ARIZONA JOURNAL OF INTERDISCIPLINARY STUDIES

VOLUME 8 - SPRING 2022

RESEARCH, INNOVATION & IMPACT
Societal Impact
confluencenter
for Creative Inquiry

ARIZONA JOURNAL OF INTERDISCIPLINARY STUDIES

VOLUME 8 | SPRING 2022

CONTENTS

INTRODUCTION FROM RII

The Office of Societal Impact, within the Office of Research, Innovation, and Impact (RII), is proud to support the Arizona Journal of Interdisciplinary Studies (AJIS). We aim to expand equitable access to high quality undergraduate research and inquiry experiences across disciplines and are committed to showcasing the efforts of undergraduate student researchers. Providing a wealth of opportunities ensures that our future is full of a diverse group of leaders forging innovative pathways, forming powerful collaborations, and making remarkable discoveries.

It has been a delight to see this undergraduate-run student initiative take on new life, beginning with its revival in 2020-21 and building greater stability this year to allow for continuity of the journal, leaving it in an excellent position to expand in new directions moving forward. Notably this year AJIS has broadened their team, keeping in place the editor-in-chief position and eight general editors while adding an assistant editor-in-chief, a graphic designer, and an additional general editor position. The journal has also joined the National Undergraduate Consortium for Science Journalism, a prestigious network of 17 STEM-focused undergraduate journals at universities and colleges across the country. Two AJIS representatives attended the National Undergraduate Conference on Scientific Journalism, organized by the Columbia Undergraduate Science Journal editorial board. The leadership, creativity, and dedication shown by the entire AJIS staff has been inspiring!

The papers represented in this 8th issue of the journal display the wide range of innovative undergraduate research occurring at the University of Arizona. Addressing large societal and environmental challenges will require interdisciplinary action, creative solutions, and top-notch scholarship. AJIS and its team of dedicated student editors have ensured that the journal showcases these qualities, providing a forum for undergraduate research while also allowing AJIS staff to develop their skills as an editorial team. In a time when many conflicting ideas and values are wielded like weapons, AJIS provides a scholarly platform to focus on understanding the challenges facing us and the work being done to create solutions.

We are dedicated to continuing our support of AJIS as it gains more solid footing in the UArizona community through faculty, staff, and student partnerships. Please join us in congratulating this year's AJIS team on all of their hard work. We'd especially like to acknowledge the commitment of senior Robert Lowell, Editor-in-Chief, who reestablished the journal in the summer of 2020. We wish Robert and all of the graduating seniors all the best in their future endeavors!

Kimberly Sierra-Cajas
Director, Undergraduate Research & Inquiry
Interim Director, STEM Learning Center
Societal Impact Office

Rebecca Lipson
Undergraduate Research and Special Projects - Coordinator
Societal Impact Office

INTRODUCTION FROM THE CONFLUENCENTER

On behalf of the Confluencenter for Creative Inquiry, we are delighted to support the Arizona Journal of Interdisciplinary Studies, Spring of 2022 issue. The journal reflects students' intellectual curiosity about interdisciplinary research and creative activities across our campus.

Many of these journal entries represent much of the rhetoric and challenging obstacles we see in modern society, including emphasizing the role that humanities, arts, and science plays with respect to the topics in this issue.

The diversity of topics, ranging from regional healthcare disparities to agricultural and biological technologies, as well as mindfulness and family care, demonstrate the varied interests of our students and reflect the many ways in which creative inquiry manifests itself on our campus.

We are proud to be part of this editorial effort which exemplifies how innovative and interdisciplinary scholarship can transcend borders, disciplines, and intellectual differences. The dedication of this journal's undergraduate editors to cross-campus collaboration, to intellectual curiosity and rigor, and to supporting and promoting the interdisciplinary work of their fellow students, reflects the core values of this journal and of Confluencenter for Creative Inquiry.

We also recognize the hard work of the student-run editorial team as they were able to expand their team and outreach this year. They received a record number of student submissions, resulting in the addition of an extra journal entry in this issue as compared to last year.

Join me in recognizing and congratulating these inspiring student editors and authors, and our campus partners, as we continue this collaboration towards supporting innovative, interdisciplinary undergraduate scholarship.

Javier Duran, Ph.D.

Director
Confluencenter for Creative Inquiry

LETTER FROM THE EDITOR

This year has been a year of organizational development for the Arizona Journal of Interdisciplinary Studies. As I, and many of the editors who helped revive the journal last year, prepare to graduate this spring, we have been working hard to create the roles and procedures necessary to ensure that the journal will have a sustainable future. I am confident that, in this, we have achieved our goal.

In addition to the continued support from the University of Arizona Libraries and the Confluencenter of Creative Inquiry, the journal entered a new partnership this year with the University of Arizona's Office of Research, Innovation, and Impact (RII). The support of our new sponsor allowed us to pay each student on the editorial team for the first time, and to create two new paid positions, a full-time Graphic Designer and an Assistant Editor-In-Chief. I am sincerely grateful to RII's Rebecca Lipson for her guidance throughout the year.

I would also like to thank Ellen Dubinsky, from the University of Arizona Libraries, Dr. Susan Crane, our faculty advisor, and Angela Martinez, from the Confluencenter for Creative Inquiry, for their enthusiastic support. Thank you to everyone on our amazing editorial team for your outstanding commitment to the editing process, to Michael Miscio, our Graphic Designer, for working so hard to design our promotional materials and our final product, and to Theodore Lowell, our Assistant Editor-In-Chief, for all your fantastic work. Finally, thank you to all of the authors who made submissions this year. None of the pages that follow would have been possible without the contributions of each of these individuals.

It is hard to believe it has already been two years since we began the process of reviving this journal—it has been a fantastic adventure. I am happy about how far we have come, and excited to see where the journal goes from here. It is my pleasure to present Volume 8 of the Arizona Journal of Interdisciplinary Studies.

Sincerely,

Robert Lowell
Editor-In-Chief, Spring 2022

TRENDS IN ALFALFA GROWTH AND GROUND-WATER LEVELS IN ARIZONA

Matthew T. Ford

001

Abstract:

The federal government has been providing significant subsidies to the dairy industry since 1933. These subsidies are important to farmers and to the industry as a whole because they keep incomes steady during fluctuations in market prices. However, federal policies can also incentivize dairy production which increases agricultural production which has negative impacts on water resources. Here, we explore the impacts of dairy subsidies on groundwater storage in Arizona. On one hand, the dry climate, abundant sunshine, and good soil make Arizona an attractive location for alfalfa farms, and alfalfa is a major source of feed for the dairy cow population. However, Arizona has very limited surface water supplies, and irrigated agriculture often relies on groundwater. Groundwater use is unregulated in many rural parts of the state, which creates the potential for unsustainable pumping to support water-intensive crops like alfalfa. We present a retrospective analysis of alfalfa and dairy expansion across the state using datasets from CropScape, which uses satellites to determine ground cover, and the United States Department of Agriculture (USDA, 2021). Using data from the Arizona Department of Water Resources (ADWR), we explored how much alfalfa is being irrigated by renewable versus-nonrenewable sources. Finally, we explored connections between alfalfa and groundwater levels.

Preliminary results show a correlation between increased alfalfa growth and declining water levels in areas where groundwater is alfalfa's main irrigation source. Future work will explore spatial patterns in alfalfa expansion and groundwater declines relative to different regulatory frameworks across the state.

Introduction

How have changing groundwater levels and use affected agricultural decisions on crop choices, irrigated acreage, irrigation intensities and energy consumption inside versus outside Active Management Areas (AMAs). Groundwater is an important water supply for irrigated crops. Crop irrigation accounts for 43% of water use worldwide and 53% in North America (Siebert et al. 2010). As surface water supplies decrease, farmers will increasingly rely on groundwater supplies to irrigate crops. Decreasing groundwater levels could threaten drinking water supplies, food supplies, and energy systems.

In the mid-1900s, significant advancements in groundwater pumping technology allowed pumps to reach greater aquifer depths and rapidly pump large volumes of water. Currently, groundwater comprises 50% of the source water for irrigation (Dieter et al. 2018). The energy demand required to pump groundwater for irrigation is usually larger than delivering surface water. Farmers consume large amounts of water and energy but have relatively small contributions to support jobs and economic growth (Nesheim et al. 2015).

Groundwater usage rates are only sometimes monitored in agricultural use therefore, finding reliable well and pumping data outside the AMAs can be challenging. In the United States only 7.5% of crop acreage is irrigated, but irrigated crops account for 55% of all crop sales (U.S. Department of Agriculture, 2014). These data suggest that high-value crops are mainly grown in large areas of irrigated farmland, especially in the Desert Southwest. Water and energy use in irrigated farmland largely occur in the western United States where energy demand is already strained and water supply is low (Hitaj & Suttles, 2016). Assessing the energy cost at which farmers start to shift from low-value crops to high-value crops could be crucial in predicting future crop shifts and how those shifts could affect the economics, food supply, and water supply in a given area.

The independent variables include changing groundwater levels over a set period of time, and the stakeholder use of that groundwater (e.g. municipal, agricultural, industrial, etc.). The dependent variables include crop choices, irrigated acreage, irrigation intensities, and energy consumption.

I hypothesize that decreasing groundwater levels lead to an increase in the amount of energy required to pump the same volume of water, which in turn will cause a shift from staple crops to non-staple crops. A staple crop is one which sells for low monetary value and is a standard portion of a person's everyday diet. I believe irrigated acreage and irrigation intensities will vary from low to high depending on the specific crops' water needs and acreage requirements. I suspect a larger increase in irrigated acreage will have occurred outside the AMAs where there is little to no groundwater regulation, compared with inside AMAs, where regulations are strictly followed. The scope of this project will be on the Phoenix, Pinal, Tucson, Santa Cruz AMAs as well as southwestern Arizona, which is not part of any AMA. All areas are located within the state of Arizona in the United States.

Definitions

Groundwater- Water that is found and stored beneath the earth's surface in small pore spaces between sand, soil and rocks.
Surface Water - Water that collects on the earth's surface such as in lakes, rivers, reservoirs, etc.
AMA- Active Management Areas. These are five watersheds the with strict groundwater regulations and conservation programs to reduce groundwater withdrawals. The goal, as stipulated by Arizona's Groundwater Management Act of 1980, is for each AMA to achieve "safe yield" (i.e. groundwater withdrawals ≤ groundwater recharge) by 2025.
Staple crops- Crops that sell for low monetary values and are a standard portion of a person's everyday diet. Some examples include: cereals, legumes, tubers and root crops.
Non-staple crops- Crops that sell for large monetary values such as vegetables, fruits, flowers, condiments and spices.

Literature Review

Numerous researchers (e.g. Harou & Lund, 2008; USDA, 2014; Dieter et al., 2018) have studied changing (i.e. dropping) groundwater levels and how this has affected farmers' crop choices, irrigated acreage, irrigation intensities, and energy consumption. Harou & Lund (2008) studied groundwater overdrafts in Tulare basin in California, where there is also sparse groundwater management or regulation. They observed that groundwater overdraft can have large system-wide consequences on the aquifer and sometimes has the potential to make water withdrawals economically impossible. To minimize groundwater overdraft, Harou & Lund (2008) described various solutions, including: increasing runoff capture and infiltration; taxes and fees; water conservation policies; relocation of high-water-use crops; using or importing other water sources; and land cycling. Land cycling is a strategy to optimize a specific crops' water needs by mimicking the natural precipitation patterns of the region. Farmers can inexpensively pump water out of the ground but at some point the groundwater will be too deep and it will become too expensive to pump (i.e. when groundwater pumping costs exceed crop sale prices). Farmers use many methods to combat this such as water trading, smart irrigation, etc., which have been extremely successful in certain areas when managed correctly. Groundwater over-drafting in this area of California could cost farmers around $31 million annually in extra groundwater pumping costs. In 2014, the USDA reported that sprinkler systems were the most common irrigation method, followed by gravity fed systems, and then drip irrigators. Dieter et al. (2018) also showed that irrigated lands in the USA have become more efficient, with a 10% increase in sprinkler use and 11% decrease in flood irrigation, meaning there is less water lost to evaporation. New technologies and infrastructure will hopefully buffer the continued overdrafting of the system.

Howitt et al. (2015) discuss how to minimize economic losses during drought. Some processes they suggest are groundwater substitution for surface water use. Drought can potentially cause fallowed land, crops losses, and job losses. Howitt et al. (2015) showed that economic hardships were not equally distributed over their study area and that areas with limited groundwater resources experience the largest economic impact during drought. Groundwater use on crops in the California central valley is more expensive than delivering surface water, but this can slightly be offset by increasing crop sale prices at harvest time. They estimated that the surface water shortage was 48% less than in a normal water year. This led to a 72% increase in groundwater pumping with a total net water shortage of around 10%. The amount of idle land was 45% greater than during a normal water year (Howitt et al. 2015).

Groundwater pumping costs increased by 75% in response to the drought. The most fascinating thing associated with this study were these statistics. One might expect crop losses to be large, but crop revenue losses accounted for only 2.6%. These minimized revenue losses are the result of the success of water trading that typically takes place with municipal water companies, which are less subsided than farm water. In fact, farms can make more money by trading their water to municipalities than growing crops.

