Branded Goods? *Just-Food.com*, March, 13–15.

Lewis, H. (2006) Identifying the Top Ten European Retailers' Current Private Label Strategies. *Just-Food.com*, March, 16–20.

Mendez J. L., Oubina, J., & Rubio, N. (2008). Expert Quality Evaluation and Price of Store Vs. Manufacturer Brands: An Analysis of the Spanish Mass Market. *J. Retailing and Consumer Services*, 15, 144-155.

Nunes, P. F. , Johnson, B. A., & Breene, R. T. S. (2004) Selling to the Moneyed Masses, *Harvard Business Review*, Vol 84, Aug/July, Iss 7/8, 94- 104.

PLMA Consumer Research Report, (2006). Star Power: The Growing influence of of Store Brands in the U.S. Ipsos MORI survey for Private Label Manufacturers Association, 2006.

Stagnito Communications (2005). *Supplier Source Book*, Thomson Gale & Goliath (2006)

Tarnowski J. (2007). Thinking Like a Brand. *Progressive Grocer*, 11/01/2007, Vol. 86, Issue 15, 16-22.

Tortola, J. O. (2007) Private Eyes. *Progressive Grocer*, 12/01/07, Vol 86, Issue 17, 14-16.

Ward M. B., Shimshack, J. P., Perloff, J. M. , & Harris, J. M. (2002). Effects of the Private Label Invasion in Food Industries. *American Journal of Agricultural Economics*, 84 (4), 961–973.

Wellman, D. (2005). Frozen in Place. *Frozen Food Age*, April, 53, 9, 18-21.

Wu, C-C., & Wang, C-J. (2005). A Positive Theory of Private Label: A Strategic Role of Private Label in a Duopoly National Brand Market. *Marketing Letters*, 16:2, 143–161.

FUTURE OF RETIREMENT BENEFITS

M. Muniz
Department of Accounting

Defined benefit plans were the main type of retirement plans offered by companies up until the mid 1980s. From the mid 1980s to the present there has been a dramatic increase in the number of defined contribution plans. Companies discovered that defined contribution plans were more beneficial to the company, which resulted in a large number of companies switching the type of retirement plan they offer. This project will explore the reasons why many companies have switched to defined contribution plans and the methods that they have used to switch plans. The switch in plans places more pressure on the employees to invest their money appropriately in order to have sufficient funds for their retirement. This project discusses the pros and cons of these two kinds of retirement plans. This research conducted by M. Muniz.

I. Introduction

The futures of retirement plans are uncertain. Defined benefit plans were the main type of retirement plans being offered by companies until the mid-1980s. From the mid-1980s to the present there has been a major shift from defined benefit plans to defined contribution plans. The main reason for the switch is that defined contribution plans are more profitable for companies. Companies are not investing as much money in defined contribution plans as they do in defined benefit plans. Another possible reason for the change may be a result of the new rule that the Financial Accountings Standards Board (FASB) issued last year. This rule requires modifications to how defined benefit plans are recorded.

Defined benefit and defined contribution plans each have advantages and disadvantages. However, with the increasing number of defined contribution plans it is evident that employers are seeing more advantages to these plans than defined benefit plans. In contrast, the main disadvantage of the increasing number of defined contribution plans is the greater amount of responsibility being placed on employees to invest their money wisely for retirement.

II. Defined Benefit Plans

Defined benefit plans used to be the traditional retirement plan that companies offered their employees. According to the Department of Labor (2007), under a defined benefit plan an employee receives a specified monthly payment when they retire. This payment usually is,

created through a formula that considers salary and service, among other items. A company under this plan accumulates some assets in equity or debt securities (Grant, Grant & Ortega, 2007). The company ensures that it will have sufficient funds to satisfy the pension obligations by investing money into financial assets (Glover, 2007).

Many corporations offer defined benefit plans. Some of these plans have been successful. An example of a successful defined benefit plan is the New York State Teachers retirement plans ("Top Funds by Type," 2007). However, there are many examples of failed defined benefit plans. According to Davolt (2006), in 2002 United Airlines filed for bankruptcy, which meant that they could not afford to pay their retirement obligations to their employees. As a result, the Pension Benefit Guaranty Corporation assumed United Airlines' pension obligations (Davolt, 2006).

II.a. *Advantages and disadvantages of defined benefit plans*

There are many advantages to defined benefit plans that companies find attractive. One of the advantages that will save a company money is the combination of the longevity risk and the insurance principle. The longevity risk is the possibility that one will outlive one's money. Defined benefit plans are only financed to the employee's life expectancy instead of to the extreme outer limit of the individual's life span (Waring & Siegel, 2007). The insurance principle is used when an individual dies prematurely and the excess money is used to pay for those individuals that live beyond the life expectancy (Waring & Siegel, 2007). The company saves money by using this excess money to pay for the retirement of an employee that lives longer than expected.

The lack of excess deferment is another advantage of defined benefit plans compared to defined contribution plans (Waring & Siegel, 2007) Under defined contribution plans if every employee's was fully funded there would be a great deal of excess deferred compensation. This means that if every defined contribution plans were financed to individual's maximum possible age the result would be a great deal of unspent money from those individuals that died before they could spend the money.

Another advantage of defined benefit plans is the forced savings for employees (Waring & Siegel, 2007). Forced savings is a hidden advantage for an employee because the savings rate is unknown to the employee. The employee pays in a certain amount of money, according to this rate, that will result in sufficient amount of money set aside for their retirement (Waring & Siegel, 2007).

Professional management is another advantage of defined benefit plans. Having investments made by professional investment managers improves an employee's chances of having the money well invested than if they did it themselves (Waring & Siegel, 2007). Since defined benefit plans use professional investor managers, they have access to investments that are not available to the average investor. Some examples of these types of investments are real estate, private equity, and hedge funds (Olleman, 2007).

Another advantage for employees is predictable payouts. The amount of money being paid out is the easily calculated, which means that planning for the future is easier for retired individuals (Wilburn, 2007). According to Waring and Siegel (2007), the fees that

employees pay through defined benefit plans are lower than defined contribution fees. The average fees for United States equity funds, which are used mostly in defined benefit plans, are under .5% a year. Retail mutual funds, which are used most in defined contribution plans, are above 1% per year (Waring & Siegel, 2007). Lastly, defined benefit plans can have extra benefits that are advantageous for employees. These extra benefits can include cost of living adjustments, retiree health coverage, and early retirement (Wilburn, 2007).

Defined benefit plans are not without their disadvantages. For employees, one of the flaws with this retirement plan is that it is not portable. The amount of money that was accrued is not able to be taken with the employee if they change jobs before they retire. In addition, employees cannot borrow money against this plan. Also, defined benefit plans cannot be bequeathed to another after that person has died.

Currently, many companies are focusing on the perceived disadvantages of defined benefit plans instead of on the advantages. Companies making the switch view defined benefit plans as dangerous to their balance sheets and income statements. Companies fear that their balance sheets and income statements are being exposed to uncontrollable volatility. In addition, companies view defined benefit plans as being more costly than the alternative, defined contribution plans (Waring & Siegel, 2007). This idea is a result of the financial burden that defined benefit plans place on the company to fund the plan to a set target amount (Glover, 2007).

III. Defined Contribution Plans

Defined contribution plans are very different from defined benefit plans. Under defined contribution plans employers and employees make contributions into the pension accumulation accounts. Employees, typically, are allowed to distribute money into a dozen or so “asset-class” options. Unlike defined benefit plans, where an individual receives payments for their entire lifetime, employees under defined contribution plans usually draw a lump sum amount when they retire from the company (Ambachtsheer, 2007). Some examples of common defined contribution plans are 401(k), 403(b), and 457 plans (Aon Consulting Inc., 2007). A 401(k) plan is a private-sector salary deferment retirement plan that is authorized in the Internal Revenue Code Section 401. 403(b) is a public-sector salary deferment retirement plan that was created in the Internal Revenue Code Section 403. 457 plans are nonprofit-sector salary deferment retirement plan that was created in the Internal Revenue Code Section 457 (History of 401(k) Plans: An Update, 2005).

III.a. *Advantages and disadvantages of defined contribution plans*

The advantages for defined contribution plans are employer and employee specific. Waring and Siegel (2007) believe employers’ value defined contribution plans because these plans have a lower out of pocket deferment than defined benefit plans. This results from the lump sum payment that employees receive as opposed to the deferred payments made by defined benefit plans. In addition, under these plans the amount of money that the employer matches is significantly less then in defined benefit plans (Waring & Siegel, 2007).

Employees' value defined contribution plans because these plans are flexible, which means that employees that cannot afford to put money into their retirement plans are not required to (Waring & Siegel, 2007). Defined contribution plans are valuable to employees because employees are allowed to borrow money against their plans (Waring & Siegel, 2007). Each defined contribution plan varies with its borrowing features.

An example of borrowing against a defined contribution plan is evident in the city of Orlando. An employee for the city of Orlando can borrow a minimum of $1500 (City of Orlando, 2004). The maximum amount of money that can be borrowed is 40% of the accumulated retirement balance or $50,000, which ever is smaller. Repaying the loan is done through payroll deduction. The amount of the deduction cannot be any more than 20% of the employee's salary (City of Orlando, 2004). In addition, a defined contribution plan is portable and the only item that can reduce the value is poor investment returns (Waring & Siegel, 2007). The portability of the plan means that the amount of money accrued in the retirement plan before that individual retires can be taken to another job if that employee leaves their current position.

Defined Contribution plans are not without its flaws or disadvantages. One major flaw is having the employee invest their own money. Many individuals are inconsistent, hesitant, and irrational when it comes to their own financial future (Ambachtsheer, 2007). There are negative effects of employees being hesitant or not having initiative when it comes to investing money on their own in defined contribution plans.

According to Olleman (2007), these effects can be seen in a study done in Nebraska. In Nebraska, state and county employees hired from 1964 to 2003 were using a defined contribution plan. During this same time school employees, state judges, and state patrols were using a separate defined benefit plan. Over the years leading up to 2002, the return from defined benefit plans were 11% while defined contribution plans only returned 6-7%. Deeper investigation uncovered that half of the defined contribution plan members were only using the default stable fund for their money when they should have had funds allocated in more risky but more profitable stock funds in order to increase the amount of the return that the individuals were receiving (Olleman, 2007).

Individuals are at a disadvantage in investing money for their own retirement. Many people do not know the basic "asset-class" building ideas. For example, an individual in their late 20s should invest more in riskier assets, like stocks, than in safer assets, like bonds because a young worker can wait out the ups and downs of the market (US Securities and Exchange Commission, 2007). An individual in their 40s should have more of their retirement invested in safer assets, like bonds, instead of in riskier assets, like stocks. In this case, this individual is closer to retirement and cannot risk their retirement in the ups and downs of the market (US Securities and Exchange Commission, 2007). In addition, most individuals are not able to identify the risks of these assets (Waring, 2007). Due to this lack of understanding today, many employees invest a great deal of their retirement money in the plan sponsor's stock (Ambachtsheer, 2007). This is a huge problem because if the sponsor's stock price drops dramatically the employee will lose a great deal of their retirement money.

As an example, consider an employee in Ford Motor Company that invested all of their retirement in the company's stock. According to the Wall Street Journal (2007), Ford's stock price in 2002 was approximately $16 a share. Five years later, on June 23 2007, Ford's stock price was about $9 a share (Wall Street Journal, 2007). This employee would have lost about half of their retirement money. If this same employee had invested all their money in a mutual fund like Morgan Stanley American Opportunities B the employee would have enjoyed a net increase in wealth. According to the Wall Street Journal (2007), in 2002 Morgan Stanley stock cost approximately $22 per share. Five years later, the stock was about $28 per share, resulting in a five year return of 27%. To obtain an optimal portfolio an employee should have diversified investments. The lack of diversification is not a problem under defined benefit plans. Investment professionals know that the best way to invest money is to have a diversified plan (Ellis, 2007).

Ambachtsheer (2007) states that other flaws workers under defined contribution plans face are information asymmetry and misaligned interest with regard to the financial service industry. Information asymmetry occurs when the employer knows more about the trustee of the pension plan than the employee. Misaligned interests occur if the employer receives a kickback for using that trustee's company. In addition, excessive fees hinder employees' abilities in reaching their pension goals. This tends to drive a wedge between workers and the retirement money that they invest. Also, defined contribution plans leave members to bear the burden of the longevity risk (Ambachtsheer, 2007); although it is possible for employees to hedge this risk by investing in annuities when they retire (Waring & Siegel, 2007). Annuities are stable investments that will provide return without the risk of losing money. Investing in annuities is important because it will give an individual the opportunity in having enough money to live on when they retire.

Another problem that employees face under defined contribution plans is having enough money to live on when they retire. One can see from the Nebraska example that if employees do not invest their money appropriately then their return will be almost half of the return from a defined benefit plan. This reduction in money can make retirement difficult. One may suggest that the solution to this problem would be to save additional money, outside the defined contribution plan, in interest bearing accounts. This is very hard for employees to do because payments in defined contribution plans are made from pre-taxed earnings (Chen & Estes, 2007). Saving money outside the plan requires more pre-taxed dollars.

IV. Financial Accounting Standards (FAS) 158

The Financial Accounting Standards Board (FASB) is making major changes in the first phase of Financial Accounting Standards (FAS) 158. According to FASB (2006), Phase I of the rule requires employers to recognize either overfunded or underfunded status of defined benefit retirement plans as either assets or liabilities. An overfunded plan would be an asset while an underfunded plan would be a liability to a company. In addition, Phase I requires employers to recognize changes in funding status, either over- or under-, in the time period that these changes occurred (FASB, 2006).