Siebert et al. (2010) discussed how groundwater use in irrigation is becoming a more widely-used practice worldwide. It accounts now for 43% of total irrigation water and around 545 (km3). The three largest groundwater users are India, China, and the USA. Irrigation water accounts for 70% of global freshwater withdrawals, and 90% of total global water use is for irrigation. Worldwide only about 38% of cropland can be irrigated with groundwater, for many reasons. Dieter et al. (2018) discussed the history of water use for irrigation in the United States, starting with groundwater pumping for irrigation, which steadily increased from 1950-1980, at which point it peaked at 150 bgal/day. From 1980 to 2005, irrigation withdrawals remained steady at 127 bgal/day, and has decreased to 116-118 bgal/day) today. The source of irrigation has shifted: surface water deliveries have decreased by 14% and groundwater withdrawals were 13% greater than in 2010 than in the 1980s. Withdrawals for irrigation were 37% of the total water withdrawals, and 17 states west of the continental divide accounted for 81% of total irrigated lands in the USA (Dieter et al, 2018). Increased reliance on groundwater for irrigation will cause groundwater tables to drop, and negatively impact the cost of pumping.

Scanlon et al. (2012) addressed groundwater depletion and sustainability of irrigation in the central versus western areas of the USA. In particular, they compared the High Plains aquifer in the Midwest with the Central Valley aquifer in central California, which are the two largest aquifers used for irrigation in the USA. Pumping for irrigation has caused a 36% and 15% decline in groundwater levels in these two aquifers, respectively. Scanlon et al. (2012) showed that depletion varies spatially: the northern regions of each aquifer had increases in groundwater table levels, in contrast to the southern regions, where pumping was dominant, and thus had groundwater declines. Modeling in the same area of the High Plains aquifer was conducted by Condon & Maxwell (2014), and included the interaction of surface water and groundwater over a simulated (hypothetical) 20-year period which was chosen because it would contain climate variability.

The model ran simulations where crops are irrigated using groundwater only, and also a constant pump model was analyzed to simulate irrigation against a base scenario where model results showed exactly what one would expect: groundwater pumping affects the temporal dynamics of water table depth by causing it to continue to drop over time rather than remaining relatively static. These modeling results were important because they confirmed that natural recharge cannot replenish the groundwater quickly enough to compensate for large-scale agricultural withdrawal. Irrigation makes groundwater systems highly variable, which makes them vulnerable, especially in times of drought when recharge to the system is low. Another interesting conclusion is that the water table in recharge areas is relatively stable, and doesn't experience a large water table decline that is common in pumping areas.

Castle et al. (2014) analyzed surface and groundwater depletion in the Colorado River basin from 2004 to 2013. Approximately 77% of all surface and groundwater water lost to evapotranspiration and consumptive use throughout the Colorado River basin is groundwater. During the 10-year period of study, the basin experienced moderate to severe drought in which surface water levels dropped only slightly at 0.9 km3/yr while groundwater levels dropped sharply at 5.6 km3/yr. This decline is concerning because there are strict usage requirements on surface water and far fewer regulations on groundwater use. Fewer regulations make groundwater more appealing to farmers, and the result is the depletion of the aquifer faster than it can replenish itself, and this has potentially severe consequences for all stakeholders who rely on the aquifer.

The goal of the current project is to explore how changing groundwater levels within and outside regulated areas have affected agricultural choices and to determine if groundwater regulations are driving crop choices and spatial distribution throughout Arizona. This information will be useful in future groundwater modeling to accurately predict drawdown due to agricultural pumping in southwestern Arizona. Groundwater level changes were analyzed throughout the state of Arizona using data from the Arizona Department of Water Resources (ADWR). Crop selection and locations were analyzed using the United States Department of Agriculture CropScape (USDA, 2021) and agricultural statistics survey (United States Department of Agriculture, 2019).

Site Description

The domain of my research project is the southwestern part of the state of Arizona and includes the counties of La Paz, Yuma, Maricopa, Pinal, and Pima, all shown in red in Figure 1. Arizona has a diverse range of physical features. The northeast corner consists of the Colorado Plateau (5,000-11,000 ft above mean sea level (AMSL)) which is mostly high desert with small areas of forests. The southwestern portion of the state contains basin and range provinces with multiple valleys in between several small mountain ranges. The desert regions have high daytime temperatures and cooler overnight temperatures and receive low amounts of rainfall (e.g. 12"/yr) compared to the higher elevation locations (e.g. 20"/yr). The transition zone between low and high desert runs from southeast to northwest and contains both desert and mountainous terrain. Figure two is a topographic map of Arizona that illustrates the transition zone, with low desert located in the southwest part of the state, depicted in deepening shades of green with increasing elevation, and high desert (950 to >3750 m AMSL) depicted in shades of light orange to red with increasing elevation. The large range of physical features in Arizona make it a fascinating location for this project.

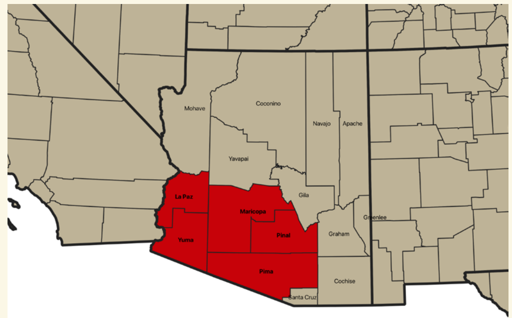


Figure 1: Arizona county map highlighting the southwestern region

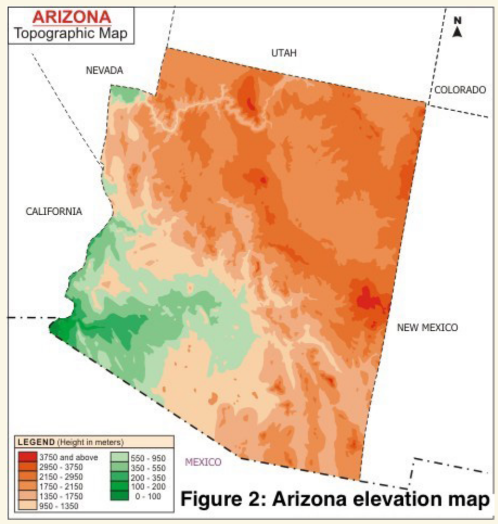


Figure 2: Arizona elevation map

I have focused much of my research on the Phoenix and Pinal AMAs, La Paz County, and the Harquahala irrigation non-expansion area (INA) (see Figure 3). Areas outside of AMAs such as Yuma and La Paz counties (see Figure 1) were also part of this study, and are important because they are the location of Arizona's largest agricultural regions outside the regulated groundwater areas of AMAs.

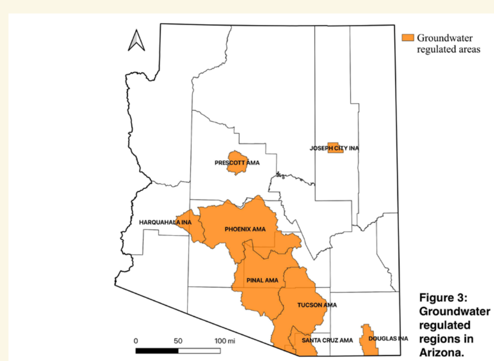


Figure 3: Groundwater regulated regions in Arizona.

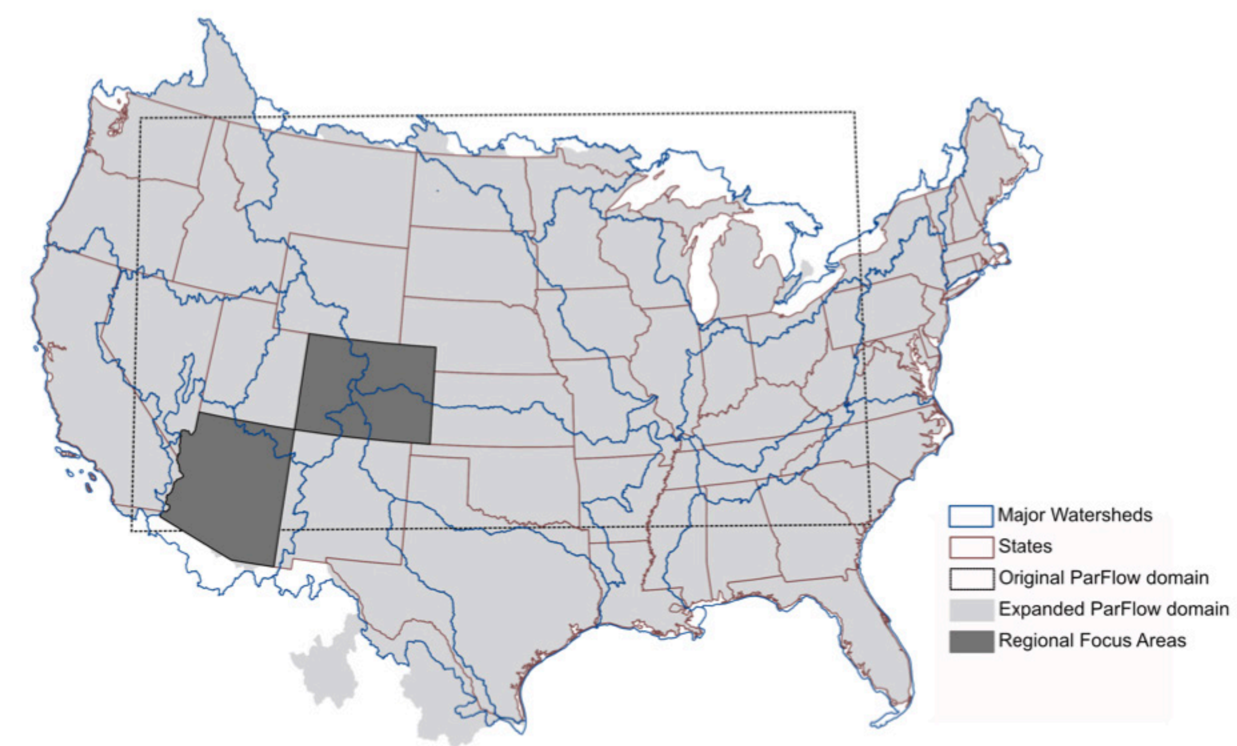


Figure 4: Map of the continental United States highlighting Arizona & Colorado

Yuma County is especially important because it is also a large agricultural region that continues to grow, with over 120,000 irrigated areas of farmland in 2010 (Yuma County Agriculture Water Coalition, 2015). The project is likely to expand into other states such as Colorado for my master's thesis (see Figure 4).

The abundant sunshine in Arizona makes it a very favorable location for growing crops. The annual precipitation in the driest parts of the state (e.g. near Yuma) is often less than 4 inches per year, which is insufficient to support most crops, and therefore crops must be irrigated using surface water or groundwater. There are two separate rainfall seasons in the Sonoran Desert region. In the winter months, steady light to moderate frontal storms passes through the area from the Pacific Ocean. Rain is the dominant precipitation type in the valleys but snow commonly occurs in locations above 4,000 ft above sea level. The other rainfall period is during the summer monsoon. During the monsoon, there is a seasonal shift in wind patterns, and moisture from the Gulf of Mexico, Gulf of California, and the Pacific Ocean are blown over state. Uneven summer heating of the land causes the moist air to rise and condense, which causes intense and violent thunderstorms that can drop inches of rain in short time periods. All other months are generally very dry and rain seldom occurs.

Materials & Methods

The objective of this research was to investigate how changing groundwater levels and use affect agricultural decisions on crop choices, irrigated acreage, irrigation intensities and energy consumption within versus outside AMAs. Data showing groundwater levels and pumping rates in Arizona were obtained from the Arizona Department of Water Resources known as the ADWR, which is a government agency dedicated to protecting and enhancing state water supplies. Groundwater levels from about 45,000 wells in Arizona were procured from the ADWR database. This study focused on depth-to-water measurements from 1910-2020 in four groundwater subbasins: Gila Bend (GIL), East Salt River Valley (ESR), Eloy (ELO), and Maricopa-Stanfield (MST) (Figure 5). This is a fairly comprehensive history of wells drilled in Arizona going back to the 1800s. Most of these wells do not have water level measurements associated with them, and were thus not relevant to this particular study. A group of 1000 index wells were selected for use in this study. Index wells have been measured at least yearly since 1984 by ADWR personnel. There is also a small group of wells that have pressure transducers in them. These transducers are automated devices that measure the water level in wells at specified intervals – in this case every 15 minutes – and record the time and associated water level in a data logger. The data from the logger is then downloaded onto a computer or wirelessly updated to the ADWR database. Water level and pumping data were downloaded and linked to each well's identification number and put into a QGIS file for a visual representation before they were analyzed.