One of the main changes in this phase involves a company's disclosure notes. Phase I requires that a company should add any additional information on the effects from delayed recognition of gains or losses on net periodic benefit costs for the next fiscal year in its disclosure notes (FASB 2006). Net periodic benefit costs are the smoothing impacts of gains, losses, changes in prior service benefits, and transition costs (Burrows, 2006). In the industry, smoothing is a company that spreads out its gains or losses over several years instead of recording them all in the period in which they were incurred (Bell, 2006). Phase I of this project, for publicly funded companies, took effect on December 15, 2006. For all other companies FAS 158 takes affect June 15, 2007 (Bell, 2007).

FASB had several goals for creating FAS 158. FASB wanted to make it easier on the users of the financial statements. Before FAS 158, having underfunded or overfunded status appear only in the disclosure notes made it more difficult for the users of the financial statements to assess the company's financial position. In addition, it was difficult to determine the company's ability to satisfy their Projected Benefit Obligation (PBO) (FASB, 2006). In addition, FAS 158 was created to improve financial reporting made by companies. FASB believes that FAS 158 makes information more reliable since it is more complete and timely. Also, it eliminates the alternative measurement date that companies in the past could have used for plan assets or benefit obligations (FASB, 2006).

There were numerous reasons that FASB felt this rule was needed. One major reason was to correct allowances made by previous rules. According to FASB (2006), previous rules allowed employers to recognize a liability that was significantly less than the amount that the retirement plan was underfunded. As a result, companies that did this had an overstated equity section of their balance sheet to make up for the understated liabilities section. Having more equity than liabilities makes the company appear more valuable. Previous rules, also, allowed employers to recognize assets in its financial statements even though the company had an underfunded pension plan. Other rules, also, allowed employers to delay recognition of certain changes in plan assets or obligations that would have affected the cost of providing pension benefits to employees (FASB, 2006).

V. How FAS 158 Works and Objections to the Rule

The major change under FAS 158 deals with the excess of Projected Benefit Obligation (PBO). According to Burrows (2006), PBO is the value of the benefits that have been accrued through the current period. These values are adjusted to reflect any expected future income increases (Burrows, 2006). If a company has a PBO that exceeds the fair market value of the assets than the excess amount will be recognized as a liability. On the other hand if a company has fair market value of its assets exceeding the PBO then that excess needs to be recorded as an asset (Burrows, 2006).

Since FAS 158 is requiring companies to recognize gains or losses in the period that they occurred there are some changes that will affect the smoothing technique. Any gains, losses, changes in prior service benefits, and transition costs should be immediately recorded in the Other Comprehensive Income account (Burrows, 2006). The Other Comprehensive Income account records any amounts that are not part of the processing or

selling of a product. The amounts in this account will be later recognized by the company (Burrows, 2006). As these items, eventually, become a part of Net Periodic Benefits there will be a balancing entry to Accumulated Other Comprehensive Income. Accumulated Other Comprehensive Income is an account of the accumulated year by year adjustments that have been made to Other Comprehensive Income (Burrows, 2006).

There are many companies that are raising objections to FAS 158. Companies are arguing that pension deficits or surpluses should be based upon already built up retirement benefits not on expected income increases. The reason that companies are against using expected income increases is that these increases are not certain and may not happen (Reilly, 2006).

Some companies have problems with how to recognize expected future payments. According to Reilly (2006), these companies argue that obligations that use expected future payments are not liabilities and should not be presented as such on the financial statements. One of the biggest objections from companies is that PBO tends to result in larger obligations than already earned obligations (Reilly, 2006). This means that already earned obligations produce lower deficits or more surpluses for companies. Based on a study by actuarial consultants Milliman, 100 of the country's largest public companies have a combined pension deficit of $90 billion dollars in 2005. If the already earned benefit method is used this deficit shrinks to about a $10 billion deficit (Reilly, 2006).

V.a Effects of FAS 158 on companies

The exact effects that FAS 158 will have on companies are not certain at this point. It is possible that FAS 158 will have a negative impact on a company's stock prices. According to Friedman (2006), FAS 158 may show companies as weaker than they were previous to this rule, which means that fewer individuals will choose to invest in the company. FAS 158 may give bond holders more rights at the expense of equity owners. This situation occurs if the company's debt covenants were based upon stockholders' equity and not on the analysts adjusted shareholders' equity amount (Friedman, 2006).

Another possibility is that FAS 158 will have no impact on company's stock price. Under current models, financial analysts have already taken into consideration accounting reform changes. Since the information was publicly available before FAS 158 professionals could just look in a company's disclosure notes to determine the funding status of the pension plans (Friedman, 2006). Since the company's funding status was public before FAS 158 this means that accounting reform changes are already reflected in the current prices of the company's stock. In addition, companies that are experiencing big pension deficits or other problems already have low stock prices (Friedman, 2006).

FAS 158 may increase a company's stock price. FAS 158 could produce less volatility in operating earnings since analysts and investors are focusing on operating costs without adjusting the financials. Under FAS 87 companies were excessively netting components in operating income. Some examples of these components would be service cost, interest cost, expected return on assets and amortizations (Friedman, 2006). With Phase II of FAS 158 the only components that will remain in operating costs will be service costs and possibly

price or service amortizations. FAS 158 could increase a company's stock prices is by promoting better analysis and decision making from corporate leader and users of financial statements (Friedman, 2006). The improved analysis and decision making will result in more accounting transparency. FAS 158 will allow a company's Chief Financial Officer (CFO) to manage the company's pension plans better. This would occur since a company's financials are better aligned with the underlying economics of the plans (Friedman, 2006).

However, Reilly (2006) believes that FAS 158 could have a negative effect on a company's net worth. Under FAS 158 many companies will now recognize a larger liability. This increased liability recognition could jeopardize the lending agreements that companies have made (Reilly, 2006). The result of losses of lending agreements is that companies could suffer a lower net worth.

VI. Switching from Defined Benefit Plans to Defined Contribution Plans

In 2003, more than 20% of companies that offered defined benefit plans had taken action to freeze these benefits or were actively considering freezing these benefits in order to switch to defined contribution plans (Chen & Estes, 2007). Companies that are considering this switch have some options to make the switch successful. According to Waring and Siegel (2007), a company could choose to soft freeze the current retirement plan. A soft freeze requires that after a certain date a company stops admitting new participants into the current defined benefit plan. It, also, requires that the company cut down the future accruals of current defined benefit plan participants (Waring & Siegel, 2007).

However, a company could take a more radical approach. A company could decide to hard freeze their defined benefit plan participants. A hard freeze occurs when on a certain date a company terminates the existing defined benefit plan and there are no further accruals for the existing plan participants (Waring & Siegel, 2007).

There are various reasons that companies cite to explain the shift from defined benefit plans to defined contribution plans. According to Ellis (2007), by switching to a defined contribution plan a company can avoid the long term liabilities of defined benefit plans. This means that a company's liability to its employees is reduced. In addition, the attendant risk of quarterly Earnings per Share (EPS) disruptions that result from unexpected changes in interest rates help push companies toward defined contribution plans (Ellis, 2007).

The main reason that companies give for switching to defined contribution plans deals with competitiveness. Defined contribution plans can help a company stay competitive domestically and internationally. Chen & Estes (2007) believe that these plans cut a company's costs and put the company more in line with their domestic competitors that do not offer defined benefit plans. In addition, by switching to defined contribution plans these companies are staying competitive with foreign firms. Many foreign companies have their pension plans covered either by the government or by the personal savings of their employees (Chen & Estes, 2007). Companies that are concerned with the volatility of defined benefit plans are choosing to switch to defined contribution plans to eliminate this issue (Geisel, 2007).

There has been a large shift in the types of retirement plans that companies offer from 1980 to the present. According to Glover (2007), the number of employees that were enrolled in defined benefit plans has dropped 27% from 1988 to 2005. During this time, defined contribution plans increased 39%. Single employer plans have, also, seen a decrease. Since 1980, the number of single employer defined benefit plans has dropped 74.4% to 28,769 in 2005 (Glover, 2007). With the decrease in defined benefit plans it is not surprising that defined contribution plans have almost twice the amount of assets as defined benefit plans. Defined contribution plans hold $2.68 trillion of assets while defined benefit plans only hold $1.58 trillion of assets (Glover, 2007).

Public industry is not the only one to see a drop in defined benefit plans. Public industry consists of companies listed on a stock exchange, required to file statements with the Securities Exchange Commission, and closely held, which means that most of the public shares are owned or controlled by a few individuals (Privately Held Company, 2007). Private industry has seen a drop in defined benefit plans, as well. From 1992 to 2005 defined benefit plans have decreased from 25% to 10%. Defined contribution plans have increased from 1992 to 2005 from 44% to 63% (EBRI, 2007).

Recently, there are numerous large companies that are choosing to switch from defined benefit plans to defined contribution plans. General Motors is one of these highly public companies making the switch. According to Larkin (2006), as of January 1, 2007 General Motors is choosing to freeze its current defined benefit plan and is beginning to shift into a define contribution plan. General Motors identified the main reason for this switch was to reduce the financial risks associated with the retirement plans (Larkin, 2006).

VII. Conclusion

It is rather important for employees, in today's market, to take an active role in understanding their retirement plan. Employees under a defined benefit plan need to recognize the large switch that does not appear to be slowing down. Instead of having a secure and predictable retirement future, with a defined benefit plan, many employees are now facing an uncertain retirement. It is up to these employees to properly invest their money in order to have sufficient income for their retirement. Without obtaining proper information or taking an active role in their retirement investing employees, under defined contribution plans, will find it very hard to live during their retirement without getting a part-time job.

The companies conducting the switch from defined benefit plans to defined contribution plans are doing so because defined contribution plans are more advantageous to their company's success. Even though many companies believe that defined contribution plans are more advantageous there are equal advantages and disadvantages to each plan. Although, many companies are make this switch the Financial Accounting Standards Board will continue to monitor and reform pension recognition made by companies.

VIII. References

(2005, February). History of 401(k) Plans: An Update. Retrieved June 20, 2007, from Facts from EBRI Web site: http://www.ebri.org/pdf/publications/facts/0205fact.a.pdf

(2007). Top Funds by Type. *Pensions & Investments*. *35*, 30.

Ambachtsheer, K. (2007). Why We Need A Pension Revolution. *Financial Analysts Journal*. *63*, 21-5.

Aon Consulting Inc, (2006). Aon Consulting's 2006 National Employee Benefits Trends Survey: Key Findings. *Forum*. 1-8.

Bell, A. (2006). FASB Completes Pension Reporting Standard. *Life & Health*. *110*, 7.

Burrows, E. (2006). Dramatic New Accounting Rules for Pension Plans. *Journal of Pension Benefits*. *14*, 5-9.

Chen, H., & Estes, J. (2007). Planning for Defined Pension Plans in Today's Environment. *Financial Service Professionals*. 55-8.

City of Orlando. (2004). *General Employee Defined Contribution Retirement Plan* (1-8). Orlando: Pension Advisory Committee.

Davolt, S. (2006). Stewardesses Reveal 'Naked Truth' About Pensions. *Employee Benefit News*. *26*, 56.

Department of Labor, (2007). Types of Retirement Plans. Retrieved June 23, 2007, from U.S. Department of Labor in the 21st Century Web site: http://www.dol.gov/dol/topic/retirement/typesofplans.htm

EBRI, (2007, June). Retirement Trends in the United States Over the Past Quarter-Century. Retrieved July 5, 2007, from Facts from Employee Benefit Research Institution Web site: http://www.ebri.org/publications/facts/

Eillis, C. (2007). Where Were We?. *Financial Analysts Journal*. *63*, 18-20.

FASB. (2006). Statement of Financial Accounting Standards No. 158, *Employers' Accounting for Defined Benefit Pension and Other Postretirement Plans—An Amendment of FASB Statements Nos. 87, 88, 106 and 132(R)*. Norwalk, CT: Financial Accounting Standards Board

Friedman, E. (2006).Will Accounting Reform Be Bad for Sponsors. *Financial Executive*. *22*, 48-51.

Geisel, J. (2006). Big Names Join Defined Benefit Exodus. *Business Insurance*. *40*, 14.

Glover, H. (2007).Defined-Benefit Plans Still Seen as Viable. *American Banker*. *172*, 1-2.

Grant, T. C, Grant, G. H, & Ortega, W. R (2007). FASB's Quick Fix for Pension Accounting is Only First Step. *Financial Analysts Journal*. *63*, 21-35.

Larkin, J. (2006). GM Annouces Restructuring of its U.S. Salaried Retirement Benefits. *Automotive Industries*. *186*, 41.

Olleman, M. (2007). Defined Contribution Experience in the Public Sector. *Benefits & Compensation Digest*. 20-5.

Privately Held Company. (2007). *Dictionary of Accounting Terms*. Retrieved June 22, 2007, from Answers.com Web site: http://www.answers.com/topic/privately-held-company

Reilly, D. (2006). Moving the Market-Tracking the Numbers/Outside Audit: Accounting Board Sticks to Its Guns. *Wall Street Journal*, p. C3.

US Securities and Exchange Commission, Securities Exchange Commission, (2007). Retrieved July 11, 2007, from Beginners' Guide to Asset Allocation, Diversification, and Rebalancing Web site: http://www.sec.gov/investor/pubs/assetallocation.htm

Wall Street Journal, (2007). Stock Charting. Retrieved June 23, 2007, Wall Street Journal: http://online.wsj.com/quotes/stock_charting.html?symbol=AMCPX&type=usfund

Waring , B. M., & Siegel, L. B. (2007). Don't Kill the Golden Goose! Saving Pension Plans. *Financial Analyst Journal*. *63*, 31-45.