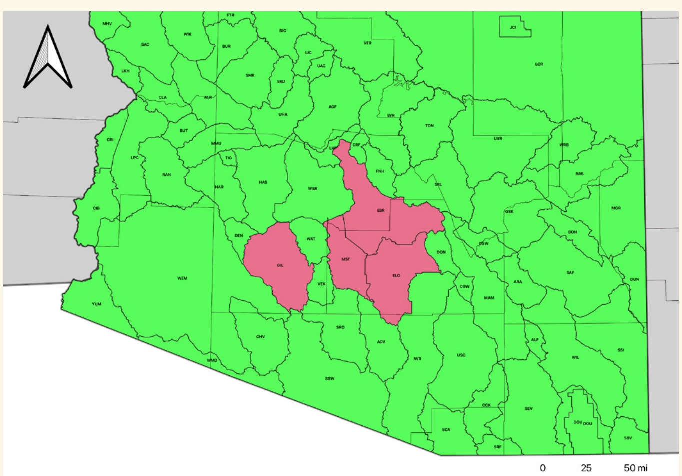


Figure 5:

Depth-to-water measurements were obtained from ADWR's Groundwater Site Inventory (GWSI) online database for the groundwater subbasins (highlighted in red).

As mentioned previously, the domain of this research project is the state of Arizona with the possibility of expanding into other states for my master's thesis research. I have focused on the Phoenix, Pinal, AMAs, and the Harquahala irrigation non-expansion area (INA), as well as Yuma and La Paz counties, which are neither part of an AMAs nor an INA. Areas outside of AMAs and INAs are especially important because they are the location of Arizona's largest agricultural regions outside the regulated groundwater areas of AMAs. In unregulated areas, reporting the quantity of water pumped is voluntary and therefore limited. The water level data in Arizona dates back to the late 1800s and continues through today but pumping data are available only from1984 to - the present. Water level measurements from the late 1800s and early 1900s were taken so infrequently that the data for this time period are not useful for this study. Therefore, the period of study was from 1984 to the present.

Our research team determined that the Arizona Department of Water Resources (ADWR) and the United States Geological Survey (USGS) would be able to provide the largest and most accurate public-domain datasets. These datasets were chosen because of a large temporal and spatial range, and both government agencies have a history of strict quality control procedures when collecting data measurements. Data showing groundwater levels and pumping rates inside Arizona were obtained from ADWR and USGS. Groundwater levels from all wells inside of Arizona were taken from Wells 55, which is a database of approximately 45,000 wells drilled between 1800-2020. A smaller group of 1,450 strategically-selected wells make up the Groundwater Site Inventory (GWSI) index wells which have been measured at least yearly by ADWR since 1984. Agricultural data for Arizona were gathered from the USDA Agricultural Statistics Survey (United States Department of Agriculture, 2019), which is completed every five years. While the complete ADWR dataset ranges from 1840-2017, our main focus was 1930- 2017 since government dairy subsidies started in the 1930s. Satellite data were acquired from CropScape to identify alfalfa field locations in Arizona; this dataset ranges from 2009-2020 (U.S. Department of Agriculture, 2021). After the data were obtained, they were input into spatial software QGIS to better visualize the spatial distribution of the data. The data were then uploaded into a data analysis package in Python called the Geospatial Data Abstraction Library (GDAL), which allowed the data to be sorted, manipulated, and graphed to analyze and better understand the patterns and trends.

Preliminary Results & Discussion

Preliminary results show a correlation between an increasing number of dairy cows and an increase in alfalfa acreage in Arizona (Figure 6). Alfalfa is a dairy cow's main food source. In 1983, the Dairy Production Stabilization Act was enacted by the United States federal government and authorized direct payments to farmers to reduce their milk production. The problem with this program was the government had strict regulations qualifying farmers for the program and direct payments were lower than dairy prices at the time. By the early 1990's most farmers decided not to participate. Figure 6 shows how dairy production increased significantly with the increase in dairy cow populations and the corresponding increase in alfalfa production.

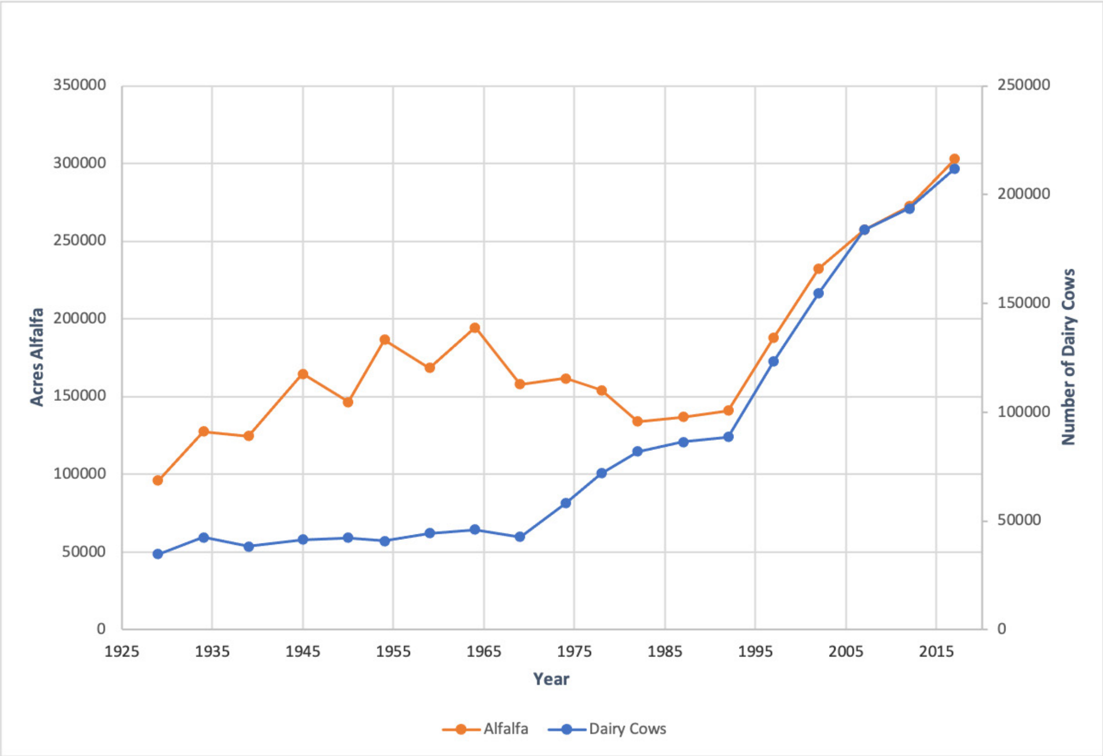


Figure 6: The dairy cow population and alfalfa acreage in Arizona from 1925-2017.

The earliest data on the spatial distribution of alfalfa in Arizona only dates back to 2008 in the CropScape database. Figure 7 is a map of southern Arizona that illustrates the increase in alfalfa acreage in the state. The new alfalfa acreage in 2020 is shown in light blue, while the acreage from previous years (2015, 2010, and 2008) is shown in increasingly darker shades of blue.

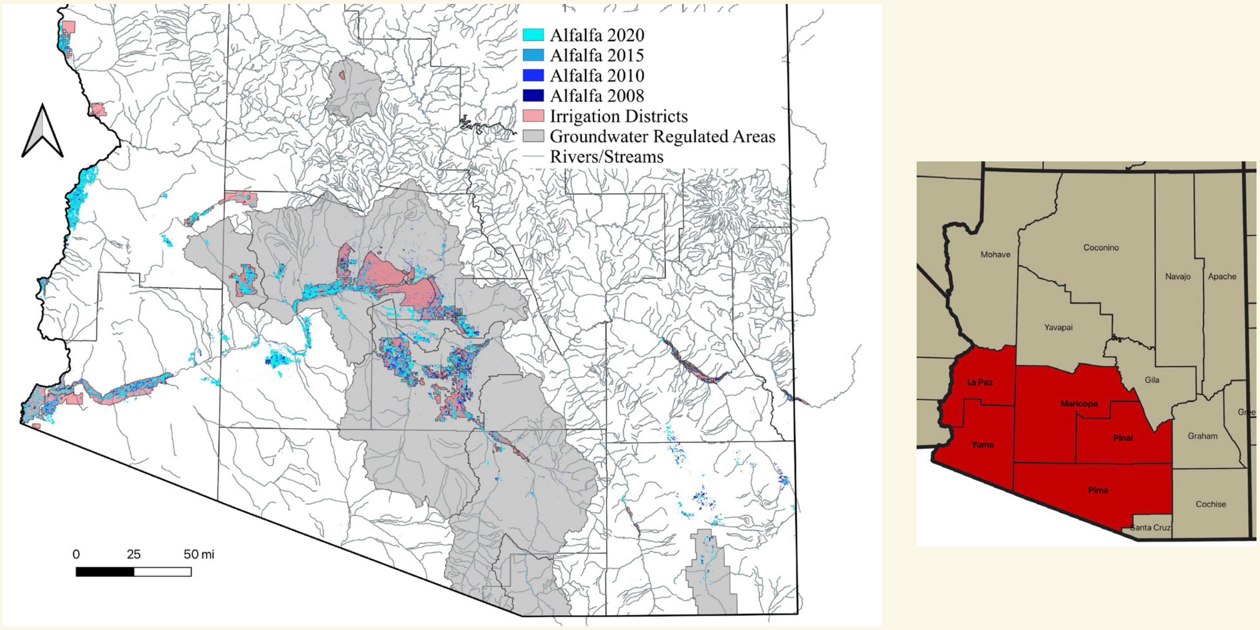


Figure 7: Spatial distribution of alfalfa in Arizona 2008-2020.

The largest increases in alfalfa acreage are evident in Maricopa County which has groundwater regulations (the gray shaded areas in Figure 7) in some but not all parts of the county. The large increase in alfalfa in this county most likely has been attributed to the fact that over 60% of the total state population lives in Maricopa County (United States Census Bureau, 2020). Dairy farms tend to be located relatively close to urban areas so that their products are trucked shorter distances to local markets, therefore reducing transportation costs to the farmers. Figure 7 also shows that alfalfa is grown within and outside irrigation districts.

Figure 8 shows the change in alfalfa acreage between 2008 and 2020. In the Figure 8, irrigation wells are shown as green dots, alfalfa acreage in 2020 is shown in light blue versus dark blue in 2008, and rivers shown in gray. This figure also shows the connection that irrigation wells are commonly found near natural waterways where shallow groundwater is. This shallow groundwater allows farmers to reduce pumping costs because pumps only have to lift groundwater a short distance to the surface. Alfalfa is commonly grown near streams and irrigation wells to provide sufficient irrigation water due to its high water requirements. Commonly near streams, the depth to water is usually significantly more shallow than in areas further from natural waterways. Streams can be one of two types; the first type is a losing stream which is more common in the desert southwest. In a losing stream, water infiltrates to the groundwater table, which causes a reduction in surface water flow.

A gaining stream is the opposite: the stream gains water from the groundwater table which thus feeds surface water flow. Alfalfa and other crops are commonly grown close to natural waterways where groundwater is shallow because this allows farmers to minimize pumping costs. It costs farmers less money to withdraw shallow groundwater versus deeper groundwater. Shallow groundwater wells are also much cheaper to drill compared with deep groundwater wells, since drillers tend to charge by the foot. Overall, growing alfalfa and other crops close to natural waterways can make it more economically friendly for farmers to withdraw groundwater.

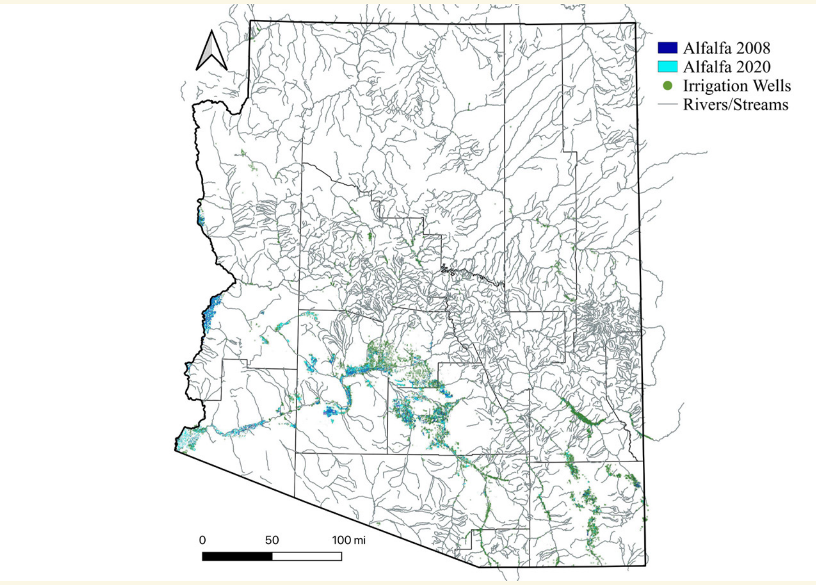


Figure 8: Spatial distribution of irrigation wells, rivers, and alfalfa locations in Arizona in 2020.

Depth to water from 1910-2020 is shown in Figure 9 for four groundwater subbasins: Gila Bend (GIL), East Salt River Valley (ESR), Eloy (ELO), and Maricopa-Stanfield (MST). These four subbasins were chosen because three of the subbasins (ESR, ELO, and MST) are located in areas where there are groundwater regulations, and the GIL subbasin provides an example of groundwater conditions where few if any regulations exist. An especially important characteristic of all four subbasins is that they collectively contain 157,530 acres of alfalfa which is 52% of Arizona's total. Since these irrigated alfalfa acres are too far from natural waterways to be irrigated by surface water their primary irrigation source must be groundwater, it can be concluded that increased alfalfa production in these areas has had a negative impact on the groundwater levels. Figure 9 shows that since 1980, groundwater levels have stayed steady or even had some recovery in the three regulated subbasins. The steady and/or recovering groundwater levels can be explained by the enactment of the Groundwater Management Act of 1980 in Arizona which created AMAs and INAs and charged the AMAs with the goal of safe yield by 2025.

Safe yield means the amount of groundwater being withdrawn is less than or equal to the amount of groundwater recharge to an aquifer. In the GIL subbasin where regulations are lacking, there is a continuing groundwater decline even after 1980. Figure 9 shows that in all four subbasins, groundwater levels have steadily declined from the early 1900s through today. Alfalfa grown in areas where groundwater is the primary irrigation source have had negative impacts on groundwater locally in those areas.

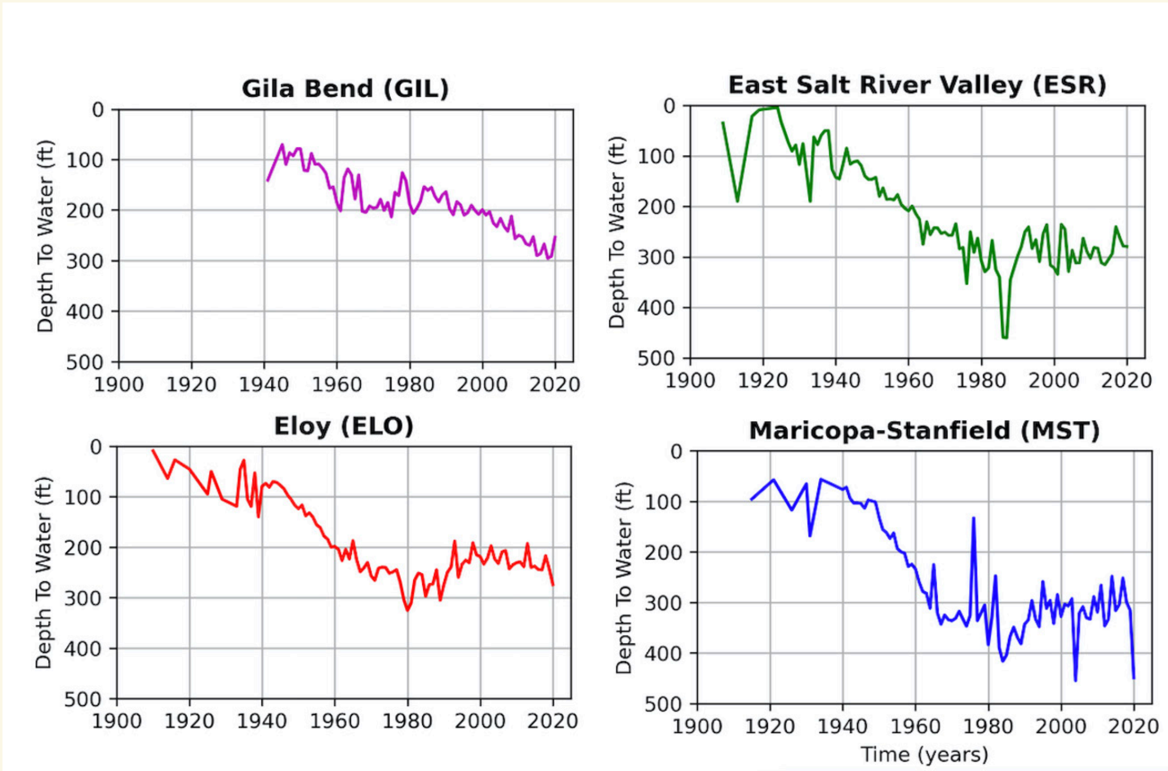


Figure 9: Time series of depth to water in four selected subbasins.

Future Research

The next step in this research is to investigate the connection between alfalfa expansion and new groundwater regulation in CA. In 2014, California signed the Sustainable Groundwater Management Act (SGMA), hoping to attain sustainable groundwater yield within the next 20 years. The goal of the SGMA is to stop groundwater overdraft in California, and reach balanced levels of pumping and recharge within the next 20 years of implementation (State Water Resources Control Board, 2020).

It is possible that after the introduction of SGMA in 2014, California farmers have moved to Arizona to take advantage of the lack of regulations outside the AMAs. I intend to explore spatial patterns in alfalfa expansion and groundwater declines relative to different groundwater regulation areas in AZ. Regulations differ slightly in AMAs compared to INAs and I will investigate how these differences in regulations affect groundwater declines. Third, I will explore discrepancies between the USDA Agricultural Statistics Service and CropScape values for alfalfa acreage (see Figure 10). CropScape estimates alfalfa acreage using Landsat imaging while the USDA Agricultural Statistics Service sequesters farmers records. For 2006-2012, CropScape estimated ~100,000 more acres of alfalfa than the USDA, and the difference has been increasing since ~2013. I hypothesize that the value generated by the USDA data is underestimated due to varying crop rotations involving alfalfa, which causes large amounts of under sampling. Lastly, I will analyze how dairy subsidies affect dairy in Arizona compared with data at the national level. A preliminary investigation (Murphree, 2018) shows that dairy cow populations have stayed constant nationally but increased in Arizona. This can be explained by the fact that milk yields per cow are higher because Arizona has more favorable environmental factors such as dry climate, abundant sunshine, and warm temperatures(Murphree, 2018). This could be an additional explanation for farmers moving to Arizona to obtain higher milk yields per cow rather than just the lack of groundwater regulations.

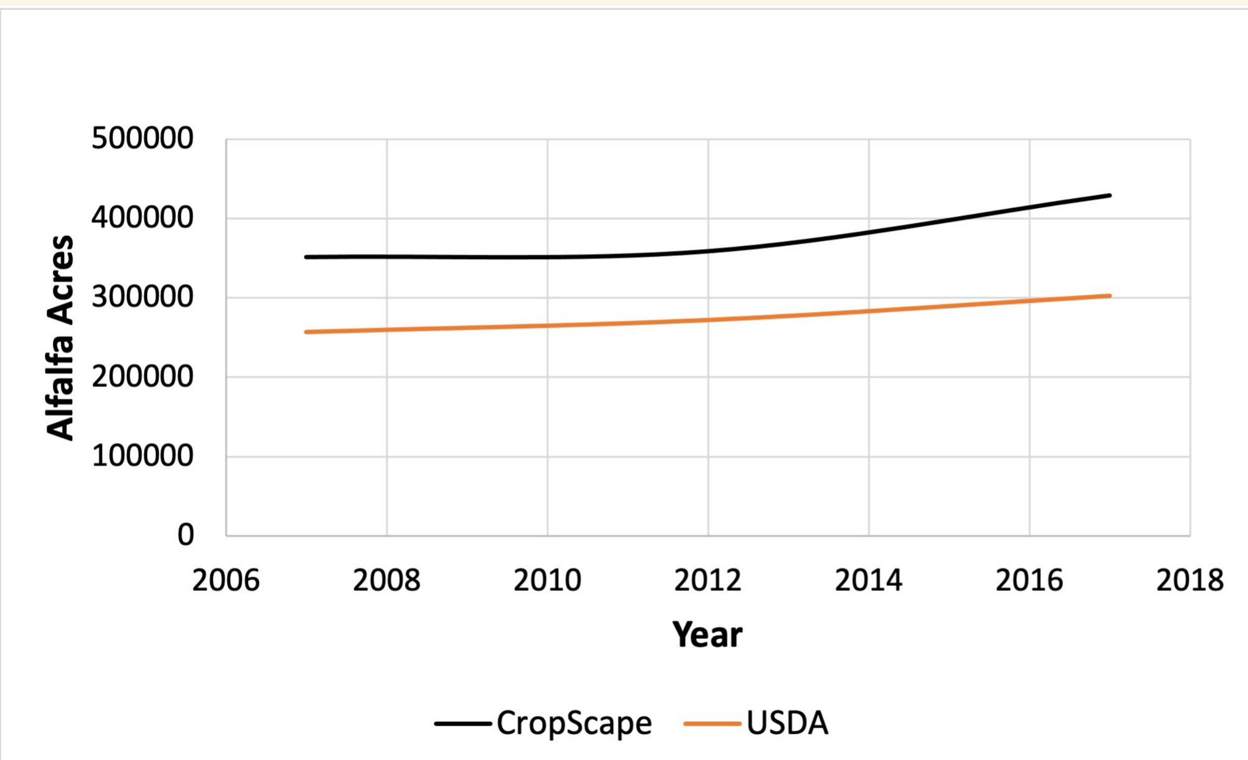


Figure 10: Alfalfa acreage CropScape vs. USDA Agricultural Statistics Service

References:

Arizona Department of Water Resources. (n.d.). Active Management Areas. Retrieved February 2, 2022, from AMA Annual Supply and Demand Dashboard:https://new.azwater.gov/ama/ama-data

Castle, S. L., Thomas, B. F., Reager, J. T., Rodell, M., Swenson, S. C., and Famiglietti, J. S. (2014), Groundwater depletion during drought threatens future water security of the Colorado River Basin, Geophys. Res. Lett., 41:5904– 5911.

Condon, L., & Maxwell, R. (2014). Groundwater-fed irrigation impacts spatially distributed temporal scaling behavior of the natural systen: a spatio-temporal framework for understanding water management impacts. Environmental Research Letters, 2014. 9(3): p. 009-034.

Dieter, C.A., Maupin, M.A., Caldwell, R.R., Harris, M.A., Ivahnenko, T.I., Lovelace, J.K., Barber, N.L., and Linsey, K.S. (2018) Estimated use of water in the United States in 2015: U.S. Geological Survey Circular 1441, 65 p.65

Freeze, A., & Cherry, J. A. (1979). Groundwater. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Harou, J., & Lund, J. (2008). Ending groundwater overdraft in hydrologic-economic systems. Hydrology Journal, 2008. 16: 1039-1055.

Hitaj, C., & Suttles, S. (2016, August). Trends in U.S. Agriculture's Consumption and Production of Energy: Renewable Power, Shale Energy, and Cellulosic Biomass. Economic Information Bulletin Number 159, pp. 1-53.

Howitt, R., Medellín-Azuara, J., MacEwan, D., Lund, J. R., & Sumner, D. (2014). Economic analsis of the 2014 drought for California agriculture. University of California, Davis, CA: Center for Watershed Sciences.

Murphree, J. (2018, June 5). The Most Interesting Facts about Arizona Dairies You'l Ever Read. Retrieved from Arizona Farm Bureau: https://www.azfb.org/Article/The-Most Interesting-Facts-about-Arizona-Dairies-Youll-Ever Read#:~:text=It%20varies%2C%20but%20about%2010,465%20%C2%BD%20gallons%20per%20month.

National Geographic Education Staff. (2014, September 2). National Geographic Resource Library. Retrieved from Staple Food Crops of the World: https://www.nationalgeographic.org/maps/wbt-staple-food-crops world/#:~:text=Food%20staples%20are%20eaten%20regularly,food%20crops%20around%20the%20world.

Nesheim, M., Oria, M., & Yih, P. (2015, Jun 17). A Framework for Assessing Effects of the Food System. Retrieved from NCBI: https://www.ncbi.nlm.nih.gov/books/NBK305168/

Scanlon, B. R., Faunt, C. C., Longuevergne, L., Reedy, R. C., Alley, W. M., McGuire, V. L., & McMahon, P. B. (2012). Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley. Proceedings of the national academy of sciences, 109(24), 9320-9325.

Siebert, S., Burke, J., Faures, J. M., Frenken, K., Hoogeveen, J., Döll, P., & Portmann, F. T. (2010). Groundwater use for irrigation–a global inventory. Hydrology and earth system sciences, 14(10): 1863-1880.

State Water Resources Control Board. (2020, June 10). Groundwater Management Program. Retrieved from California Water Boards: https://www.waterboards.ca.gov/water_issues/programs/gmp/

U.S. Department of Agriculture. (2014). Results from the 2013 Farm and Ranch Irrigation Survey. Census of Agriculture Highlights ACH, 12-16.

U.S. Department of Agriculture. (2021). CropScape - Cropland Data Layer. United States Department of Agriculture National Agricultural Statistics Service. https://nassgeodata.gmu.edu/CropScape/ Accessed May 5, 2021.

U.S. Department of Agriculture. (2019). Census of Agriculture. Retrieved from National Agricultural Statistics Service: https://www.nass.usda.gov/Publications/AgCensus/2017/Full_Report/Census_by_State/Arizona/index.php

U.S. Census Bureau. (2020). QuickFacts. Retrieved from 2019 Arizona Census : https://www.census.gov/quickfacts/maricopacountyarizona

Yuma County Agriculture Water Coalition. (2015). A Case Study in Efficiency – Agriculture and Water Use in the Yuma, Arizona Area. Arizona Department of Water Resources (ADWR).