Wilburn, D. A. (2007). A Bad Idea Defined. *NEA Today*. *25*, 48.

SOFONISBA ANGUISSOLA: A TALENT THAT TELLS THE TRUTH ABOUT WOMEN

Sara Getz
Art Department

Due to scientific theories about reproduction, women during the sixteenth century were thought of as inferior to men, unable to create, and unequipped to create new and unique artwork (Jacobs, 1999). During the sixteenth century, Sofonisba Anguissola was one of the first women to raise awareness about the falsity of these ideas. Her artistic talents impress male artists such as Giorgio Vasari and Michelangelo (Jacobs, 1999). Even more significantly, she is credited by Vasari for creating work that is "truly lifelike" (Jacobs, 1999). Sofonisba's talent gives her status among other male artists and gives the women around her hope for gaining an equal status among men.

Since the beginning of art history women have struggled to find equality in the world, including equality in the practice of art making. Due to scientific theories about reproduction, women have been thought to be inferior to men, unable to create, and unequipped to create new and unique artwork (Jacobs, 1999). During the sixteenth century, Sofonisba Anguissola was one of the first women to raise awareness about the falsity of these ideas.

In the first art history class I ever took we discussed the beautiful expressionist paintings of Monet. We analyzed the brilliant ideas behind Picasso's cubism. We gazed in awe at the carvings of Michelangelo and admired his fresco ceiling. These brilliant men crowded our art history textbook, but where were the women artists while all of this history was being made? Among all of these talented men I knew there must be some women to match them. So, why don't we hear more about women artists throughout history and were those who did exist at all significant?

Surprisingly, I found that science was a driving factor in the oppression of women artists during the sixteenth century. Thoughts about the inferiority of women were guided by Aristotle's theory about reproduction. Aristotle says that there are four causes involved with reproduction (Jacobs, 1999). Here are the causes that he brings up: "The efficient cause, or the impetus to the making of that thing, the formal cause, or that which gives form to matter, the material cause, or matter that receives the form; and the telic cause, or goal it endeavors to reach" (Jacobs, 1999, p.31). The efficient, formal and telic causes each deal with the role of the male. During reproduction, these include the creation of sperm and the passing of sperm from the male to the female. The man is therefore thought of as a giver and creator of life.

The material cause, on the other hand, is the one and only cause attributed to the female (Jacobs, 1999). This cause puts emphasis on the passiveness of the woman. She is the one who is given the sperm by the male and passively accepts it into her body (Jacobs, 1999). Her only role is to give it protection during the creation of a child. Basically, Aristotle is saying that the man is in control of the reproductive process and, unlike women, can create and give life to new beings.

Unfortunately, Artistotle's theories about reproduction are applied to other aspects of life as well. For example, at this time women are very restricted with what they are allowed to do in the world of art. For the most part, art is thought of as a man's profession as it involves the act of creating, something Aristotle says is not a woman's activity; in fact women do not even posses the ability to create (Jacobs, 1999).

However, during the sixteenth century, when women did insist on picking up the paintbrush, there were additional restrictions that they were expected to follow. For example, painting the figure was absolutely prohibited. It was certainly not proper for a lady to use nude models, which was a necessity in order to learn how to paint the figure (Chicago & Lucie-Smith, 1999). Therefore, it was also rare to have a woman painting portraits because this was also a form of figure painting (Chicago & Lucie-Smith, 1999).

The most forbidden painting genre among women was the highest-ranking form of figure painting, called history painting (Chicago & Lucie-Smith, 1999). These paintings tended to be highly complex images, loaded with detail of narrative and full of religious content (Chicago & Lucie-Smith, 1999). Lower-ranking forms of art such as landscape painting and still-life painting were the most accepted genres among women artists (Chicago & Lucie-Smith, 1999).

However, Sofonisba Anguissola, knowing the set conventions for women artists and aware of Aristotle's theories, overcomes and challenges these obstacles through her subject matter and her level of skill. One painting from the 1550's, called *Pieta*, shows that Sofonisba is quite able to be creative (Fig. 1) (Perlingieri, 1992). She copies *Pieta* from a painting that her teacher has already made with the same title (Perlingieri, 1992). Then, a man named Lionello Venturi comes across these two paintings at their original exhibition in Milan (Perlingieri, 1992). He says:

> In a small room of the same gallery, with a similar background of gloomy cliffs, the central group of Bernardino Campi's *Pieta* was copied by Sofonisba Anguissola. The luminous close-up imagery and *sfumato* [a smoky shading and coloring from light to dark] of the countryside, on the right, is thus enhanced (Fig. 1). Her teacher's work is scrupulously and truthfully copied. But Sofonisba shows more signs of Parmigianesque influence than Bernardino, in the pointed hands of the Virgin, in the golden color, in the delicacy of the pointed face and tiny features (Perlingieri, 1992, p.47).

Venturi describes Sofonisba's work as having a higher quality and better usage of techniques such as sfumato and detailing than even her own teacher who was, of course, a male. Sofonisba's *Pieta* (Fig. 1) is an exact copy of her teacher's painting of the same title;

however, she drastically alters it to make it even more successful than the original (Perlingieri, 1992).

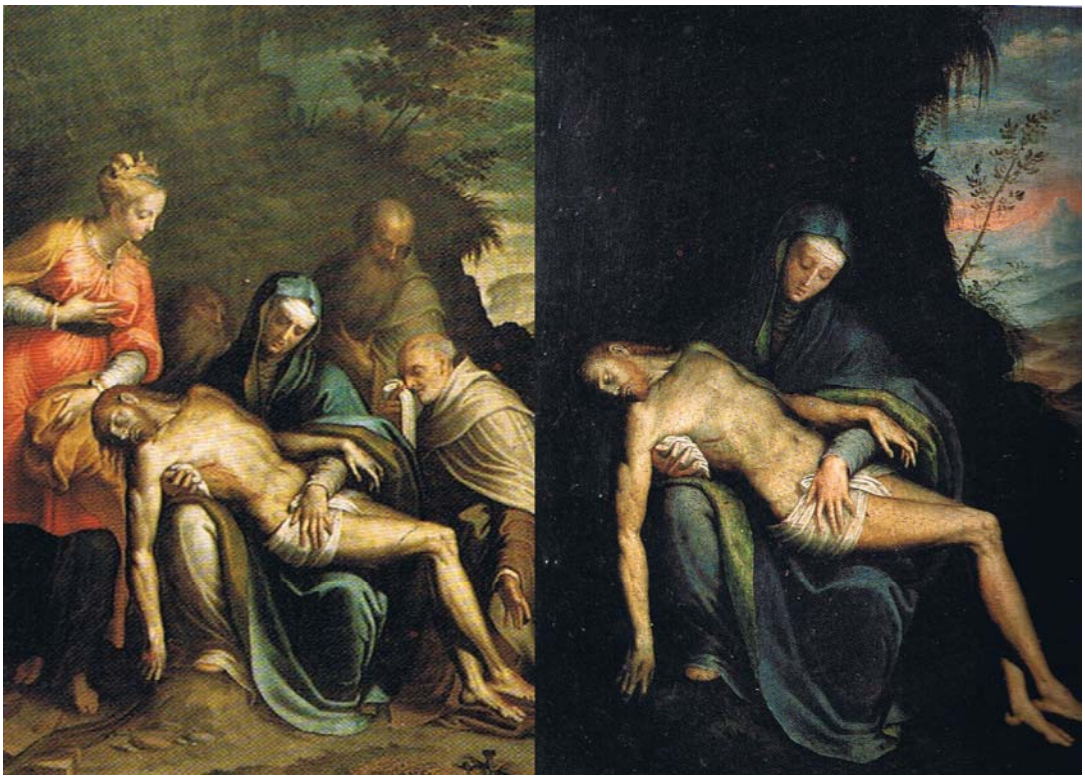


Fig. 1: *Pieta*, by Bernardino Campi, 1550's on left and *Pieta*, by Sofonisba Anguissola, 1550's on right.

Sofonisba's talents are even able to impress the famous sculptor and painter, Michelangelo. Receiving recognition and guidance from this master alone speaks volumes about Sofonisba's abilities. Furthermore, after seeing Sofonisba's drawing of a girl smiling, Michelangelo requests that she should also draw a boy weeping (Jacob, 1999). Although impressed by her first drawing, Michelangelo believes that this would be a much more difficult task (Jacobs, 1999).

Sofonisba then accepts this challenge and creates a drawing called *Asdrubale Bitten by a Crayfish*, in 1557 (Fig. 2) (Jacobs, 1999). Author Fredrika Jacobs describes the level of difficulty involved with creating such a drawing:

> Anguissola's drawing of her tearful brother and smiling sister may also be seen as a successful resolution of the problem of differentiating responses that, while generally expressive of opposite emotions, are strikingly siilar in their physiognomic manifestations (Jacobs, 1999, p.57).

Fig. 2: Sofonisba Anguissola, *Asdrubale Bitten by a Crayfish*, 1557.

He explains that the facial differences between laughing and weeping can be very hard to distinguish for an artist (Jacobs, 1999). The slight differences that appear on the face are hard to differentiate; therefore, it is a huge accomplishment to create each of these facial expressions.

Upon seeing Sofonisba's paintings, another artist, Giorgio Vasari, in addition to Michelangelo, was also impressed by her talent. After reviewing Sofonisba's work, he says that "If women know so well how to make living men, what marvel is it that those who wish are also so well able to make them in painting" (Jacobs, 1999, p.38)? He is actually crediting Sofonisba as being able to give life to men as well as her paintings. Her painting, *The Chess Game*, from 1555 is an example of her ability to give life to painting (Fig. 3) (Perlingieri, 1992).

Perlingieri, the author of *Sofonisba Anguissola: The First Great Woman Artist of the Renaissance*, describes this painting as "a complex and ambitious project" (Perlingieri, 1992, p.82). Sofonisba paints three of her siblings playing chess along with a woman that appears to be

Fig. 3: Sofonisba Anguissola, *The Chess Game*, 1555.

the maid. She creates an outdoor setting that was most likely constructed from her imagination using some elements from the city and landscape around her (Perlingieri, 1992). Similar to *Pieta*, here she uses sfumato in order to create a soft blue landscape in the background that nicely contrasts with the decorative and elaborate clothing of the foreground figures (Perlingieri, 1992). Vasari, again, complements Sofonisba's work, explaining that "It was done with such diligence and quickness that they all seemed truly alive and only lacking in speech" (Perlingieri, 1992). He is referring to the figures in this painting as being lifelike, which is, again, a huge compliment for any woman artist to receive.

Furthermore, the interaction between the figures makes the game of chess seem real. The youngest sibling appears to be laughing at the sibling on the right as she raises her hand in defeat. Meanwhile, the oldest sibling looks out at the viewer as if asking for praise and congratulations as she holds her sisters black queen, indicating that she has won the game of chess (Gerrard, 1994). Perhaps these gestures and expressions are what cause Vasari to credit Sofonisba with the ability to create life.

The Chess game's main subject is also important to Sofonisba's message for this painting. Chess had undergone many rule changes as it was introduced and popularized in Italy in 1510 (Gerrard, 1994). The new game of chess, much like the kind played today,

gave the pawns new abilities. For example, in order to make the game move along faster, the bishop pawn gained the ability to move an infinite number of spaces diagonally, when previously it could have only moved one diagonal space at a time (Gerrard, 1994). Every other pawn also had its own new set of rules. The most significant rule change for Sofonisba's painting, however, belongs to the Queen. Author Mary D. Gerrard (1994) states that "The new status and power of the queen-now greater than that of the king himself--was evidently the most noteworthy result of the rules change…" As the oldest sister grasps her sibling's black queen in her left hand, she signifies more than the fact that she has just won the game; the queen, as a symbol, implies that the status of women should at least be equal to the status of men in society (Gerrard, 1994). As the status of the chess queen has risen, so should the status rise of all women in society.

Sofonisba did not create life-like paintings by accident. She knew that what she was doing was significant. In her *Self-Portrait*, from 1554, her pose and what she holds in her hand are important signifiers for her message (Fig. 4) (Perlingieri, 1992). She sits in a three quarter pose and gazes out at the viewer in a challenging stare (Perlingieri, 1992). She dresses in a simple, yet elegant gown adorned with lace at her collar and sleeve openings. Her dress is simple, but shows that her status in society is not of lower-class.

Fig. 4: Sofonisba Anguissola, *Self-Portrait*, 1554.

The book that Sofonisba holds in her left hand is a sign of her intelligence and success (Chicago and Lucie-Smith, 1999). On it is written, "Sofonisba Anguissola virgo se ipsam fecit 1554 (Sofonisba Anguissola, the unmarried maiden painted this herself, 1554)" (Perlingieri, 1992, p.78). Here she makes a statement about having never been married. Therefore, she is also saying that she has never belonged to a man and never given birth to a child (Perlingieri, 1992). She detaches herself from the theories of Aristotle and in a way rejects them while also taking on some manlike qualities. She also writes in a very matter-of-a-fact tone which puts emphasis on the fact that and unmarried, childless woman painted the portrait.

Women during this time had trouble gaining the full respect from men that they may have deserved. In order to gain this respect, women who wanted to paint could either choose to paint with some restrictions, or paint on their own accord but gain no respect. Artist and author Judy Chicago and Edward Lucie-Smith (1999) describe the situation as follows:

> Essentially a female artist could either elect to be a gentlewoman who painted, with the restrictions which this implied in terms of both daily conduct and longterm social mobility, or she could resign herself to being thought of as a not quite repectable outsider. A gentlewoman artist had access to a particular sort of patronage – she could become a favorite with the female members of a court or aristocratic circle. But this limited the scope of her art. This seems to have been the choice made by the gifted mannerist portraitist Sofonisba Anguissola (Chicago & Lucie-Smith, 1999, p.115).