ROBOT-ASSISTED SURGERY: AN ANALYSIS OF THE LEGAL AND ECONOMIC IMPLICATIONS

Helen Witharm

002

Abstract:

The legal and economic issues that arise as a result of including artificial surgery in the medical field will be analyzed. The questions of liability will be addressed by exploring how to define surgical robots in legal terms, including artificial agents and electronic personhood. The incorporation of additional parties in surgical procedures complicates attributing responsibility, and court cases involving such complications are discussed. In the past, failed surgeries were the fault of the surgeon and hospital and rarely any other parties. Now that surgeon-operated surgical robots are being used to assist in procedures, software engineers, manufacturers, and machine producers can also be held liable for any malpractice. Finally, the financial component of purchasing this technology is contrasted with the benefits and revenue achieved with robot-assisted surgery. A number of surgical robots, such as probes used in procedures like endoscopies and bronchoscopies, aren't used for incisions and are usually less expensive and entail less liability. Another selection of surgical robots are used as tools controlled by the surgeon for performing procedures; these machines can cost millions and face more legal liability. Robot-assisted surgery can cause higher efficiency, shorter recovery times, and minimal scarring, but it also raises legal and financial questions regarding the risk and feasibility of incorporating surgical robots into the medical field.

Keywords: **liability, economics, artificial intelligence, surgical robots**

Introduction

Technology has advanced greatly in the past twenty years and has gradually been integrated into medicine, making the process of testing, charting, and organizing surgery schedules and patient records smoother and more precise. As technology has grown in its ability, questions have arisen on incorporating artificial intelligence with surgery. The aspiration is that surgery can be made faster and less invasive which can lead to shorter hospital stays allowing medical care to be more accessible to all. The shortened hospital stays incur less expense on the patients, lowering a barrier to obtaining health care further. However, the perceived benefits of robot-assisted surgery are accompanied by several questions about the risks and liability of a robotic third-party in addition to the economic analysis of investing in the expensive machines.

Integrating robots with surgery complicates liability. Typically, in cases of malpractice, the only parties who face liability are the surgeon and the hospital. When artificial intelligence is involved many more parties can be held responsible in the case of surgical complications or fatality. Both the surgeon and hospital can be held liable in addition to the machine's manufacturer and software engineering company as well. The process of designing machines capable of surgery is extensive because it requires machines who depend on consistency and predictability to perform in an environment where predictability cannot be promised. If a complication occurs during surgery, it will require prompt action from the surgeon and may cause an increased risk of serious damage than if performed in a traditional manner.

The question of efficacy of robot-assisted surgery being feasible also depends on economic restraints. Due to the complex software, mechanical tools, and liability involved, the machinery utilized is very expensive and is not possible for many hospitals to afford. Furthermore, the ability to use these robots requires extensive training and therefore educational facilities will increasingly be expected to train upcoming surgeries in using the machines.

This puts a financial burden on hospitals and institutions to invest large sums of money towards procedures that will become more popularized and increase the number of facilities that perform robot-assisted surgery.

It is assumed that the ability to perform more surgeries in an amount of time will increase income, but it must also be assured that a completion of more procedures generates enough revenue to cover the cost of the machines.

There are still many developments to be made in the production of machinery and artificial intelligence for robot-assisted surgeries. These developments include making the machines more capable of handling complications that may arise during surgery to ease the worry of legal issues and lowering the cost for hospitals to adopt the new equipment. Through the continued pursuit of bettering surgical artificial intelligence and machinery, the risk of liability and financial burden can be lessened. Therefore, more hospitals are able to participate in this minimally invasive approach.

The Legal Implications of Robot-Assisted Surgery

The laws and regulations related to robot-assisted surgery are being discussed worldwide due to its complexity and novelty. An article written by Aída Ponce Del Castillo for the European Trade Union Institute (ETUI) details how the European Parliament has been approaching the subject. In the case of robot-assisted surgery, it becomes more complicated to apply laws and regulations due to machinery not having the same legal attributes as humans. To address this issue, it is important to clearly define robot-assisted surgery in legal terms as well as determine how to attribute personhood and liability. Through this process, the ETUI established a way to appropriately categorize the machines and hold them accountable in an increasingly technological surgical environment.

The complex computer software used in artificial intelligence creates greater barriers to understanding the actions of the machinery, examples of which occur in cases of computers interpreting MRI scans. These programs, created to interpret findings from scans and images are "modeled after the human brain, ... when given additional data, the neural network can modify its decision-making process for a more accurate response, without any explanation of how it has done so" (Sullivan & Schweichart, 2019). This situation is known as a "black box", referring to the lack of clarity in how conclusions are drawn by artificial intelligence (Sullivan & Schweichart, 2019).

This poses another legal issue because of the way liability in the medical practice currently stands. Producers of medical devices can be held liable for the damages caused by the device, but the issue becomes less clear when the producers of a machine are unaware of how the artificial intelligence will behave in practice.

To combat the issue of adapting laws to apply to artificial intelligence, it is important to clearly define what is meant when referring to robotic surgery or robot-assisted surgery. For the sake of legality, robot-assisted surgery cannot be referred to in broad terms. A largely accepted definition of robots is "an artificial device that can sense its environment and purposefully act on or in that environment; an embodied artificial intelligence; or a machine that can autonomously carry out useful work" (Winfield, 2012). This does not apply to robotic surgery. Instead, it is the "operation by telemanipulation of dexterous robotic tools through small incisions'' where the "surgeons sit at a console near the operating table and utilize joysticks to perform complex procedures" (Yang et al., 2020). The machines used in robot-assisted surgeries do not fall under the definition of robots due to their lack of autonomy, but they cannot be considered mere machines.

Ponce Del Castillo, a lawyer with a focus on regulatory issues of human genetics, goes on to explain how to properly define these machines by addressing their unique characteristics. She states that "the term 'artificial agent,' understood as a spectrum concept, makes sense because it covers a wide diversity of...decision-making algorithms, auto-mated machines, digital agents, hybrid multi-agents, Internet bots, robots, nano-robots, drones, etc" (Ponce Del Castillo, 2017). Since the machines are fully designed and created by humans they fall under the category of "artificial," and their ability to act as directed by the surgeon categorizes them as 'agents.' For this reason, when discussing robot-assisted surgery, it is in reference to surgeons performing procedures with the assistance of these artificial agents.

Most artificial agents are not considered autonomous due to their inability to act with intention or awareness of consequence, trying them legally as an electronic person is impossible. The machine itself cannot be prosecuted. In most cases, the surgeon has complete control over the artificial agent and therefore any malfunction is the responsibility of the software developer and manufacturer, not the machine (Kaplan, 2016).

A software developer of a surgical robot faces many difficulties such as mapping a large variety of individual cases, complicated surgeries causing a limited ability to operate, and the possibility of stains on the camera lens obstructing the surgeon's ability to navigate the procedure (Hu et al., 2020). Refining these possible liabilities can be achieved by further testing and thorough research. It is proposed that "real-time automated surgical video analysis could facilitate the objective and efficient assessment of surgical skills and provide feedback on surgical performance" in an article written by Yang et al. In addition, generating "timelines displaying tool usage during a surgery" can help with the data collection to make the artificial agents more capable of handling complex procedures. Familiarity with the way the machines operate, and the incorporation of data collected from live surgeries will benefit the performance of the artificial intelligence and ease the risk of legal action brought against producers of the machines.

Many lawsuits have been brought against producers of these surgical robots. A lawsuit involving a robot-assisted surgery was brought in 2016 against Intuitive Surgical on the grounds of a failed surgery performed with their da Vinci Robot (Kaplan, 2016). The same article goes on to explain that the plaintiff claimed that the machine was the cause of her postoperative infection, but the performing surgeon claimed that the fault was on her for ignoring post-operation instructions. In this case, Intuitive Surgical settled with the plaintiff, but other lawsuits had been brought against Intuitive, claiming they withheld knowledge of the machine's malfunctions (Kaplan, 2016). One attorney claimed that "in less than 1 percent of cases, patients or doctors reported that the insulation covering the 'wrists' of the robotic arms cracked, allowing electricity to shoot out and burn the patient" (Kaplan, 2016). When producers of surgical machines are involved in robot-assisted surgeries, they assume liability and face consequences for under-developed and refined machinery, as demonstrated in these allegations.

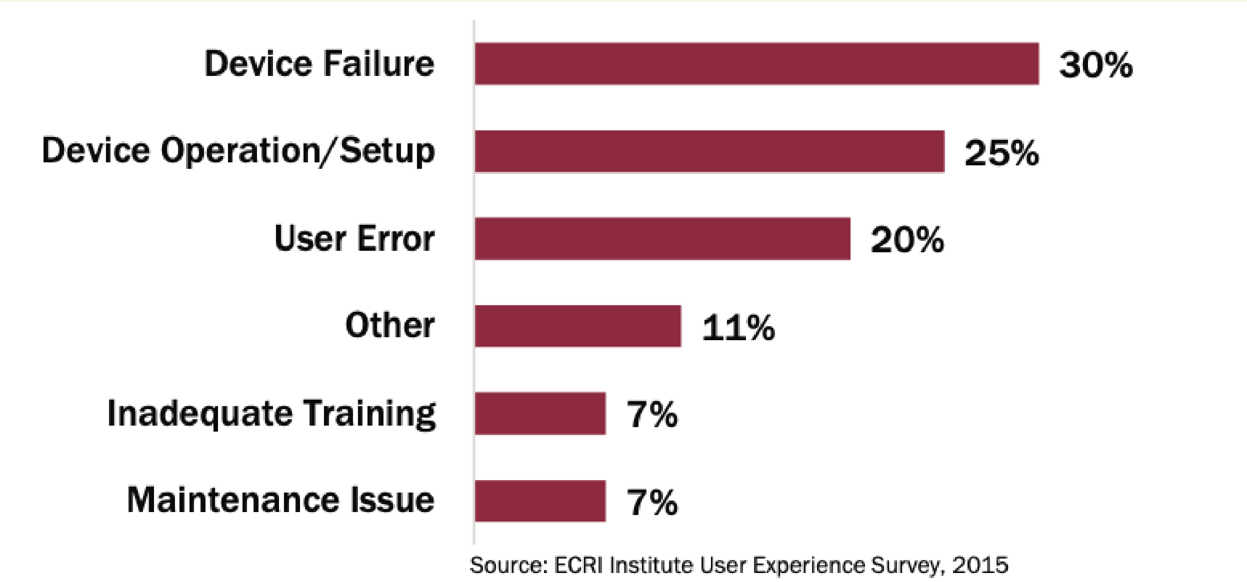


Fig. 1: ECRI Institute User Experience Survey, 2015

To determine who was most likely to blame for flawed performances by the da Vinci Robot, a survey was conducted by the ECRI Institute in 2015. Statistics gathered showed that over 50% of problems were due to the device's failure and flaws in operation and setup, while the remainder of issues were due to other causes including user error. This does indicate that when surgical robots are involved, device failure and liability is a considerable possibility. This nature of liability entails its own struggles as current law grapples with addressing artificial intelligence being incorporated with medicine.

The complex and diverse nature and usage of these surgical robots raise many questions about how liability ought to be applied. In the AMA Journal of Ethics, Hannah R. Sullivan and Scott J. Schweikart wrote that "it is challenging to find a responsible party, as so many different entities—software developers, hardware engineers, designers, and corporations—go into the creation of AI systems." (Sullivan & Schweichart, 2019) One solution proposed in this article is to "confer 'personhood' on the artificially intelligent machine itself, viewing the machine as an independent 'person' under the law." (Sullivan & Schweichart, 2019) This AI personhood wouldn't apply to machines like da Vinci where the surgeon has complete control over the device (Sullivan & Schweichart, 2019). For machines like da Vinci the responsibility if a mistake is made will either be attributed to the manufacturer of the machine or the surgeon/hospital in most cases. AI personhood would apply to programs made to interpret medical findings based on data collection. This example reveals the importance of distinguishing what type of machine is being referred to when speaking about surgical robots. There's a great variety of artificial intelligence, procedures, and parties involved, making liability a complex issue to tackle.

The Economic Implications of Robot-Assisted Surgery

Some artificial agents are already in use for surgeries and the data collected has indicated that despite the price tag, the number of surgeries performed with the help of artificial intelligence has increased dramatically. The most widely known machine used in robot-assisted surgery is the da Vinci Robot. The surgeon controls the arms of the da Vinci Robot while looking at a screen that projects a magnified view of the surgical field.

Laparoscopic surgery is well-known for its ability to decrease recovery time and scarring and da Vinci has become increasingly popular.