The book that Sofonisba holds shows that she has joined this circle of intellect and sacrifices many freedoms in order for her work to be respected and well-known (Chicago & Lucie-Smith, 1999).

Sofonisba was one of the first women to begin the fight for equal rights among men and women. She fought for the right to be taken seriously as an artist in the same way that men were taken seriously as artists. Sofonisba turned her knowledge and determination into master works of art. Credited by Vasari as have "truly life-like" images, and praised by Venturi for improving upon her own teacher's work, Sofonisba has accomplished more than art; she has helped rise the status of women to a position of equal importance as men (Jacob, 1999).

References

Chicago, J. & Lucie-Smith, E. (1999). *Women and Art: Contested Territory*. New York: The Ivy Press Limited.

Gerrard, M. D. (1994). *Here's Looking at Me: Sofonisba Anguissola and the Problem of the Woman Artist*. Renaissance Quarterly, 47(3), 556-622. Retrieved Spring, 2008, from WilsonWeb database.

Jacobs, F. H. (1999). (Pro)creativity. *In Defining the Renaissance Virtuosa: Women Artists and the Language of Art History and Criticism* (pp. 27-63). Cambridge: Cambridge University Press. (Original work published 1997).

Perlingieri, I. S. (1992). *Sofonisba Anguissola: The First Great Woman Artist of the Renaissance*. New York: Rizzoli.

Photographs

Anguissola, Sofonisba. “Self-Portrait”. Oil on Canvas. (1554). Kunsthistorisches Museum, Vienna. Web Gallery of Art. Ed. Emil Kren and Daniel Marx. 1996. 28 Mar. 2008 <http://www.wga.hu>.

Anguissola, Sofonisba. “Asdrubale Bitten by a Crawfish”. Black chalk and charcoal on brown paper. 1554. Pinacoteca Nazionale, Siena. Web Gallery of Art. Ed. Emil Kren and Daniel Marx. 1996. 28 Mar. 2008 <http://www.wga.hu>.

Anguissola, Sofonisba. “Pieta”. Oil on Canvas. (1550's). Pinacoteca di Brera, Milan. Sofonisba Anguissola: The First Great Woman Artist of the Renaissance. By Ilya Sandra Perlingieri. New York: Rizzoli, 1992. Plate 23. 50.

Campi, Bernardino. “Pieta”. (1550's). Pinacoteca di Brera, Milan. Sofonisba Anguissola: The First Great Woman Artist of the Renaissance. By Ilya Sandra Perlingieri. New York: Rizzoli, 1992. Plate 23. 50.

Sofonisba Anguissola: “Lucia, Minerva and Europa Anguissola Playing Chess” (1555) National Museum, Poznan; Photo credit: Erich Lessing/Art Resource, NY. Art Resource. [26 March 2008], http://www.artres.com/.

Adoption: A Literary Exploration

Claire A. Kavanaugh
Counseling and Psychological Services Department

Adoption is a legal process in which a child is raised by someone other than his or her biological parents and can be accomplished with varying degrees of openness. Adopted children often experience a period of grief at some point in their lives related to their adoption and their loss of their biological origins. Bibliotherapy is one method of helping children cope with adoption related grief and though most of the adoption books in print are authored by adoptive parents and fail to address the issue of the child's birth parents in any detail. In response I authored a children's picture book intended to supplement the available literature—recreational, educational and therapeutic—written from the perspective of a birthmother that focuses on the love involved in adoption. The purpose of the book is to function as a bibliotherapy tool and aid in the healing process of adopted children as well as facilitate a better understanding of the often stigmatized practice of adoption.

I. Openness in Adoption

There are several types of adoption that differ in the way the legalities are handled, either through private attorneys or through adoption agencies, and in the amount of contact allowed between the birth and adoptive families. In closed adoptions there is no contact between the child and his adoptive family and his birthparents; all access that an adopted child may have to his biological origins is cut off until he is an adult and able to search for his birth family himself. Open adoptions, on the other hand, allow for open contact between the two families. This contact can include letters, pictures and visits. In between closed and open adoptions are semi-open and semi-closed adoptions that have varying degrees of openness and contact.

The most important benefit of open adoptions is the access the adopted child has to his entire history. Children of open adoption have the opportunity to learn about the history of their biological family, their own medical history, ethnicity, genealogy and even whose eyes they have or from whom they inherited their talents—all luxuries children in closed adoption do not have. Adopted children run the risk of not being able to develop a complete sense of self and that risk is even higher in adoptions where the child is not able to explore freely his biological origins.

No matter what the circumstances or openness of an adoption, many adopted children struggle with related grief and loss. While open adoptions can aid in minimizing these feelings by providing immediate answers that only birth parents can provide, the grief

process is one that most adopted children will most likely have to work through at some point in their lives.

II. The Grief of Adopted Children

In our western society we operate under the assumption that children do not grieve. We believe them to be resilient and able to bounce back from any loss or injury, be it emotional or physical. Often we conclude that if we simply do not mention loss or, at the very most, give it only a passing acknowledgment and stay strong, children will do the same and just move on; in essence, if we as adults skirt the issue of grief and loss with children then they will remain protected from sorrow and pain. We want so badly to protect the fleeting innocence of children that we have convinced ourselves that the best course of action is to brush aside their pain with phrases like "you're ok," "don't feel bad," and "don't worry," and not address their grief directly, or at all, because, after all, they are only children and, therefore, cannot possibly comprehend loss to the extent that an adult is able to. The latter comment may have a kernel of truth, in that children rarely have to deal with the logistical aftermath of a loss, such as death, but that does not mean, however, that they do not grieve that loss as deeply or personally as their adult counterparts. When we shelter children from grief we are really protecting ourselves from addressing our incapability to relate to and help a child through a loss (Fiorini and Mullen, 2006).

The fact of the matter is children do grieve. Though their understanding and concept of loss matures with time (a nine-year-old will recognize that death is permanent while a three-year-old may ask when the deceased relative is coming back) they are capable of internalizing and recognizing a loss no matter what their age. What eludes adults and leads us to believe that children do not grieve is the way in which children grieve. They will often act out, or alternate between being sad or angry and being outwardly "fine", or they may escape to a world of fantasy (Fiorini and Mullen, 2006). Children are not adults and as such we do not expect them to approach the world in the same way we do, so why should we expect them to approach loss as we do or grieve as we do?

Adopted children face a unique type of loss. Their loss occurred very early in life; in many cases right after they were born. This pre-existing loss was one they could not deal with or even comprehend as a newborn, but this does not mean that it is one they will never feel the ramifications of. Adoption is a process, not simply a single, isolated event and will be a large part of a child's story and sense of self; while it does not define a person, it is, nevertheless, a part of who they are (Silber, 1990).

Loss can be defined as: "the permanent or temporary removal of an important object, person, or event, or failure to achieve a coveted goal" (Van Gulden and Bartels-Rabb, 1993, p. 20). Under this definition the loss of an adopted child is the loss of a person, or persons; his birth family. It is also a loss of origins and of a complete self identity (Joshua and DiMenna, 2000). The lack of shared physical traits with family members could create feelings of detachment as physical similarities are an easily recognizable, outward sign of belonging (Van Gulden and Bartels-Rabb, 1993). This feeling of isolation is further compounded by societal conceptualizations that define procreation and parenthood as indistinguishable entities, thereby facilitating the notion of "real" parents. Adopted children face the dilemma of who their "real" parents are, those who gave birth to them or those who raised them. Their grief may also stem largely from their conflicting feelings as they struggle to integrate the two sides of themselves: their adopted family, which to them is

their "real" family, and their biological family, which society has deemed their "real" family (Leon, 2002). Family is where you go to belong and feel like an intrinsic part of something. Adoptees, however, lose that biological connection and so may feel as if they do not truly belong anywhere.

As a child grows and matures so does his concept of adoption, which may trigger feelings of grief. Adoption loss is certainly a loss that can change and reemerge over time. As children grow they will understand information regarding their adoption in new and different ways. Things that didn't make sense to them as a child may make sense as an adolescent and they may need to grieve and reprocess their feelings in order to reconcile their new understanding. This cycle can easily take a lifetime and may never be fully resolved.

Preschool aged children often grieve by rejecting the fact that they were born to someone other than their adoptive parents and mourn the loss of not only their biological family but also the loss of a genetic tie to their adoptive parents (Silber, 1990). Adoption related grief typically becomes most noticeable in school aged children at which point they are cognitively able to better understand and internalize adoption and apply its implications to their own lives. They may become angry with their adoptive parents for not being their "real" parents or they may feel as if they "should have been" born to the parents they live with like many of their friends were (Silber, 1990; Van Gulden and Bartels-Rabb, 1993, p. 34). At this age children may also begin to feel guilty about being adopted; they may feel there was something wrong with them and that is why they were unwanted by their birthparents and will view the relinquishment as abandonment and rejection (Joshua and DiMenna, 2000). In a study conducted by Triseliotis (1973, as cited in Watkins and Fisher, 1993) the main thing interviewed adult adoptees said they would have liked to have been told was that they were not rejected but wanted and loved by their birthparents.

As children begin to make sense of their own adoption story they may experience many secondary losses including:

Security — if their first parents gave them up why wouldn't their second (adoptive) parents who they are not even biologically related to, in most cases, do the same? Also, what if their biological family came to take them away from the only family they've known?

Control — someone else chose who their family would be and what their life would be like.

Sense of normalcy — if few, or any, of their other friends are adopted, thereby making them "different."

Innocence — they are realizing that life isn't always easy and that sometimes bad or sad things happen—especially if their adoption was in response to drugs, alcohol, abuse, abandonment, economic hardships, etc.

Perspective — they may view their adoptive parents differently. They are no longer their only set of parents. They may even idealize their birthparents and make up stories about their past if not enough information is available to them.

> Identity — they may feel like they don't know who they are or where they belong, especially if they are from another country or culture. They may feel disconnected from themselves because they don't know anything (or little) about their ethnic, medical and genetic background.

They may also grieve the fact that their questions may never be answered. They will have to come to terms with the fact that they may never have the opportunity to know "why" or have the chance to speak with a birthparent or someone else who can answer at least some of their questions. This unknowing may leave them with a feeling of emptiness. Those of us who are fortunate enough to have been raised by our biological parents undervalue the easy access we have to information regarding our past. We even take for granted the fact that a person's past doesn't begin with their birth but instead encompasses events that happened long before they were born, events that involved our family members and ancestors whose actions laid the pathway to our present. For adopted children their adoptive parents' past is not their past and their own story often only involves the events of their lifetime. Curiosity about one's origins is natural but, unfortunately, it is this natural inquisitiveness that adoptive children usually feel guilty about having.

The adoption process involves extremely complex emotions on all sides of the triad (the adoptive parents, birth parents and adopted child) and therefore warrants the same level of openness that is necessary in other relationships (Silber, 1990). They may not want to approach their adoptive parents (the only link to information they have) with questions regarding their birth family as they are worried they may hurt their parents' feelings. By asking about their adoption story and genealogy they fear that it may suggest to their adoptive family that they do not love them or view them as their true family. They may even swallow their grief or questions in an attempt to fit in with their adoptive family due to fear of further rejection as sometimes adoptees feel as if they owe their adoptive families for accepting them when their biological families did not (Eldridge, 1999). It is, therefore, up to the adoptive parents to be receptive to their children's grief and curiosity and encourage conversation to address issues and concerns related to their adoption that may arise. Maria Trozzi outlines the three major functions that caregivers of grieving children have:

1. To foster honest and open relationships with children
2. To provide a safe and secure space in which children can mourn
3. To be role models of healthy mourning

(Trozzi, 1999, p. 11)

Children grieving the losses associated with adoption or questioning their self identity and origins need their feelings to be recognized and validated; they need to be assured that what they are going through is normal. Due to possible fear and guilt of the child adoptive parents often need to initiate conversation or at least let children know that it's ok to talk about their adoption. It is important for children to know that they can approach their adoptive parents with their questions and that they will share in their grief and search for information. Adoptive parents themselves are not immune to feelings of grief and loss of their own related to adoption. They feel the loss of being able to have biological children, loss at having to "share" their children with another set of parents, loss at having to prove to others that they are ready for and would be good at parenting. By sharing in this grief and mourning together parents and children can begin to cope together and will only

strengthen the bond between them. After all, "grief shared is grief diminished" (Trozzi, 1999, p.11).

III. Bibliotherapy as a Coping Tool

Experiences that elicit grief and loss are ones that will cause a child to question how the world works; they will have encountered new experiences and emotions that will ultimately alter their perception (Manifold, 2007). Children generally lack the verbal skills to express complex emotion and thus may rely on their imagination as a safe haven in which they can work through their experiences. Emotionally painful experiences can be brought to mind by a sensory stimulus and by understanding the triggering event can learn about the emotional response. Picture books can be the perfect means for initiating a deeper emotional understanding in children by providing the exact tools that enable the imaginative mind to interpret painful experiences—rhythm, metaphor, image and simple narrative—and thus tap into a primal way of dealing with emotional stress (Manifold, 2007). Egan created a framework for the stages of imagination that a person will go through as they grow (Manifold, 2007). These stages are somatic, mythic, romantic, philosophic and ironic thought with somatic being the earliest stage and ironic thought the final stage. Elementary aged children usually relate best to books that utilize the characteristic elements of the somatic and mythic stages: rhythm, metaphor and narrative. Even in older children, adolescents or even adults, when a deeply stressful situation arises the instinct is to regress to the mythic and somatic stages, regardless of current imaginative developmental stage, to begin making sense of their experience (Manifold, 2007). Due to their ability to guide us to personal understanding no matter what our age, picture books can be an important bibliotherapy tool used by both parents and professionals in initiating conversation with children who have experienced a loss.