The da Vinci Robot is not a small investment. One machine alone costs upwards of two million dollars and each surgery costs anywhere from three to six thousand dollars for the patient (Scott, 2016). Over a thousand hospitals have adopted this device and the number of procedures done by the da Vinci Robot "rose from 1500 in the year 2000 to more than 20,000 in 2004" (Kumar R., 2005). While this increase may seem substantial, experts claim that "hospitals must perform anywhere from 150 to 310 procedures within six years to offset upfront and ongoing costs" (Scott, 2016). Unless developments are made to improve the performance of the da Vinci Robots, it is impractical for most hospitals to make this investment because although it does reduce scarring and recovery time, it requires many surgeries to be performed in order to be profitable. This problem is not only caused by a possible lack of surgical candidates, but the feasibility of providing enough hospital beds and equipment to support a large number of procedures.

Most procedures using the da Vinci Robot are surgeries that are more invasive, hence benefitting from the minimally invasive nature of robot-assisted surgery. Gynecology makes up 52% of all procedures performed with da Vinci (ECRI Institute, 2015), a field where minimal scarring is highly preferable. In a field, however, where procedures are already minimally invasive, there is very little incentive to invest in machines like da Vinci where procedure outcomes are not improved by its use. An example of this can be observed in the field of ophthalmology. Dr. Richard Lindstrom, an ophthalmologist in Minnesota, states that "eye surgery is already minimally invasive with very small incisions and nearly no blood loss. Also, visualization is usually excellent," but the final reason he shares for why the da Vinci Robot is undesirable in the field of ophthalmology is due to its cost (Lindstrom, 2021). Adopting this machine is not justified when "robotic systems are expensive, have a steep surgeon learning curve and usually require more time per procedure" (Lindstrom, 2021). He continues to explain how this also becomes unreasonable for educational facilities to afford because to master the use of these machines a student must perform between 150 to 250 procedures (Lindstrom, 2021).

There are several surgical robots on the market that are economically feasible, and the Monarch by Auris is among those. The Monarch is a robotic bronchoscope that is controlled by the surgeon with a controller that closely resembles that of an Xbox or PlayStation. This machine, unlike the da Vinci, isn't meant for surgical procedures that require incisions or resections.

The Monarch is designed to perform Bronchoscopies, an exploratory procedure intended to identify cancerous nodules in the lungs. While this procedure doesn't remove the nodules, it does have the ability to "detect malignant tumors at an earlier stage, therefore significantly increasing lung cancer survival rates'' (C.F. Graetzel et al., 2019). This surgical robot costs about $500,000 (Densford, 2018), which is significantly less than the da Vinci Robot. Bronchoscopies usually cost about the same as laparoscopic surgeries but are performed over 500,000 times a year, making the cost of the Monarch much easier for hospitals to recover from (Guglik, 2018). Robotic scopes like the Monarch incur less liability due to their benign nature and are typically less expensive than machines like the da Vinci.

To some hospitals, though, the fact that the cost of a surgical robot may not be recovered isn't an issue. A study conducted shows that in California, the profit margin for hospitals averaged 8%, which from the years 2013 to 2017 resulted in over 5 billion dollars in revenue being earned each year (Belk, 2021). Not all hospitals have profit margins of this extent, but this does illustrate that for some hospitals the attention brought by integrating a surgical robot into their facility can be worth the cost of machines like da Vinci. Additionally, having surgical robots allows for opportunity to facilitate education: another form of revenue. Less invasive surgery and shorter hospital stays are attractive to patients because they can save money on hospital bills and reduce the necessary amount of time off from work. As robot-assisted surgery becomes increasingly popular, hospitals may become more willing to accept the high cost of expensive machinery with the hope that the attention and opportunity it offers make up for the disparity.

Conclusion

Technology, as it's grown in scope and capability, has proven greatly beneficial in the medical field. Surgical robots have been integrated into many surgical programs and have resulted in less invasive surgery and shorter recovery time. However, it has created new obstacles of liability as there are now several more parties involved in the case of malpractice. Software developers and manufacturers are now involved in a field that previously only had two primarily responsible parties: the surgeon and the hospital.

In cases of highly advanced artificial intelligence, tort laws have been prompted to add terminology to address electronic personhood and create precise definitions to address whether a machine used in surgery is a mere tool or an artificial agent. Furthermore, economics impact the ability for hospitals and institutions to integrate surgical robots in the operating room. Robot-assisted surgery includes much more expensive machinery as well as hundreds of practice procedures to reach mastery. While surgical robots do assist with recovery time and minimal scarring, hospitals must weigh whether the publicity and financial benefits are worth the millions of dollars required to purchase these machines. Robot-assisted surgery is growing in popularity, but as technology advances rapidly, legal and economic discussions are being had regarding the feasibility of integrating these machines.

References:

Belk, D. (2022, January 17). Hospital Financial Analysis. True Cost of Healthcare. Retrieved March 6, 2022, from https://truecostofhealthcare.org/hospital_financial_analysis/

Densford, F., & Redick, M. (2018, November 28). Auris health raises $220M for medical robotics platform. The Robot Report. Retrieved March 6, 2022, from https://www.therobotreport.com/auris-health-220m-monarch-medical-robot/

ECRI Institute. (2015). Robotic Surgery. Robotic Surgery Infographic. Retrieved March 6, 2022, from https://assets.ecri.org/PDF/Infographics/Robotic-Surgery-Infographic.pdf

Graetzel, C. F., Sheehy, A., & Noonan, D. P. (2019). Robotic Bronchoscopy Drive mode of the Auris Monarch Platform*. 2019 International Conference on Robotics and Automation (ICRA). https://doi.org/10.1109/icra.2019.8793704

Guglik, C. (2018, January 18). Bronchoscopy's hidden issue . Bedside Bronchoscopy. Retrieved March 6, 2022, from https://www.bedsidebronchoscopy.com/bronchoscopys-hidden-issue/

Kaplan, T. (2018, May 18). Robot surgery: Intuitive surgical settles litigation over robotic arms. Oroville Mercury-Register. Retrieved March 6, 2022, from https://www.orovillemr.com/2016/04/20/robot-surgery-intuitive-surgical-settles-litition-over-robotic-arms/

Kumar, R., & Hemal, A. K. (2005, October). Emerging role of robotics in Urology. Journal of minimal access surgery. Retrieved March 6, 2022, from https://www.ncbi.nlm.nih.gov/pubmed/21206664

Lindstrom, R. (2021). Ophthalmic robotic surgery may be most applicable in retina. Healio. Retrieved March 6, 2022, from https://www.healio.com/news/ophthalmology/20210412/ophthalmic-robotic-sugery-may-be-most-applicable-in-retina

Ponce del Castillo, A. (2017). A law on Robotics and artificial intelligence in the EU? SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3180004

Scott, C. (2016, August 10). Is da vinci robotic surgery a revolution or a ripoff? Healthline. Retrieved March 6, 2022, from https://www.healthline.com/health-news/is-da-vinci-robotic-surgery-revolution-or-ripoff-021215#Dude,-wheres-my-robot?

Sullivan, H. R., & Schweikart, S. J. (2019). Are current tort liability doctrines adequate for addressing injury caused by Ai? AMA Journal of Ethics, 21(2). https://doi.org/10.1001/amajethics.2019.160

Winfield, A. (2012). Robotics: A very short introduction. Oxford University Press.

Yang, C., Zhao, Z., & Hu, S. (2020). Image-based laparoscopic tool detection and tracking using Convolutional Neural Networks: A review of the literature. Computer Assisted Surgery, 25(1), 15–28. https://doi.org/10.1080/24699322.2020.1801842

THE HITCHHIKER'S GUIDE TO MENTAL CONTRASTING: EXPLORING THE INTERSECTION OF MINDFULNESS & MENTAL TIME TRAVEL

Kate Chambers

003

Abstract:

The adaptivity of mindfulness as a personality trait and the beneficial impact of mindfulness interventions on mental health are both substantiated by research. However, the phenomenon of mindfulness as a quality of consciousness remains ambiguous. This is especially true in regards to its relationship with time. Mindfulness emphasizes intentionality; has been proposed to represent healthy time perspectivity; and is correlated with goal conceptualization, pursuit, and achievement. As a preliminary measure in closing this research disparity by probing mindfulness and future-oriented cognition, the goal conceptualization strategy mental contrasting emerged as an intuitive addition to the body of research. This strategy professes three tenets: goal inception, obstacle generation, and solution cultivation. As present-minded action is accentuated, mindfulness is anticipated to be compatible with this model. The Mindful Attention Awareness Scale (MAAS) assessed subjects (n=94) instructed with Think Aloud protocol during which they vocalized organically-emerging thoughts for ten-minute recorded intervals. A qualitative scoring analysis was then applied so as to discern the proportion of the goal conceptualization strategy mental contrasting within raw test subject transcripts.

Linear regression models demonstrated a positive correlation ($p < 0.001$) between trait mindfulness and the utilization of the full three tenets of mental contrasting which emphasizes overcoming obstacles in anticipation of goals. Fantasizing, a fixation on goals without emphasizing obstacles, possessed an inverse relationship with trait mindfulness (p-value: 0.012), while failing to generate solutions to goal obstacles was also negatively correlated with trait mindfulness ($p < 0.001$). Mindfulness as an intervention may be calibrated to encompass such adaptive temporal implementation.

The Adaptivity of Mindfulness and Time Perspective Orientation

Though empirically embraced and clinically purported as intervention (Keng, 2011), secular mindfulness has been critiqued for its disproportionate emphasis on the present while neglecting its application to alternate tenses (Dreyfus, 2011). Given the extensive bedrock of literature pertaining to the adaptive sensorial correlates of trait mindfulness such as enhanced visual perception (Kharlas, 2016), superior motor acuity (Tarrasch-Lilach, 2017), interoception capabilities (Gibson, 2012), and both superior attentional awareness and low sensory sensitivity (Hebert, 2016), it is logical to anticipate that mindfulness may correspond with an adaptive relationship with sense of time. Corroborating this assertion is the proposition that mindfulness exists as a healthy time perspective (Zimbardo, 2014). Existing time perspectivity research suggests mindfulness is correlated with contemplating future ramifications of one's current actions. (Seema et al., 2014). Mindfulness and present-mindedness additionally correspond with self-directedness (Smalley, 2009), goal generation (Sobol-Kwapinska, 2009), pursuit (Brown, 2007) and subsequent achievement (Sobol-Kwapinska, 2009). Trait mindfulness is additionally associated with authenticity of goals characterized by their self-concordant properties which align with the identity of the goal-setter (Sheldon et al., 1999). Such goal authenticity is adaptive in that it emphasizes internal locus of control and is furthermore correlated with goal attainment (Koestner, 2002). Trait mindfulness also predicts goal internalization over the duration of their pursuit (Smyth, 2020). The bulk of research therefore suggests a positive relationship between goal quality and mindfulness. As such, we deemed it logical to examine intentional goal-related cognition's relationship with trait mindfulness.

Effective goals are additionally characterized by concreteness and by the proximal, more immediately attainable nature of subgoals (Cochran, 1996). In this context, this may be interpreted as a connection between future aspiration and the present-moment actions necessary in acquiring said aspiration. The goal conceptualization strategy mental contrasting (Oettingen, 2012) emphasizes actionable, present-oriented solutions designed to circumvent obstacles which threaten the achievement of a goal. This three-tenet framework was found to positively correspond with the achievement of goals (Kappes, 2021). Mental contrasting itself is characterized by first richly envisioning a desired future. This is followed by considering obstacles which impede the achievement of this reality. Mental contrasting then ultimately reconciles these obstacles by compiling contingency plans and exploring avenues by which the obstacles may be overcome so that the future in question may become reality. This method synthesized the present and future by combining the intentionality associated with mindfulness (Rothaupt & Morgan, 2007) with future-oriented ambition applied to actionable intention within the present tense. This structure is reminiscent of mindfulness' moment-minded nature. It would therefore be logical to expect a positive correlation between trait mindfulness and mental contrasting.

It is important to consider that merely fantasizing about a longed-for alternative timeline does not constitute mental contrasting or effective goal conceptualization. Mental contrasting was developed as a means of circumventing such indulgent cognition (Oettingen et al., 2010). Remaining entrenched in fantasizing boasts a host of undesirable correlates such as poor problem-solving capability, life dissatisfaction and low rates of behavioral transformation (Oettingen, 2013). It was therefore anticipated that such dissatisfied individuals who engaged in high rates of fantasizing would possess lower trait mindfulness, which we postulate will manifest in adaptive proclivities in regard to time perspectivity and goal generation. This adaptivity would specifically manifest as structured orientation towards future-oriented cognition in the form of mental contrasting.

Other individuals are distinct in that they succeed in progressing within the tenets of mental contrasting from fantasizing to generating obstacles which impede the procurement of their desired future. However, in response to such barriers, while some advanced to the third and final tenet of cultivating solutions, others ruminated on the obstacles without such progression. Such subjects' transcripts were characterized by discouragement, anxiety, and complacency. These non-adaptive problem-solving orientations are in direct contrast to the adaptive orientations mindful individuals displayed in the literature regarding goal conceptualization (Sobol-Kwapinska, 2009). It was therefore hypothesized that these individuals would emerge possessing low levels of trait mindfulness within the analysis.

In terms of individuals who did not engage within goal-directed cognition, mental contrasted or otherwise, a noteworthy relationship with mindfulness was not anticipated. This is because we are interested in probing ways in which mindful individuals engage with goal orientation only when they deem it necessary to do so. We are not probing goal emphasis itself, and subjects were not instructed to engage within goal-oriented cognition within this task.