Bibliotherapy is broadly defined as the use of books to address emotional or behavioral issues (Barancik, online). It is meant to create a therapeutic interaction between the audience and the literature (Grier, 2006). When using picture books as bibliotherapy, in which the child is often read to, the child not only interacts with the book itself (both text and illustrations) but also with the reader. It is this added dimension that makes bibliotherapy an ideal way for a parent to open up communication with their child on topics such as their adoption and birth family. The advantage of bibliotherapy is that it allows children to step back and take a third-party, objective view of their own personal situation and test out their different emotions and reactions in a safe, risk-free environment (Grier, 2006). A good bibliotherapy book will gently lead the reader to a better, deeply personal understanding of not only their experience or situation but also of themselves. The revelation may be as simple as realizing that, like the character in the book, the reader also isn't a very good listener all the time or as complex as understanding that being adopted means that they were chosen by their adoptive parents and will always be loved by their birth parents (Pehrsson, 2006). Books can provide a voice for children who are unable to express what they think and feel and even promote an enhanced self awareness and worth, especially valuable to children dealing with adoption losses. By reading about how others dealt with similar situations or by being introduced to another perspective children can identify and develop their own healthier, more informed method of coping (Pehrsson, 2006). Perhaps most importantly is that by using bibliotherapy to approach topics related to their child's adoption, adoptive parents are not only validating their child's feelings,

reassuring them that it is ok to talk about their adoption and aiding in their grieving process by sharing in it with them, but they are also strengthening the parent-child bond that will further their emotional growth and healing.

Simply reading an appropriate bibliotherapy book is not enough to elicit a therapeutic effect in a child. Rather it is the resulting interaction, including questioning, exploration, and conversation, that is incited by the book that will allow the child to apply what they have been read to themselves and their own situations. If the child is not invited to relate the story to themselves or make sense of their own situation by questioning different aspects of the book it will have no therapeutic value.

Bibliotherapy can also be used in more than just a one-on-one setting. Classrooms are ideal venues for introducing larger groups of children to different points of view. Some grief issues in adopted children arise from being treated as different by their classmates who do not understand adoption. By using adoption related bibliotherapy to inspire discussions in class, students can become educated about adoption thus leading to a greater general understanding and acceptance among an adopted child's peers (Manifold, 2007).

Not all picture books are appropriate as bibliotherapy. In order for a picture book to be a useful therapeutic device it must have not only meaningful text and engaging illustrations but a successful marriage of the two (Manifold, 2007). The illustrations must highlight and reinforce the important messages of the text and the sophistication of both the illustrations and text must match precisely. For example, if the vocabulary used in the text is appropriate for elementary aged children the illustrations should not be geared toward preschoolers. The child audience will use the illustrations as an interpretive guide for the story he is being read. Also, the author must approach the central question of the book with dignity and avoid presenting the resolution as the only possibility to encourage imaginative interpretation by the child (Manifold, 2007). Finally, selection of an appropriate bibliotherapy book must be done with the child (e.g. his age, maturity and education level, interests, and specific problems) in mind. If the child is unable to relate to or is simply uninterested in the book he will be unable to internalize its meaning and it will fall short of its therapeutic, healing goal.

It is with the coping process of children and the goals of bibliotherapy in mind that I conceived and designed my picture book so that it would be a useful bibliotherapy tool for children dealing with adoption related grief and loss. The story is meant to explain one possible adoption scenario and explain why birth parents might relinquish their child. Unique among most adoption books available it focuses primarily on the birth parents and the dual love of the child's two families. Adoption, while emotionally difficult at times, is a beautiful practice that can lovingly create families and should be celebrated as such.

IV. My Piece

This is the story
Of how you came to be
Part of not one
But two families

There are few little boys
As lucky as you
Instead of one family

Why, you, sir, have two

It all started way back
One mid winter day
When I did find out
You were on the way

I called up my love
To tell him the news
So unexpected
“Well, what should we do?”

We took stock of our lives
Took a good look around
With school, bills, and youth
Nothing promising found

No yard you could run in
No room for a crib
Wallets with cobwebs
For bottles, toys, bibs

We thought and we thought
And came up with a plan
Adoptions the ticket!
For our little man!

We’ll find him a family
With all that we lack
A castle to live in
Fun toys by the sack!

This big decision
Was no easy one
We wanted forever
To be with our son

But we knew we could not
Give you the whole world
For you deserve life,
Before you unfurled

We read through the stories
Of ladies and lads
Who all had one wish:
To be mothers and dads

We looked at the pictures
Read you every word
To see which of the couples
You liked or preferred
We read you of dogs,
And apple picking
And knew they were it
By your joyful kicking

After nine fleeting months
And long labor and strife
You came to try out
This thing they call ‘life’

The three of us played
You, your birthdad and I
For two wonderful days
‘Till the time for “good bye”

An excited young couple,
Your parents-to-be,
Had flown cross the country
Their son for to see

When your parents arrived,
Looking elated and smart,
Into their arms
We handed our heart

Through pictures and letters
We watch you grow
The smiles and adventures
Your happiness shows

Every so often
A child comes along
So unique and so special

Too little love would be wrong

The love of one family
Just will not do
For one so amazing
Double the number by two

Though so far away
We love you so much
There's not a day that goes by
That our hearts you don't touch

So remember, my love
We whisper each night
To reach you by moonbeams
"We love you, good night"

IV.a. *My piece as bibliotherapy*

I believe this book to be a useful bibliotherapy tool because it will facilitate and encourage conversation about and a deeper exploration of a child's adoption history. The piece itself is lively yet heartfelt, simple yet intricate and focused on adoption but anchored on the universal topic of love. It explains the reasons a birthparent may relinquish their child and reassures the reader that it was not out of malice or disappointment or due to any fault of the child's; instead it was a decision made out of love and wanting nothing but the best for that child, which the birthparents could not provide. The central theme is not sacrifice and the tone is not maudlin or depressed. A children's book needs to speak to children on their level. It needs to touch them through both text and illustrations. If they cannot relate to the book or if it simply does not hold their interest it fails to convey its message. The strength of my work is that it gently conveys complex emotions in a positive, interesting way that a child can easily relate to. It will, however, be up to the counselor or parent to engage the child in the story by perhaps asking questions along the way that promote a more complex understanding of the meaning of the text or by going back through the story once it has been read through completely in order for the book to reach its full therapeutic potential.

V. Acknowledgements

I would like to thank Dr. Jody Fiorini for her encouragement and support in preparing this piece and Dr. Maureen Curtin for her vast contributions to my project as a whole.

VI. *References*

Barancik, S. (2006). Best children's books. Find, read or write. *Bibliotherapy: Good book happier kid.* Retrieved January 9, 2008, from http://www.best-childrens-books.com/index.html.

Eldridge, S. (1999). *Twenty things adopted kids wish their adoptive parents knew.* New York, NY: Dell Publishing.

Fiorini, J. J. and Mullen, J. A. (2006). *Counseling Children and Adolescents through Grief and Loss.* Champaign, IL: Research Press.

Grier, L. O. (2006) Bibliotherapy: Coping with loss. *Louisiana Libraries, 69* , 12-20.

Joshua, J. M. and DiMenna, D. (2000). *Read two books and let's talk next week: Using bibliotherapy in clinical practice.* New York, NY: John Wiley and Sons, Inc.

Leon, I. G. (2002). Adoption losses: Naturally occurring or socially constructed? *Child Development, 73*, 652-663.

Manifold, M. C. (2007). The healing picture book: An aesthetic of sorrow. *Teacher Librarian,34*, 20-26.

Pehrsson, D. E. (2006). Benefits of utilizing bibliotherapy within play therapy. *Play Therapy 1.2,* 10-14.

Silber, K. (1990). *Children of open adoption and their families.* San Antonio, TX: Corona Publishing Co.

Trozzi, M. (1999). *Talking with children about loss.* New York, NY: Berkley Publishing Group.

Van Gulden, H., and Bartels-Rabb, L. M. (1993). *Real parents, real children: Parenting the adopted child.* New York, NY: Crossroad Publishing Co.

Watkins, M. and Fisher, S. (1993). *Talking with Young Children About Adoption.* London: Yale University Press.

ALLEGORIES OF VAMPIRE CINEMA

Jeremy Magnan
Department of English & Creative Writing

Allegories of Vampire Cinema is a theoretical film essay involving the issue of spectator relations to vampire films before, during, and after viewings. The piece closely examines which character the spectators are truly meant to connect with. This is an interesting and important issue to raise as it offers a new analysis that had not previously been explored, aligning the spectators not with the protagonists of these stories, but with the vampire itself. In my research, I gathered dozens of books, magazine articles, and journal entries to delve deeply into the horror genre and vampire subgenre. I also screened over three dozen vampire films, though only a handful are cited directly. The essay was pieced together from the beginning of January through March when, upon completion, I presented my findings at the 2008 PCA/ACA National Conference in San Francisco. Implications that are brought to light upon the revelation that the spectator is being aligned with vampires include the notion that the vampire film may not be an isolated case. With further study, theories and analyses may bring about spectator relations and alignments with not only a myriad of other antagonistic horror icons, but antagonists throughout the entire scope of film.

Many authors have sought to lend insight into the metaphorical relationship between the vampire, their victims, and even their spectators. On the spectators of horror films in general, Joseph Biggs and Dennis Petrie offer that "...one goes to the horror film in order to have a nightmare... a dream whose undercurrent of anxiety both presents and masks the desire to fulfill and be punished for certain conventionally unacceptable impulses (Biggs & Petrie, 2008, p. 484)." It is their position that the spectators of horror view these films due to a subconscious desire to see their "unacceptable impulses" played out by the monster (in our discussion, vampires) and to be punished for the surrogate actions that the monster plays out in our stead. In regards to the vampire, Jorg Waltje sees our clear alignment with the vampire as soon as we sit down in the theater. He explains:

> "The vampire only comes out in the dark and spends the rest of the time in his coffin. The spectators voluntarily sit in a coffin (the darkened cinema), watching a screen on which not only light but also (within and between every frame) darkness is projected (Waltje, 2000, p. 29)."

While I agree that this is a startlingly clear example of our relationship to the vampire, this vampire-spectator relationship can be further clarified through a common iconographical object in most of these films in a way that has not as yet been established.

Lacan's famous mirror stage is one of his pillars of seeking out the moment when the identity of a child in relation to itself begins to develop. "The child... can already recognize as such his own image in a mirror. This recognition is indicated in the illuminative mimicry of the *Aha- Erlebnis*... This event can take place... from the age of six months... up to the age of eighteen months (Lacan, 2004, p. 441-442)." *Aha,* you may say, *but the vampire casts no reflection, does it not?* Stoker himself, Dracula's keeper, has been the catalyst for your exclamation: "This time there could be no error, for the man was close to me, and I could see him over my shoulder. But there was no reflection of him in the mirror (Stoker, 2003, p. 30-31)!" So what would Dracula's answer to Lacan's mirror stage be in fact? Fiona Peters states:

> "Vampires have no need for an unconscious- nor can they be seen in mirrors because they do not need to rely on the process of identifications that Lacan describes; in other words they have not become formed as human subjects, and in the case of those who become vampires after being human... they have evaded the symbolic order... (Peters, 2006, p. 180)"

In Peters' argument, humans who become vampires have separated and transcended themselves from the symbolism that is the vampire to become one of them. Interesting... My question for Peters would be *What if someone was a vampire and didn't know it? Must they still graduate from the fully-fledged human's mirror stage?* I believe they do. But who ever heard of someone not knowing that they are in fact a vampire? Perhaps my line of questions has no value... I believe Slavoj Žižek had it right when he said, "It is therefore clear why vampires are invisible to the mirror: because they have read Lacan and, consequently, know how to behave... (Žižek, 19992, p. 126)"

Christian Metz's groundbreaking work in *The Imaginary Signifier* is the starting point from which I will make clear the metaphorical truth behind the absence of the vampire from the mirror. He theorizes that "...film is like a mirror... (Metz, 2000, p. 410)." He goes on to explain that, "...although... everything comes to be projected, there is one thing, and one thing only that is never reflected in it: the spectator's own body (Metz, 2000)." If this is true, then perhaps we have not developed our identification inside of this film-mirror through Lacan's mirror stage. Metz responds, "... what *makes possible* the spectator's absence- is the fact that the spectator has already known the experiences of the mirror... (Metz, 2000, p. 411)" Later he adds that because of this, "The spectator has the opportunity to identify with the *character* of the fiction (Metz, 2000, p. 411)." I'm not entirely convinced though that it is a simple identification that we are meant to make.

It is through Žižek and Metz though that my claim is ready to be revealed: the mirrors in which vampires cannot be seen are analogous to the film-mirror that we encounter when we go to the cinema to view one of these films. As such, it is clear that not only are we aligned with the vampire through the space we enter and the darkness we become enveloped in as Waltje has claimed earlier, but we *are* the vampires that we see in front of us. It is not a mythic, undead man with phallic teeth that we are being warned against; the vampire is our subconsciously primal sexual and violent desires, and we are seeing our mirrored selves in its eyes. *We* do not identify with the fictional character as supposed by Metz though; it is the vampire who is identifying with us.

Our vampire cannot be seen inside of the mirror.

Metz also adds that there is not only some sort of relationship between spectators and characters in the films but also with the equipment that films employ as well, "...the spectator can do no other than identify with the camera, too, which has looked before him at what he is now looking at... (Metz, 2000, p. 413-414)" This can be seen as an explanation of our absence from the film-mirror as well. If it is true that we see what the camera before us has seen, then the camera is, in fact, a surrogate for our sight in our absence from the set. What is projected upon the screen then, is our vision returning to us. This is only appropriate in regards to our absence because, supposing we were there to witness the acts being displayed for us on the screen, we wouldn't be able to see ourselves then either. So, if the screen, or film-mirror, is actually casting our own reflection when we see the vampire, then it is safe to assume that when the vampire looks into the mirror, s/he

must see us, and, when they do, they often react violently upon this reflection, frequently shattering the glass. But why?

If the vampire is meant to be "the embodiment of human evil (Wright, 1974, p. 45)," and/ or "the incarnation of unbridled sensuality (Wright, 1974, p. 45)," as Judith Hess Wright claims, then perhaps the vampire destroys the mirror because it sees in that instant that it is only one fragment of who we are and/ or who we can become and the idea that we have a choice to leave the theatre and its darkness behind is more than the vampire in us can bear. We have let the vampire in us escape into the screen for a few hours and when we drag it kicking and screaming, pushing it back down into our subconscious realm, it reacts in the same way a two year old reacts when hearing the word "no."

Through this, Žižek's joke about vampires having read and/ or at least having gone through Lacan's mirror stage holds more weight than he probably surmised when adding it to the page because, in fact, *we* all have. Waltje's earlier claim is unfinished. He goes on to say, "Having turned themselves [the spectators] into vampires, they are waiting for the film-vampire to come out and join them (Waltje, 2000, p. 29)." This is actually a half-truth. We spectators are merely waiting for the vampire within us to have its fun and then *rejoin* us once we see that side of us punished for its desires by the protagonists that we thrust it against. Matthew Bunson explains the vampire's aversion to mirrors in *The Vampire Encyclopedia*:

> "Folklore for this aversion stem from the concept that a mirror also reflects a soul, and evil beings have no soul to reflect. It has also been argued that the bloodsuckers actually exist in two worlds, that of the living and that of the dead. As it is in neither world completely, it will not be seen in a mirror (Bunson, 1993, p. 176-177)."

Without us, the vampire wouldn't have a soul to reflect, as it is nothing more than a decimal without our complete presence. Also, as I have just outline, the vampire does indeed live in two worlds. Apart from us, acting out its desires inside of the screen, it is dead. It cannot actually live without its true host. Before we release it into the screen and after we trap it once again after the film, it is a part of our whole and, as such, is alive with us.

It is also interesting to note that this spectator-screen-vampire relationship has not gone unnoticed and that films since Coppola's *Bram Stoker's Dracula* have actually taken this dynamic a step further than only showing us the vampire's absence from the mirror. Patrick Lussier's *Dracula 2000* includes a scene in which a young, voluptuous reporter (Valerie Sharpe) and her cameraman are attacked by Dracula. As the attack begins, the camera man sees Valerie seize up, and her neck is suddenly sliced open though no cause can be seen through the lens of the camera. Sharpe flees into the news van and watches in horror as Dracula manhandles her colleague. As she watches on the video monitor receiving the feed

from the camera, Dracula is absent from the screen. He is absent from this film-mirror just as we are absent from the film-mirror in front of us.

Valerie Sharpe is attacked.

The cameraman is being attacked by Dracula though he cannot be seen on the screen.

In *Shadow of the Vampire* (E. Elias Merhige, 2000), Max Schreck is fictionalized as being a real vampire during the filming of Murnau's *Nosferatu*. In a pivotal moment of the film, Schreck encounters a projector on his own, without the interference of Murnau and his crew. Stacey Abbott describes the scene:

> "Like a child amongst toys, he curiously begins to crank the lever resulting in an image of a sunrise being projected onto the wall. While he is transfixed by the sight of the first sunrise he has seen in centuries, the sequence changes meaning as soon as Schreck instinctively places his hand before the lens in order to protect his shadow on the screen. This equipment captures and projects a part of himself (Abbott, 2004, p. 3)."

The vampire, in this scene, is seeing what it is to be the whole without us. His shadow, the part of himself that Abbot is referring to, is representative of the vampire as part of us cast upon the screen.

It is not a surprise that in spite of this essay's claims, people will continue to flock to the theatres to unleash their inner vampires every time a new vampire film is released. For lovers of these films, it is a necessary evil, a period of time when they can allow these subconscious desires to manifest themselves before their eyes, relieving the tension that bottling these desires creates. Nina Auerbach shares that "...what vampires are in any given generation is a part of what I am... (Auerbach, 1995, p. 1)" Do not be afraid of the vampires that reveal themselves to you on the screen. Be afraid if you find yourself trapped in the darkness of the theatre, unable to bottle them back inside once the credits have rolled.

References

Abbott, S. (2004). "Spectral Vampires: Nosferatu in the Light of Technology." Horror film: Creating and Marketing Fear. Ed. Steffen Hantke. Jackson: University Press of Mississippi. 3-20.

Auerbach, N. (1995). Our Vampires, Ourselves. Chicago: University of Chicago Press,

Biggs, J. M., and Petrie, D. W. (2008). The Art of Watching Films. 7th ed. New York: McGraw Hill.

Bram Stoker's Dracula. Dir. Francis F. Coppola. Perf. Gary Oldman, Winona Ryder, Anthony Hopkins, and Keanu Reeves. DVD. Columbia Pictures, 1992.

Bunson, M. (1993). The Vampire Encyclopedia. New York: Crown.

Dracula (2000). Dir. Patrick Lussier. Perf. Gerard Butler, Christopher Plummer, Jonny Lee Miller, and Justine Waddell. DVD. Dimension Films.

Lacan, J. (2004). "The Mirror Stage as Formative of the Function of the I as Revealed in Psychoanalytic Experience." Literary Theory: An Anthology. Ed. Julie Rivkin and Michael Ryan. Malden: Blackwell, 441-446.

Metz, C. (2000). "The Imaginary Signifier." Film and Theory: An Anthology. Ed. Robert Stam and Toby Miller. Malden: Blackwell. 408-436.

Peters, F. (2006). "Looking in the Mirror: Vampires, the Symbolic, and the Thing." Vampires: Myths and Metaphors of Enduring Evil. New York: Rodopi Press. 177-187.

Shadow of the Vampire. Dir. E. E. Merhige. Perf. John Malkovich, Willem Dafoe, Udo Kier, Cary Elwes. DVD. Lions Gate Films, 2000.

Stoker, B. (2003). Dracula. New York: Barnes and Noble Classics.

Waltje, J. (2000). "Filming Dracula: Vampires, Genre, and Cinematography." Journal of Dracula Studies. Official Publication of the Canadian Chapter, Transylvanian Society of Dracula. 24-33.

Wright, J. H. (1974). "Genre Films and the Status Quo." Jump Cut. Jump Cut. 42-50.

Žižek, S. (1992). Enjoy Your Symptom!: Jacques Lacan in Hollywood and Out. New York: Routledge.

Further Reading

Bauer, R. N. (1999). "Hearing Dracula: Sound as Sign in Film." Romance Languages Annual. West Lafayette: Purdue Research Foundation 9, 138-143.

Bordwell, D. and Thompson, K. (2008). Film Art: An Introduction. 8th ed. New York: McGraw Hill.

Braudy, L. and Cohen, M. (1999). Film Theory and Criticism. 5th ed. New York: Oxford University Press, 610.

Clover, C. J. (1993). Men, Women, and Chainsaws: Gender in the Modern Horror Film. Princeton: Princeton University Press.

Dyer, R. (2001). "Dracula and Desire." Film/ Literature/ Heritage: A Sight and Sound Reader. Ed. Ginette Vincendeau. Berkley: University of California Press, 91-97.

Heldreth, L. G. (1999). "Vampires in Film & Television." Journal of Popular Film & Television. Washington: Heldref Publications.

Henderson, C. J. (2001). The Encyclopedia of Science Fiction Movies: From 1897 to the Present. New York: Checkmark Books.

Holte, J. C. (1999). "Not All Fangs are Phallic: Female Film Vampires." Journal of the Fantastic in the Arts. Armonk: Sharpe, Inc., 163-173.

Jenkins, H. (2000). "Reception Theory and Audience Research: The Mystery of the Vampire's Kiss." Reinventing Film Studies. Ed. Christine Gledhill and Linda Williams. London: Arnold, 165-182.

Kittler, F. (1997). Literature, Media, and Information Systems. New York: Routledge.

Konigsberg, I. (1998). "How Many Draculas Does it Take to Change a Lightbulb?" Play it Again, Sam: Retakes on Remakes. Berkley, University of California Press, 250-275.

Senf, C. A. (1979). "Dracula: The Unseen Face in the Mirror." The Journal of NarrativeTechnique. Ypsilanti: Eastern Michigan University Press, 160-170.

Siegel, S. and Siegel, B. (2004). The Encyclopedia of Hollywood: An A-Z Guide to the Stars, Stories, and Secrets of Hollywood. 2nd ed. New York: Checkmark Books.

Worland, R. (2007). The Horror Film: An Introduction. Malden: Blackwell.

Žižek, S. (1991). For They Know Not What They Do: Enjoyment as a Political Factor. New York: Verso.

REPLY TO WARD CHURCHILL

Andrew Feldheim
Philosophy Department

In this paper, I reply to Ward Churchill's contention that, in struggles against tyrannical regimes and oppressive political systems, nonviolent resistance is ineffectual without either corresponding violence or the threat of violence. My response attempts to show why nonviolent resistance is an effective method in its own right, and can be superior to violent alternatives in terms of accomplishing both short and long term objectives. Finally, I address a peculiar aspect of Mr. Churchill's position that, while insulating it from falsifiability, simultaneously limits both its credibility and usefulness.

In "Pacifism as Pathology: Notes on an American Psuedopraxis," Ward Churchill condemns nonviolent resistance as ineffectual unless accompanied by either violent resistance or the threat of such violence. In his words, "The essential contradiction inherent to pacifist praxis is that, for survival itself, any nonviolent confrontation of state power must ultimately depend either on the state refraining from unleashing some real measure of its potential violence, or the active presence of some counterbalancing violence of precisely the sort pacifism professes to reject as a political option" (Churchill, 1998, p. 44). His contention is that all nonviolent resistance must necessarily result in one of two outcomes: "1.) To render themselves perpetually ineffectual (and consequently unthreatening) in the face of state power. In which case they will likely be largely ignored by the status quo and self-eliminating in terms of revolutionary potential; or 2.) to make themselves a clear and apparent danger to the state, in which case they are subject to physical liquidation by the status quo and are self-eliminating in terms of revolutionary potential" (Churchill, 1998, p. 44). In other words, he claims that nonviolent resistance is either ineffective or extinguished before becoming effective.

I believe that Mr. Churchill has based his argument on a false dichotomy. Between the extremes of the impotent and the vanquished lie those movements, many of them nonviolent, who have achieved varying degrees of success on behalf of the disenfranchised and oppressed. The purpose of this paper is to prove the viability of this largely nonviolent middle ground and to highlight the flaws in Churchill's argument. My argument will consist of four parts. The first will propose an alternative to violent action in the form of nonviolent coercion. The second will show how Gene Sharp (2002) makes use of such nonviolent methods to construct a strategy designed to systematically undermine and, in some cases, disintegrate, tyrannical regimes. The third highlights some of the reasons why violent resistance may be less effective than its nonviolent counterpart. The fourth and final part shows that Mr. Churchill's contentions, while containing *some* elements of truth, have serious structural flaws that may cause one to question the strength of his conclusions.

Churchill seems to confuse nonviolent action with *inaction*. This is clearly not the case. As Kurt Schock makes clear, "[N]onviolent action is *active*- it involves activity in the collective pursuit of social or political objectives - and it is *non-violent* - it does not involve physical force or the threat of physical force against human beings" (Schock, 2005, p. 705). And herein lies the only limitation. All other forms of coercion, except physical violence, remain as viable options for the nonviolent proponent of social change. As will be outlined in the next section, the concerted application of economic and social pressure against repressive political systems can be of enormous value in achieving a greater share of justice for the oppressed, without resorting to physical violence or the threat of such violence.

In From Dictatorship to Democracy: A Conceptual Framework, Gene Sharp conducts a comprehensive examination, probing for theoretical and historical weaknesses within tyrannical regimes, that can be effectively, and nonviolently, exploited to undermine and, in some cases, disintegrate, these unjust systems. Sharp's logical strategy is clear and powerful: Locate the points where an oppressive regime is most vulnerable, and the points at which this system depends on the cooperation of the very people it oppresses. Then apply pressure to those points in order to further weaken its overall structure and exploit these dependencies.

In terms of well thought out strategies and tactics, Gene Sharp provides an impressive list that should be effective at weakening, and eventually dissolving, repressive authorities, and which also aims at the successful establishment of democracies after the dictatorships have fallen. It is noteworthy that Sharp's methods for accomplishing both short and long term goals have consistently proven to be effective in many different instances, without resorting to the violent tactics that Churchill insists are necessary. Notable examples can be found among the nonviolent actions taken in Chile during the late 1980s, when resistors exposed the illegitimate acts of torture perpetrated by their government (Deats, n.d.), and the Argentinean women who courageously and peacefully protested the illegitimate acts of kidnapping perpetrated by their government (Ruddick, 1989, p. 226).

Any overview of nonviolent struggles, however brief, seems incomplete without some mention of the paradigm examples of the nonviolent movements headed by Mahatma Gandhi, on behalf of his fellow Indians' struggle for equality and self-rule, and Dr. Martin Luther King, Jr., on behalf of the African-American struggle for equal rights in the United States.