Attention-Deficit/Hyperactivity Disorder (ADHD) is inversely correlated with trait mindfulness (Smalley, 2009). As we expect mindfulness to possess a positive relationship with a utilization of the full three tenets of mental contrasting, it was therefore logical to expect an inverse relationship between ADHD and this goal conceptualization strategy. As for fantasizing and the failure to generate solutions to emerging obstacles, a positive relationship between these and ADHD were deemed rational hypotheses. People with ADHD are unlikely to endorse performance-approach goal orientation (Barron, 2006). Having ADHD is also linked with a chronic sense of underachievement, feeling one should be farther along in life than one is, and a history of not living up to potential in school or work (Firmin, 2007), which further bolsters this assertion.

Finally, depression corresponds with the subjective perception of a slow passage of time (Zimbardo, 2008) and a fixation on past-oriented cognition (Hawkins, 1988). As past- orientation and goal conceptualization are mutually exclusive, it was anticipated that cognition colored by depression would inversely correlate with mental contrasting. In terms of fantasizing, due to the negative life outcomes associated with this future-oriented stance, a positive correlation between these two traits was hypothesized. As for the relationship between depression and a failure to expound upon obstacles with solutions, a positive correlation was expected due to emphasis on futility. A relationship of note was not expected between depression and cognition devoid of mental-contrasting.

Methods

Subjects (n = 78, 45 women, 30 men, 1 non-binary, 2 declined to report; range = 18-28 years) recruited from the University of Arizona's undergraduate student body were assessed utilizing a battery of questionnaires including the Mindful Attention Awareness Scale (MAAS), a metric designed to measure an individual's attunement to the present moment across everyday experience (Park, 2013).

The seven-item Generalized Anxiety Disorder self-report screening tool (GAD-7) was used to probe levels of anxiety within the subjects' previous two weeks of overall mental health (Williams, 2014). Levels of Attention Deficit / Hyperactive Disorder were measured with the Adult ADHD Self-Report Scale (ASRS-v1.1), a six-item instrument which assesses both inattention and hyperactivity (Hesse, 2012) . The Patient Health Questionnaire (PHQ-9) (Kroenke & Spitzer, 2002) assessed for depression severity. Subjects were then instructed to voice aloud all organically-emerging cognition unrelated to any task, a "linguistic fingerprinting" procedure known as Think Aloud (Duncker, 1945; Pennebaker, 1999). During this protocol, subjects were given the instruction to freely vocalize emergent thoughts with the sole prompt of "Now think" on their computer screen guiding them. Such thought content could include depiction of mental imagery, narration of external stimuli, and attunement to bodily sensation. Self-censorship was discouraged. A preliminary one-minute Think Aloud sample was collected in the presence of a researcher in order to ascertain the subjects' ability to vocalize cognition in tandem with the established experimental parameters. After participants demonstrated an ability to engage in Think Aloud, a subsequent ten-minute sample was recorded with the subject in isolation from third parties. The 16 participants who did not comply with the Think Aloud protocol were excluded from the final analysis.

An experimental testing room in the Psychology Building at the University of Arizona was used for all Think Aloud procedures. Participants were instructed to put phones away for the duration of the study so as to remove competing stimuli and reinforce the lack of task-orientation of the cognition. Upon completion of Think Aloud, participants then used a gradient scale to indicate how similar the vocalized thoughts were as compared to ones experienced in their day-to-day life.

The Think Aloud audio files were then transcribed and delineated into distinct thoughts per rigorous protocol for the sake of expedient coding and subsequent analysis. In accordance with this methodology, three raters delineated the transcripts into strong and associational transitions between thoughts. The interrater reliability in terms of average thought length and thought quantity demonstrated consistency (Cronbach alpha = .97, CI95=[.95; .99] and Cronbach alpha = .86, CI95=[.81; .91] respectively) with a 95% confidence interval in all contexts with the exclusion of temporal and perceptual orientation.

A qualitative codebook analysis (Fig. 1) was developed to assess the proportion of thoughts containing the goal conceptualization strategy mental contrasting within the raw Think Aloud transcripts.

This protocol was developed to mirror the structure of mental-contrasting (Oettingen, 2013) with emphasis on momentum of thought within goal conceptualization in conjunction with goal actionability. The paradigm was refined to emphasize present-moment actionability of goals within distinct thoughts. After transcripts were delineated into distinct thoughts per strong and associational transitions, a numeric value corresponding to a tenet of mental-contrasting was then ascribed to each thought in its entirety. This analysis allowed for assessing both proportion of distinct thoughts and for the proportion of thoughts as defined by respective word count. Linear regression analyses were then performed in conjunction with data gleaned from this protocol and the results from the battery of questionnaires associated with respective participants.

Linear regression models revealed a positive correlation (Fig. 9. $p < 0.001$, r: 0.509) between trait mindfulness and the utilization of the full spectrum of mental contrasting within the transcripts. Fantasizing possessed an inverse relationship with trait mindfulness (Fig. 7. p-value: 0.012, r: -0.289). Failing to generate solutions to goal obstacles was also inversely correlated with trait mindfulness (Fig. 8. $p < 0.001$, r: -0.467). No significant relationship was found between mindfulness and non-engagement with mental contrasting tenets. All these correlations were true in terms of mindfulness and its relationship to the numbers of thoughts indicated by the protocol, the ratio of thoughts containing the mental contrasting tenets as compared with the total number of thoughts, and in terms of content ratio indicated by word count. Depression as measured by the (PHQ)-9 possessed statistically insignificant inverse relationships with both the full spectrum of mental contrasting and with cognition entirely unrelated to goals. Slight positive relationships emerged between depression and both fantasizing and the failure to generate solutions to obstacles.

0	-**No** mention of a goal or future one wishes to experience by means of their own action. -**No** assertion of intentionality to engage in goal-related behavior.
1	+ A goal the subject would like to achieve. (e.g. "I need to start tanning more") without mentioning potential obstacles or steps to implement towards reaching that goal. + The goal may be desirable due to its pleasant qualities, positively valenced in nature. (e.g., "I want to go to the beach") + The goal may be desirable due to its contrast with a less preferred outcome (avoidant). (e.g., "I never want to forget my towel at the beach again") This does **not** apply to: • A future-oriented thought in which an individual describes a future which is outside their control. (i.e. the weather or the actions of another independent entity) Or • A statement of fact pertaining to a future event which will occur whether the individual acts or not (e.g., "The football game is this Saturday"; "My parents are visiting this weekend")
2	+ A thought which expresses one or more obstacles which must be overcome in order to either achieve a specified preferred future or to avoid the nonpreferred future. ("I want to go to the beach, but I don't have any more vacation days for a trip") + These obstacles may entail barriers, anxieties, limits, negative emotional states or personality traits of the self, or any third party described which obstructs attainment of the goal in question. + It is common for MC2 thoughts to end in tentative statements such as: "But it's whatever", "But I'm just not going to worry about it", "But yeah…", "So…" "I don't know", "Guess I just have to get over it", "Nothing we can really do about it" … As if the subject is trailing off before completing the entirety of their thought. *Please note that if a solution is mentioned tentatively, contradicted, or immediately negated by the generation of additional incompatible solutions, this is still to be labeled MC2. This is because the focal point of the thought is on obstacles and not in overcoming.
3	+ Specific solution(s) which may be employed in overcoming an obstacle(s) so that the specified future may be experienced. The cognitive momentum culminates in a solution which the speaker endorses without contradiction. ("I want to go to the beach, but I don't have any more vacation days for a trip. If I plan something over Labor Day weekend, that should work") + Within MC3 thoughts, obstacles may be implicit. If an individual is depicting actions they intend to engage in so as to pursue a goal, the obstacles are unspoken but evidently present in the speaker's mind. If no obstacle exists, there is no need to plan.

Figure 1:Mental Contrasting Coding Protocol

Results

Linear regression models revealed a positive correlation (p <0.001) between trait mindfulness and the utilization of the full spectrum of mental-contrasting within the transcripts. Fantasizing possessed an inverse relationship with trait mindfulness (p-value: 0.012). Failing to generate solutions to goal obstacles was also inversely correlated with trait mindfulness (p <0.001). No significant relationship was found between mindfulness and non-engagement with mental-contrasting tenets. All these correlations were true in terms of mindfulness and its relationship to the numbers of thoughts indicated by the protocol, the ratio of thoughts containing the mental contrasting tenets as compared with the total number of thoughts, and in terms of content ratio indicated by word count. Depression as measured by the (PHQ)-9 possessed statistically insignificant inverse relationships with both the full spectrum of mental-contrasting and with cognition entirely unrelated to goals. Slight positive relationships emerged between depression and both fantasizing and the failure to generate solutions to obstacles.

		MAAS (n = 75)		MAAS (- SS with no goals, n = 70)	
		r	p	r	p
Thought number	MC0	-0.006	0.957	-0.006	0.959
	MC1	-0.289	0.012	-0.309	0.009
	MC2	-0.467	<.001	-0.483	<.001
	MC3	0.509	<.001	0.523	<.001
Thought ratio	MC0	0.091	0.44	0.105	0.386
	MC1	-0.281	0.015	-0.305	0.01
	MC2	-0.486	<.001	-0.501	<.001
	MC3	0.484	<.001	0.495	<.001
Content ratio	MC0	0.212	0.085	0.242	0.058
	MC1	-0.345	0.004	-0.365	0.004
	MC2	-0.547	<.001	-0.563	<.001
	MC3	0.519	<.001	0.535	<.001

Figure 2: Mindfulness & Mental Contrasting Levels

		PHQ9 (n = 75)		PHQ9 (- SS with no goals, n = 70)	
		r	p	r	p
Thought number	MC0	-0.129	0.27	-0.117	0.333
	MC1	0.078	0.506	0.136	0.261
	MC2	0.179	0.125	0.218	0.07
	MC3	-0.337	0.003	-0.317	0.008
Thought ratio	MC0	-0.041	0.728	-0.12	0.321
	MC1	0.124	0.291	0.193	0.11
	MC2	.124 (.308)	.065 (.008)	.250 (.346)	.037 (.004)
	MC3	-0.212	0.068	-0.19	0.116
Content ratio	MC0	-0.046	0.711	-0.148	0.251
	MC1	0.154	0.215	0.219	0.088
	MC2	0.204	0.099	0.243	0.057
	MC3	-0.261	0.033	-0.236	0.064

Figure 3: Depression & Mental Contrasting Levels

		ADHD (n = 75)		ADHD (- SS with no goals, n = 70)	
		r	p	r	p
Thought number	MC0	0.082	0.485	0.094	0.438
	MC1	0.14	0.229	0.217	0.071
	MC2	0.252	0.029	0.303	0.011
	MC3	-0.464	<.001	-0.442	<.001
Thought ratio	MC0	0.075	0.523	-0.01	0.933
	MC1	0.091	0.439	0.172	0.154
	MC2	0.281	0.015	0.326	0.006
	MC3	-0.402	<.001	-0.386	0.001
Content ratio	MC0	-0.062	0.621	-0.193	0.132
	MC1	0.265	0.03	0.355	0.005
	MC2	0.345	0.004	0.402	0.001
	MC3	-0.469	<.001	-0.447	<.001

Figure 4: ADHD & Mental Contrasting Levels

		GAD (n = 75)		GAD (- SS with no goals, n = 70)	
		r	p	r	p
Thought number	MC0	-0.003	0.976	0.016	0.389
	MC1	0.129	0.27	0.213	0.076
	MC2	0.255	0.027	0.315	0.008
	MC3	-0.317	0.006	-0.292	0.011
Thought ratio	MC0	-0.013	0.913	-0.117	0.334
	MC1	0.137	0.242	0.232	0.053
	MC2	0.209 (.292)	.072 (.012)	.260 (0.426)	.030 (<.001)
	MC3	-.264 (-.361)	.022 (.002)	-.239 (-.341)	.046 (.004)
Content ratio	MC0	-0.055	0.658	-0.194	0.131
	MC1	0.191	0.122	0.283	0.026
	MC2	.224 (.308)	.069 (.012)	.282 (.372)	.027 (.003)
	MC3	-0.302	0.013	-0.275	0.031