Mahatma Gandhi, the great pioneer of nonviolent resistance, showed, through his incredible courage and relentless determination, the true power and effectiveness of nonviolence. Referring to Gandhi, Dr. Martin Luther King, Jr. said, "As I read his works, I became deeply fascinated by his campaigns of nonviolent resistance…As I delved deeper into the philosophy of Gandhi, my skepticism concerning the power of love gradually diminished, and I came to see for the first time that the Christian doctrine of love, operating through the Gandhian method of nonviolence, was one of the most potent weapons available to oppressed people in their struggle for freedom" (King, 1963, p. 72).

Churchill's assumption that, when a nonviolent group becomes a viable threat to an oppressive government, this group would be destroyed or become self-eliminating, does

not match the evidence, since both Dr. King's movement in the United States and Gandhi's movement in India survived their successes.

In many of the cases under discussion, violent resistance simply does not make good practical sense, and would result in unacceptable losses, while contributing relatively little to the overall cause. One reason for the inappropriateness of violent action in these circumstances is stressed repeatedly by Sharp and others. Implicit in almost all the situations under discussion is the assumption that the group *doing the oppressing* has more weapons, soldiers and all the other necessary ingredients for violent action than the *group being oppressed.* In such cases, if the subjugated group insisted on using violent means as their primary mode of operation, their movements would effectively amount to suicide, and the actual benefit to their cause would probably be negligible. It may be objected at this point that some groups and individuals protesting in this way, and willingly giving their lives, could inspire others through their martyrdom. Their sacrifice could possibly elicit the sympathy of other groups that may be willing to help their cause. In answer to this, one only has to consider the response that people are likely to elicit using this method in a nonviolent, as opposed to a violent, way. The Buddhist monks who immolated themselves in protest during the Vietnam War, and the followers of Gandhi who were beaten and killed, represent this method done nonviolently. They sacrificed their lives while making sure to harm no one else. In contrast, consider the suicide bomber who kills herself in the middle of a crowded marketplace, taking as many people as possible along with her. She is representative of martyrdom conducted violently. Both are types of martyrdom; but to the former, we assign virtue, while, to the latter, condemnation. The nonviolent variation is far more likely to elicit the support of other groups and even nations. It is practically superior.

There is also something to be said about the difference in results that are obtained when a dictatorship is overthrown through violent means, as opposed to nonviolent means. Gene Sharp notes that, essential to the removal of a dictatorship and the establishment of democracy, is a fundamental redistribution of the governmental power structure. Violence may be less conducive towards this goal. According to Sharp,

> A military coup d' etat against a dictatorship might appear to be relatively one of the easiest and quickest ways to remove a particularly repugnant regime. However, there are very serious problems with that technique. Most importantly, it leaves in place the existing maldistribution of power between the population and the elite in control of the government and its military forces. The removal of particular persons and cliques from the government positions most likely will merely make it possible for another group to take their place. (Sharp, 2002, p. 5)

Sharp feels that, unless the dictatorial power structure is changed to a more democratically oriented power structure, the stage is set for *another tyrannical group* to simply take the place of the deposed one.

At this point, in order to avoid the same reliance on absolutes that I find inappropriate in Churchill's argument, it is important to recognize the crucial role that the uniqueness of every situation has in determining the proper methods to be employed on behalf of an oppressed or subjugated group. There are kernels of truth in Churchill's contentions that

have more or less import in accordance with the specific situation. There are cases, both historical and theoretical, for which violent action seems the only logical alternative. The extermination of the European Jews during World War II seems like such an instance. One important difference between this example and the oppression of Indians by the English, or African-Americans by the United States, is that, in the case of the Nazis, it was not an instance of the same type of repression. They did not wish to subjugate or exploit the Jews; they simply wished to kill them.

Fortunately, however, the Nazi example is the exception rather than the rule. Most cases of oppression stem from a wish to subjugate a population in order to profit unfairly from their labors, or to usurp their property. To give Mr. Churchill his due, even in cases such as these, there may be factors, specific to particular situations, which call for violent resistance or a mixture of violent and nonviolent resistance. Each situation must be evaluated on its own merits. My contention is not that there is *no truth* in Churchill's position. Rather, it is his use of absolutes, his insistence that violence or the threat of violence is always necessary, that demands a refutation.

Churchill presents his contention, that violence or the threat of violence is a necessary constituent of successful resistance to tyranny, in a way that makes it unfalsifiable. This, however, does not add to its merit. To potential counterexamples, Churchill simply relies on the presence of groups which may be *potentially* violent. The very nature of tyranny, however, naturally encourages feelings of resentment and hostility on the part of the oppressed. If one looked hard enough, he could always find some indication of *potential* violence, even if not overt. Churchill's argument will, in this sense, always be true, but gives us no more actual information than a tautology. Also, since Churchill supposes a causal relationship between violent resistance and the defeat of dictatorships, and this construct is placed within an historical context, we can never know what would have happened if there had been no violence or the threat of violence, but only nonviolent resistance. While it may not be possible to prove Churchill's argument unsound, its very nature makes it of limited utility.

VI. References

Churchill, W. (1998). "Pacifism as Pathology: Notes on an American Psuedopraxis [sic]" Pacifism as Pathology. Winnipeg, Manitoba: Arbeiter Ring. 29-69.

Deats, R. (n.d.). "The Global Spread of Active Nonviolence." From the Author's website. http://deatspeace.tripod.com/global1.html

King, M. L. (1963). "Pilgrimage to Nonviolence." Christian Century Foundation, 1960. Rpt. In Instead of Violence: Writings by the Great Advocates of Peace and Nonviolence Throughout History. Ed. Arthur and Lila Weinberg. New York: Grossman. 69-75.

Ruddick, S. (1989). "Notes Toward a Feminist Peace Politics." Maternal Thinking. New York: Ballantine.

Sharp, G. (2002). From Dictatorship to Democracy: A Conceptual Framework For Liberation. Boston: Albert Einstein Institution.

Schock, K. (2005). Unarmed Insurrections: People Power Movements in Nondemocracies. Minneapolis, MN: University of Minnesota Press.

BARBARA KRUGER, *YOUR BODY IS A BATTLEGROUND*

Katherine Calak

Women's Studies Program

Barbara Kruger explores feminist theory through artistic expression. Her continued use of black and white imagery combined with red blocks of text derive from her background in art and design and her previous careers designing and editing magazines such as *Mademoiselle* and *House and Garden*. Kruger voices her concerns over feminist issues through bold images and text and in consequence, she is able to gain the viewer's attention in a manner that differs from that of many other artists.

This paper discusses patriarchal values, discourses of the gaze, and the crises in consumerism in relation to Kruger's work. The problematic social norms that arise from patriarchy are exposed through her designs. In much of Kruger's work, she reveals concepts with such an impact that the viewer is forced to further consider her or his own opinions and perceptions in relation to these issues. Kruger also works with the subject of the gaze and its relation to male power. Through the gaze, men are able to gain control over the bodies of women, creating a tension between the two sexes. Kruger ties this to a critique of consumerism, in that everything in Euro-American culture can be bought, sold, and owned. This ideology extends to relationships amongst individuals further causing struggles over control and power.

Art is not truly art unless it stands for something, gives a point of view or relays a message in someway. There is no possible way to create a visual piece that does not involve communication. A piece of artwork sends some sort of message for the viewer to process, even if only for a split second. When thinking of political or social views, one can often bring to mind an artist or a work of art that attempts to convey the message as well. Mixed media artist and graphic designer, Barbara Kruger sends powerful messages concerning feminism and women's rights through her artwork (Kruger, 1990). In much of her work, she uses bold, sometimes controversial black and white imagery along with text that is laid over blocks of vivid red color to catch the viewer's eye to evoke a concern over a particular topic. In her montage, *Untitled (Your Body is a Battleground)* from 1989, Kruger voices her opinion to protect women's rights through an image that also raises issues of power, patriarchy, stereotyping, and consumption.

Kruger's untitled work known as *Your Body is a Battleground*, shown in her book Love for Sale, depicts a bold black and white photograph with its meaning emphasized through red blocks filled with white text. The image is of a women's face split symmetrically along the vertical axis. There is a play of positive and negative space between the two halves of the image, highlighting ideas of "positive versus negative, white

versus black, good versus bad" (Kruger, 1990, p. 87). The figure's characteristics such as her hair and makeup display that of a 1950s style. Her gaze is directed straight ahead, making eye contact with the viewer. The design was executed in support of the 1989 march on Washington to support the rights of women with an emphasis on abortion (the right to choose) and birth control rights. Kruger stresses her own commitment to these issues by stamping the words "Your Body is a Battleground" down the center of the page.

In this piece, the words "Your Body is a Battleground" relate to an array of political and social stances. Ideas of power and structure in society are often referenced in Kruger's work. In the book Love for Sale, Kate Linker explains, "Power cannot be centralized; rather, it is diffused, decentralized, and in consequence, anonymous: it exists less as a 'body' than as a network of relations unifying social apparatuses and institutions" (Kruger, 1990, p. 27). Since power is something that cuts through all aspects of society, women are forced to defend themselves and their livelihoods in society. The words "Your Body is a Battleground," reference this constant fight in which women take part. The fight over a woman's right to choose what happens with her own body is one that intensely revolves around power. In her essay "Biologically Correct," feminist writer Natalie Angier writes that there is a struggle between men and women "over the same valuable piece of real estate—the female body" (Angier, 2003, p. 10). There is a constant push and pull between the sexes: women fight to have control over their physical bodies as well as their places in society, while men fight to maintain their dominance over women in society.

Thoughts of patriarchy are spoken through Kruger's artwork. According to Jane Pilcher and Imelda Whelehan, authors of Fifty Key Concepts in Gender Studies, "Literally, patriarchy means rule by the male head of a social unit (a family or tribe for example). The patriarch, typically a social elder, has legitimate power over others in the social unit, including other (especially younger) men, all women and children" (Pilcher, 2008, p. 93). As reviewed in psychologist and writer, Carol Gilligan's essay "Sisterhood is Pleasurable: A Quiet Revolution in Psychology," the traditional ways of society construct a barrier that forces women to be inferior to men. Patriarchy is formed when men take the dominant role in the family. If a man is the head of the family, then his ideas diffuse down through the inferior members of the family, causing these members to take his ideas and opinions as their own. In this way, the man's views dominate not only in his immediate family, but also in society because they become not just the dominant views, but also the only views (Gilligan, 2003, p. 98-100). Linker states, "Kruger's mission is to erode the impassivity engendered by the imposition of social norms: hence the gist of a work from 1982, in which the words 'We have received orders not to move' are superimposed on an image of an immobile woman's body, pinned against a wall" (Kruger, 1990, p. 28). In this image, Kruger portrays the controlling means of patriarchy by depicting a silhouette of a woman with sewing pins literally fastening her to the background (Kruger, 1990, p. 28). The pins reference "women's work" and help to show how women are kept in the home, or "in their place," and away from prestigious, male dominated careers.

Kruger bases much her design work on stereotyping, referring to its "domain as that of 'figures without bodies" (Kruger, 1990, p. 28). She attempts to seize and manipulate the stereotypes that Euro-American culture relies upon so greatly. Kruger has found inspiration for many a work in the words of British writer John Berger, who states in Ways of Seeing from 1972, "*Men act* and *women appear*. Men look at women. Women watch themselves

being looked at" (Berger, 2003, p. 38). The male gaze turns women into objects that are not permitted to think for themselves. Their thoughts, opinions, and behaviors stem from the constant judgments of men. In the work *Untitled, (Your Body is a Battleground),* Kruger challenges this issue by depicting the female figure in confrontation with the viewer. The subject's eyes glare directly at the viewer in defense. Kruger takes the opposite approach to this very subject in her piece, *Untitled (Your gaze hits the side of my face)*. In this piece, she shows a porcelain-like female head facing parallel to the viewer with text stacked down the left edge of the page reading, "Your gaze hits the side of my face." Linker suggests, "She may be referring to the power of the gaze to arrest—literally petrify—its object" (Kruger, 1990, p. 62). This implies that the male gaze prevents a true liberation of women. Women are still tied to the implications of patriarchy and the dominant male figure. In this piece, Kruger does not show what she wishes would happen when women are confronted as she does in *Untitled (Your Body is a Battleground),* but she confronts the viewer with what she feels are the despairing realities amongst which we live.[1]

As a slight break from most of her work, Kruger created a piece named *Untitled (You construct intricate rituals which allow you to touch the skin of other men),* which is featured in her book Love for Sale. Kruger portrays male figures as opposed to her typical female figures and confronts an issue from the opposite direction. In this piece Kruger pokes fun at the male persona by depicting a group of men who seem to be jokingly wrestling with text running across the right side of the page stating, "You construct intricate rituals which allow you to touch the skin of other men" (Kruger, 1990. p. 52). This piece takes the assumption that affection between men is generally unacceptable because by showing affection, one is taking on feminine personality traits. This leads to the stereotypes that women are frail and need a strong, sturdy masculine figure in order to perform in society.

Kruger relates her work to the consumer driven society in which we live. "It is evident in her use of red enameled frames to commodify her images, announcing their market status and pointing to the market as the irrefutable condition that no object—least of art—can evade" (Kruger, 1990, p. 76). Commodity is built into every aspect of our culture, even the purest objects or ideas—even the purest pieces of art are eventually given a value or a price tag. By having her work itself reflect consumerism, Kruger indirectly portrays women as being for sale as well. In this way, she shows how women are generally portrayed in society—as belonging to someone else. This idea of ownership is a direct result of consumer society.

Ideas of consumerism take more direct forms in work such as Kruger's *Untitled (I shop therefore I am)*. This piece shows a generic hand grasping a red block of color which contains the words, "I shop therefore I am." In Kruger's book, Love for Sale, there is a photograph of a woman carrying a handbag or shopping bag with this image printed onto it. In reference to this piece, Linker states, "…the consumer world, as Kruger remarks of 'Pee-wee's Playhouse,' is a place where *things* reign supreme" (Kruger, 1990, p. 78). Society

[1] Over time, theories of the gaze have developed, becoming more intricate. Feminist, Bell Hooks has addressed issues of the gaze in reference to race and sex (Hooks 94–105) and artist, Mary Kelly has analyzed notions of the gaze through psychoanalytic theory (Kelly 72–76).

does not take into account who we are, but instead what we have. This, in a way, causes people to be dormant things rather than actual beings, leaving us with a completely plastic society.

In a video named "Consumption," artists Michael Ray Charles, Matthew Barney, Andrea Zittel, and Mel Chin collaborate to show and explain work of their own which references a consumer driven society encompassed with mass production. Kruger and tennis star and sports commentator, John McEnroe, created the introduction to this video. McEnroe is shown briefly explaining what the viewer is about to see, while being periodically interrupted by flashes of Kruger's red and white blocks of text displaying messages such as "Love art, Buy art, and Sell art" and "Feed me, Love me, Buy me, Sell me" (Consumption, 2002). "Kruger's text addresses the viewer in much the same way advertisers sway a consumer to buy a product" ("Consumption and Contemporary Art"). The text also interrupts McEnroe's monologue similar to how our lives are constantly interrupted by intrusive advertisements.

Kruger uses photomontage as well as other forms of media to question the viewer with ideas of feminism and the ways in which society treats and often mistreats women. Generally, Kruger's work is united in that most of it stands for the same set of morals and values. In particular, her piece, *Untitled (Your body is a battleground),* portrays notions of power, patriarchy, stereotyping, and consumption. This design was originally created in support of a women's rights march on Washington. In the piece, the face of a woman directly confronts the viewer. The text plays off of ideas that the female body causes disputes over control between men and women. The red framing of the image as well and the bold text give feelings of advertising and commodity. Concepts such as these are dispersed throughout Kruger's work presenting her political and social views to society.

References

Angier, N. (2003). "Biologically Correct." Sisterhood is Forever: The Women's Anthology for a New Millennium. Robin Morgan. New York: Washington Square Press,

Berger, J. (2003). "Ways of Seeing." The Feminism and Visual Culture Reader. Ed. Amelia Jones. New York: Routledge.

Consumption and Contemporary Art." art:21. 2005. Art21, Inc. 18 March. 2007 <http://www.pbs.org/art21/legal_html>

Consumption. Wesley Miller. Perf. John McEnroe, Michael Ray Charles, Matthew Barney, Andrea Zittel, Mel Chin. FLAT, 2002.

Gilligan, C. (2003). "Sisterhood Is Pleasurable: A Quiet Revolution in Psychology." Sisterhood is Forever: The Women's Anthology for a New Millennium. Robin Morgan. New York: Washington Square Press.

Kruger, B. (1990). Love for Sale. New York: Harry N. Abrams, Inc.

Pilcher, J., and Whelehan, I. (2008). Fifty Key Concepts in Gender Studies. London: Sage Publications.

AN ETHIC OF CARE CRITIQUE

Thea Hassan
Women's Studies Program

The foundation upon which ethics should be based is debatable in the field of philosophy. Nel Noddings proposes "An Ethic of Caring" which offers a feminist approach to ethics based on caring. Noddings' ethic of care is not beneficial to feminism because it encourages caring as the sole basis of ethics, hinders a woman's ability to become autonomous, and reinforces traditional gender roles. Critiques offered by Victoria Davion, Jean Keller and Virginia Held offer insightful additions and modifications of Noddings' ethic of care. Noddings' ethic of care, while beneficial to general ethical guidelines, is not appropriate for feminism because it encourages traditional gender roles for women, ignores virtues besides care, and inhibits a person from becoming autonomous. The ethic of care could be improved by the addition of autonomy and justice into the theory.

Philosophers have long debated a moral foundation on which ethics should be based. In the light of feminism, a code of ethics which encourages women's liberation is sought as well. While Lawrence Kohlberg (1984) proposed a model for moral development with the basis of justice, many feminists reject this idea for women. Philosopher Nel Noddings (1995) rejects the idea of universal justice because it is based on an idea of sameness, and she believes that most situations are neither similar nor equal; therefore justice is not applicable. Nel Noddings proposes "an ethic of caring" which offers a feminist approach to ethics based on caring. Noddings' ethic of care is not beneficial to feminism because it encourages caring as the sole basis of ethics, hinders a woman's ability to become autonomous, and reinforces traditional gender roles (Noddings, 1995).

Noddings believes that ethics are motivated by a sense of duty. The sense of duty that Noddings suggests is the duty to care, and she splits care into two categories, "natural caring" and "ethical caring." Noddings suggests that natural care is an innate characteristic, citing that even animals take care of their offspring, yet we do not call them moral (Noddings, 1995, p. 9). Her conclusion that ethical care is dependent on natural care is an interesting statement, because it is true that it is sometimes difficult to decipher whether the motivation behind certain actions is due to an innate compulsion to act or a sense of obligation to act. However, suggesting that care is innate is to imply that it is a woman's nature to be compassionate and caring all the time. The ethic of care ignores the other virtues of a woman and replaces them simply with the virtue of care. In this way, Noddings' ethic of care is naming the woman as the eternal nursemaid.

Victoria Davion offers an insightful critique of Noddings' ethic of care, suggesting that moral principles are based on more than the virtue of caring (Davion, 1993, p. 169). She

indicates a variety of reasons why other virtues aside from care are necessary for an outline of ethics. One suggestion is that total engrossment in another individual could be dangerous, as well as a crutch to one's own morals. She gives the example of a woman supporting her Ku Klux Klan husband, thus displacing her own values by supporting his and becoming evil herself. This engrossment is obviously not beneficial to the development of a women's ethics or to thinking for herself. Jean Keller points out that heterosexual love often involves the woman's adopting her man's ethics as her own, thereby not keeping her own beliefs (Keller, 1995, p. 153). The ethic of care promotes total engrossment and displacement of one's own values in another individual, which is a risky act. Kohlberg defines his highest stage of morality as "following self chosen ethical principles" but if a woman is following another individual's ethical principles, then according to Kohlberg, she will never become completely morally developed (Kohlberg, 1984, p. 177).

Davion also suggests that justice is necessary in an ethical guideline (Davion, 1993, p. 171). Davion discusses Noddings' example of the woman who is torn between standing by her racist family or her black friend who is fighting against her family. She does not want to break the ties with the family, so eventually sides with them (Davion, 1993, p. 171). Ultimately, she is making the choice of acting immorally by displacing her own values for her family. The incorporation of other virtues into her decision-making in the light of certain dilemmas should be included, such as in the case an idea of justice in that all humans are equal. Therefore, Noddings' ethic of care is not very useful in areas such as racism or homophobia (Davion, 1993, p. 163). Another example is if a woman's husband is a homophobe and asks her to vote against same-sex marriage, she is expected under the ethic of care to do as he says in order to not break the caring relationship. The problem is, discrimination against homosexuals is wrong and unjust; therefore it is questionable whether she is really acting ethically by acting caringly.

Davion also offers an insightful suggestion that the one-caring relationship model proposed by Noddings is not suitable for adult relationships. Adult relationships should be based on reciprocity, not one-way caring (Davion, 1993, p. 167). This idea of one-caring entraps a woman into the role of permanent caretaker, with her ethics based on her ability to care. Furthermore, according to Noddings, if a woman is in a relationship that is not reciprocal and she wants to end this relationship, this goal is not ethical because she is terminating a caring relationship. It is not a fair accusation to call someone immoral for ending a relationship in which there is not reciprocity of care. A relationship should also be allowed to be ended when it interferes with one's morals. Davion states, "I believe that what is missing from Noddings' account is an account of the individuals within caring relations as important in themselves" (Davion, 1993, p. 175). In fact, if Noddings had included a higher priority on the position of the individuals, her theory may have been more suitable for feminism, allowing women to claim a self-identity and therefore become autonomous. Noddings does state, perhaps in realizing her lack of reciprocity in the one-caring relationship, that the care ethic "does not separate self and other in caring, although, of course, it identifies the special contribution of the one-caring and the cared-for in caring. (Noddings, 1995, p. 26). She is saying that if all relationships are one-caring, including the relationship with oneself, then self- interest will be kept in mind. This is not necessarily

true, because if one is in a caring relationship that is harmful to health or spirit, according to the ethic of care, it is wrong to end the relationship. The only person benefiting from the relationship is the cared-for, while the one-caring remains as a doormat. It is also ironic that Noddings mentions reciprocity in regard to animals; she does believe we should not have a one-caring relationship with an animal because there is no potential for it to show reciprocity (Noddings, 1995, p. 10). This view is slightly hypocritical, because surely there are people with whom we may engage in a one-caring relationship and who also show no potential for reciprocity, and yet this relationship is still justified by the ethic of care.

Jean Keller's suggesting that care ethics is undermining to women's autonomy is indeed supportive of the claim that Noddings' ethic of care is not suitable for feminism (Keller, 1995, p. 153). Care interrupts autonomy by interrupting a woman's development of being able to think on her own. Under the obligation to care, the ability of a woman to choose her relationships is diminished (Keller, 1995. p. 157). Keller also points out that it is unclear whether women's caring is due to socialization or to innate tendencies, as Noddings suggests. Although it is true that certain hormones in females after birth encourage nurturing behavior, it seems unrealistic to suggest that a female of any other species aside from our own would continue to have responsibility to care for the offspring when it is capable of caring for itself (Flemming et al. 1997, p. 145). Furthermore, if caring in women is merely a development of the socialization process that women go through, it would be difficult to abandon their socialization to "act in accordance with what our authentic self wants," and thus become autonomous, without simultaneously becoming uncaring (Keller, 1995, p. 145). Noddings suggests that the development of the "ideal self" should be "developed in congruence with one's best remembrance of caring and being cared-for" (Noddings, 1995, p. 22). Therefore, she is rejecting the idea of autonomy.

Care ethics also reinforces the idea of traditional roles of women as the homemakers, the caretakers, and the self righteous (Keller, 1995). In an ethic of care, a woman is expected to be the one-caring in all situations, thus forcing her to remain in the position of sole caregiver. By reinforcing these gender roles, the woman is caught in the role of a subservient person, caring for others but not for herself. It also reinforces an obligation to care while forgetting one's own needs. Noddings does state that if a woman is one-caring, she will care for her own needs, yet the ethic of care implies a certain neglect of the needs of the one-caring person. The one-caring is expected to give unconditional caring with the prospect of potential future reciprocity, no matter what her personal desires may be. According to Carol Gilligan, a woman who sacrifices her needs for the needs of others would be at level two of her moral development model for women. Gilligan explains her level two as "defining the self and proclaiming its worth on the basis of ability to care for and protect others (Gilligan, 1977, p. 496). Therefore, according to Gilligan, this woman would not yet be fully morally developed.

Noddings' ethic of care could be greatly enhanced by the addition of other virtues aside from care, such as justice or room for autonomous growth. This expansion would help boost her ethic of care to be an influential philosophy for feminism as well as well-rounded ethical guide for all persons. Her introduction of care into ethics is beneficial, and certainly much of ethics should be based on care. However, the problem with her one-caring model

is that she neglects other aspects of morality. Virginia Held offers a "meshing of care and justice" in order to improve social relations and public policy, stating that the two are both compatible and a useful combination (Held, 1995, p. 128). Noddings, however, cites specifically that justice should not be considered in moral decisions because moral judgments are not absolute, and moral judgments should be driven by obligation, not justice (Noddings, 1995, p. 22). However, a certain amount of justice is necessary, because not every person will feel obligated to care and therefore should rely on a sense of justice. Without justice, issues such as racism would still persist. Furthermore, Noddings tries to sneak some self-worth into her ethic of care, but her theory is mostly encouraging non-autonomous behavior. If women are to be liberated, they must be able to think for themselves and make autonomous decisions. Therefore, the ethic of care needs an element of self-development and autonomy incorporated into it.

In conclusion, Nel Noddings' ethic of care does make a positive contribution to ethical guidelines, by incorporating care into moral behavior. However, her ethic of care is not appropriate for feminism because it encourages traditional gender roles for women, ignores virtues besides care, and inhibits a person from becoming autonomous. An appropriate aim for feminists' ethics should focus on the morals of an autonomous woman and equality of men and women in ethics, giving neither the handicap.

References

Davion, V. (1993). "Autonomy, Integrity, and Care." Social Theory and Practice*:* 161-182.

Flemming, A., Ruble, D. Krieger, H. and Wong, P. Y. (1997). "Hormonal and Experimental Correlates of Maternal Responsiveness during Pregnancy and the Puerpium in Human Mothers." Hormones and Behavior : 145-158.

Gilligan, C. (1977). "In a Different Voice: Women's Conceptions of Self and Mortality." *Harvard Educational Review*: 481-517.

Held, V. (1995). "The Meshing of Care and Justice." Hypatia 46: 128-133.

Keller, J. (1995). "Autonomy, Relationality, and Feminist Ethics." Hypatia: 128-133.

Kohlberg, L. (1984). The Psychology of Moral Development. New York: Harper.

Noddings, N. (1995). "Caring: A Feminine Approach to Ethics and Moral Education." Rpt. in Justice and Care, Ed. Virginia Held. Boulder, CO: Westview: 7-30.