Figure 5: Anxiety & Mental Contrasting Levels

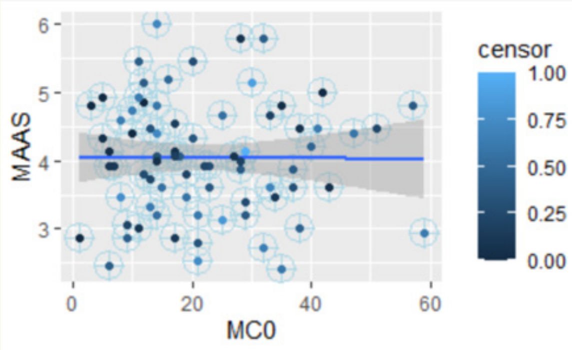


Figure 6: Trait Mindfulness & No Mental Contrasting

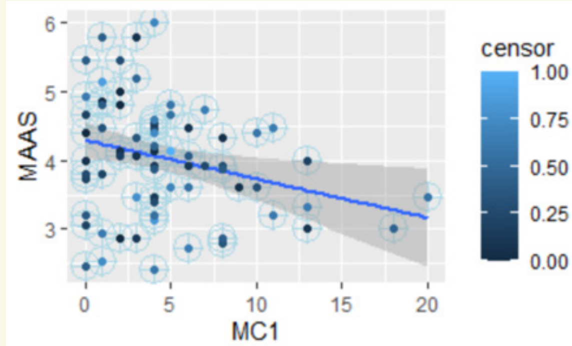


Figure 7: Trait Mindfulness & Fantasizing

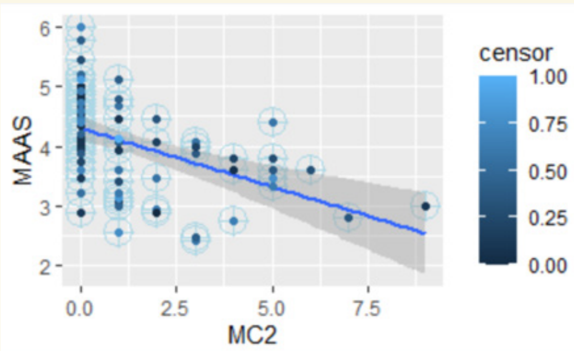


Figure 8: Trait Mindfulness & Obstacle Generation

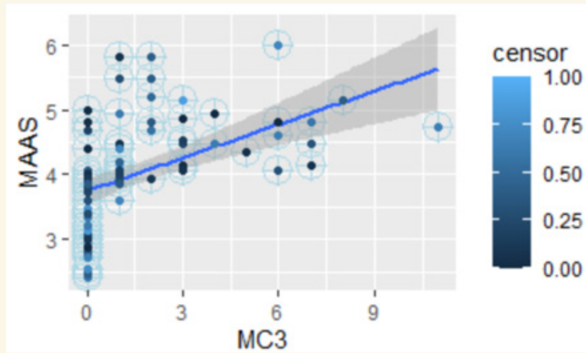


Figure 9: Trait Mindfulness & Mental Contrasting

ADHD had an inverse relationship with mental-contrasting ($p < 0.001$, r-value: - 0.464); whereas positive relationships emerged in relation to ADHD and the failure to expound upon obstacles by listing solutions (p-value: 0.029, r-value: 0.303). Specifically in terms of proportion of thought content, ADHD also possessed an inverse relationship with fantasizing (p-value: 0.03, r-value: 0.265). Depression as measured by the PHQ-9 indicated an inverse relationship with mental-contrasting (p-value: 0.03, r-value: -0.337). Anxiety possessed an inverse relationship with mental contrasting (p-value: 0.006, r-value: -0.317) and a positive relationship with obstacle fixation without generating solutions (p-value: 0.027, r-value: 0.255). When a lack of goals was controlled for, anxiety was additionally found to correspond with fantasizing (p-value: 0.076, r-value: 0.213).

Discussion

These relationships provide a bedrock for future research pertaining to the correlates of mindfulness as it relates to time perspectivity and goal generation, refinement, pursuit, and ultimate achievement. Within this task, subjects abundant in trait mindfulness implemented the goal conceptualization strategy of mental contrasting in greater proportion as compared with their less mindful counterparts. These results suggest that not only are mindful individuals more preoccupied with the present as corroborated by the body of literature, but there is an emphasis on curating feasible present moment action in regards to its future ramifications and alignment with a goal. Within this study, both the failure to generate solutions to obstacles impeding goals and a fixation upon the future without accounting for prerequisite actions ushering in such goals were correlated with depression and ADHD.

A central limitation of this study pertains to the MAAS questionnaire being characterized by a focus on attentional awareness and not soliciting self-report pertaining to the unconditional acceptance component of trait mindfulness (Trousselard et al., 2010).

While we have probed the attentional awareness component of mindfulness distinct from self-acceptance, future research may be benefitted by emphasizing self-acceptance or even by isolating it in relation to goal conceptualization. For instance, the authentic nature of mindful goals (Sheldon, 1999) may be specifically attributed to this unconditional acceptance trait as opposed to mindfulness as a collective. Scaffolding upon this study's results, future endeavors may wish to assess how mindful individuals create goals in both spontaneous, unstructured contexts and in prompted ones. This would help us ascertain delineations or shared patterns mindful individuals may profess in goal generation. Study designs curated to assess adaptive relationships or strategies other than goal conceptualization employed by mindful individuals when relating to future-oriented cognition would be additionally desirable. Another intuitive addition to the body of research would be assessing the manner in which mindful individuals relate to the past. Though nonadaptive rumination is associated between the mentally unwell and past-oriented cognition (Raffaelli, 2021), little is known of adaptive means of relating to the past and how that relates to trait mindfulness specifically.

It may also be beneficial to pivot from trait mindfulness to mindfulness as a conscious state by probing the impact of mindfulness meditation or intervention upon goal strategies employed, their quality, and their successful execution.

In response to these findings within an implementation context, mindfulness as an intervention may be calibrated to encompass adaptive temporal leanings. This will aid clinically-guided goal pursuit and the training adaptive future-oriented cognition. Mental contrasting through a lens of mindfulness may serve as an adaptive alternative to anxiety-laden catastrophizing, and framing it as such within a therapeutic context may increase overall life satisfaction. This may be of particular benefit to individuals struggling with psychopathology, namely those afflicted with depression or ADHD, which were unveiled in our study as inversely associated with mental contrasting and with mindfulness itself. Within the experiment, such suffering individuals were inclined to ruminate upon obstacles without generating solutions, and intentionally crafting therapies to circumvent cognition could foster emotional fortitude.

Should we desire to gain a comprehensive view of mindfulness as a quality of consciousness, it is necessary to view it through a temporal prism. By isolating the influence this trait exerts upon time perception, we gain invaluable insight into individual differences and novel means of operating. This may beget healthy strategies for operating within our complex world and subsequent interventions designed to foster such healthy strategies. For those entrenched within maladaptive patterns, cognitively grafting themselves to such beneficial paradigms is a prerequisite measure to achieve mental health, catalyze resiliency and contribute to humanity.

References:

Barron, K. E., Evans, S. W., Baranik, L. E., Serpell, Z. N., & Buvinger, E. (2006). Achievement goals of students with ADHD. Learning Disability Quarterly, 29(3), 137-158. doi:10.2307/30035504

Brown, K. (2007). Addressing Fundamental Questions About Mindfulness. Retrieved Fromhttps://www.tandfonline.com/doi/full/10.1080/10478400701703344?scroll=top&needAccess=true

Brown, K. W., Ryan, R. M., & Creswell, J. D. (2007). Mindfulness: Theoretical Foundations and Evidence for its Salutary Effects. Psychological Inquiry, 18(4), 211-237. doi:10.1080/10478400701598298

Deng, Y., Li, S., & Tang, Y. (2012). The relationship between wandering mind, depression and mindfulness. Mindfulness, 5(2), 124-128. doi:10.1007/s12671-012-0157-7

Dreyfus, G. (2011). Is mindfulness present-centred and non-judgmental? A discussion of the cognitive dimensions of mindfulness. Contemporary Buddhism, 12(1), 41-54. doi:10.1080/14639947.2011.564815

Duncker, K. (1945). On problem-solving. Germantown, NY: Periodicals Service Company.

Firmin, M. (2007). Correlations among symptoms of ADHD and peer relationships, academic performance, and self-image. Professional Issues in Counseling, 7, 1-6.

Gibson, J. (2019). Mindfulness, Interoception, and the Body: A Contemporary Perspective. Frontiers in Psychology, 10. doi:10.3389/fpsyg.2019.02012

Gil, S., & Droit-Volet, S. (2009). Time perception, depression and sadness. Behavioural Processes, 80(2), 169-176. doi:10.1016/j.beproc.2008.11.012

Hawkins, W. L., French, L. C., Crawford, B. D., & Enzle, M. E. (1988). Depressed affect and time perception. Journal of Abnormal Psychology, 97(3), 275–280. https://doi.org/10.1037/0021-843X.97.3.275

Holt, J. (2018). When Einstein walked with Gödel: Excursions to the edge of thought. New York: Farrar, Straus and Giroux.

Hebert, K. R. (2016). The association between sensory processing styles and mindfulness. British Journal of Occupational Therapy, 79(9), 557-564. doi:10.1177/0308022616656872

Kappes, A., Singmann, H., & Oettingen, G. (2012). Mental contrasting instigates goal pursuit by linking obstacles of reality with instrumental behavior. Journal of Experimental Social Psychology, 48(4), 811-818. doi:10.1016/j.jesp.2012.02.002

Keng, S., Smoski, M. J., & Robins, C. J. (2011). Effects of mindfulness on psychological health: A review of empirical studies. Clinical Psychology Review, 31(6), 1041 1056. doi:10.1016/j.cpr.2011.04.006

Kharlas, D. A., & Frewen, P. (2016). Trait mindfulness correlates with individual differences in multisensory imagery vividness. Personality and Individual Differences, 93, 44-50. doi:10.1016/j.paid.2015.09.027

Kirmayer, L. J. (2015). Mindfulness in cultural context. Transcultural Psychiatry, 52(4), 447-469. doi:10.1177/1363461515598949

Koestner, R., Lekes, N., Powers, T. A., & Chicoine, E. (2002). Attaining personal goals: Self-concordance plus implementation intentions equals success. Journal of Personality and Social Psychology, 83(1), 231–244. https://doi.org/10.1037/0022-3514.83.1.231

Kong, R. (2019). The Role of Time Perception in Depression. Journal of Depression and Anxiety, 08(04). doi:10.35248/2167-1044.19.8.348

Korponay, C., Dentico, D., Kral, T. R., Ly, M., Kruis, A., Davis, K., . . . Davidson, R. J. (2019, August 19). The Effect of Mindfulness Meditation on Impulsivity and its Neurobiological Correlates in Healthy Adults. Retrieved from https://www.nature.com/articles/s41598-019-47662-y

Lernia, D. D., Serino, S., Pezzulo, G., Pedroli, E., Cipresso, P., & Riva, G. (2018). Feel the Time. Time Perception as a Function of Interoceptive Processing. Frontiers in Human Neuroscience, 12. doi:10.3389/fnhum.2018.00074

Mattes, J. (2019, December 06). Systematic Review and Meta-Analysis of Correlates of FFMQ Mindfulness Facets. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/31866899

Mayo, J. P., & Sommer, M. A. (2013). Neuronal correlates of visual time perception at brief timescales. Proceedings of the National Academy of Sciences, 110(4), 1506 1511. doi:10.1073/pnas.1217177110

Mcvay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). Psychological Bulletin, 136(2), 188-197. doi:10.1037/a0018298

Oettingen, G., & Schwörer, B. (2012). Mind wandering via mental contrasting as a tool for behavior change. Frontiers in Psychology, 4. doi:10.3389/fpsyg.2013.00562

Ondobaka, S., Kilner, J., and Friston, K. (2017). The role of interoceptive inference in theory of mind. Brain Cogn. 112, 64–68. doi: 10.1016/j.bandc.2015.08.002

Race, E., Keane, M. M., & Verfaellie, M. (2013). Living in the moment: Patients with MTL amnesia can richly describe the present despite deficits in past and future thought. Cortex, 49(6), 1764-1766. doi:10.1016/j.cortex.2013.02.010

Raffaelli, Q., Mills, C., de Stefano, N.-A., Mehl, M. R., Chambers, K., Fitzgerald, S. A., Wilcox, R., Christoff, K., Andrews, E. S., Grilli, M. D., O'Connor, M.-F., & Andrews-Hanna, J. R. (2021). The think aloud paradigm reveals differences in the content, dynamics and conceptual scope of resting state thought in Trait brooding. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-98138-x

Rothaupt, J. W., & Morgan, M. M. (2007). Counselors' and counselor educators' practice of mindfulness: A qualitative inquiry. Counseling and Values, 52(1), 40–54. https://doi.org/10.1002/j.2161-007X.2007.tb00086.x

Park, T., Reilly-Spong, M., & Gross, C. R. (2013). Mindfulness: A systematic review of instruments to measure an emergent patient-reported outcome (PRO). Quality of Life Research, 22(10), 2639-2659. doi:10.1007/s11136-013-0395-8

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. Journal of Personality and Social Psychology, 77(6), 1296–1312. https://doi.org/10.1037/0022-3514.77.6.1296

Seema, R., Quaglia, J. T., Brown, K. W., Sircova, A., Konstabel, K., & Baltin, A. (2014). The Estonian Mindful Attention Awareness Scale: Assessing Mindfulness Without a Distinct Linguistic Present Tense. Mindfulness, 6(4), 759–766. doi: 10.1007/s12671-014-0314-2

Shakespeare, W., & Gibson, R. (2005). Macbeth. Cambridge, UK: Cambridge University Press.

Sheldon, K. M., & Elliot, A. J. (1999). Goal striving, need satisfaction, and longitudinal well-being: The self-concordance model. Journal of Personality and Social Psychology, 76(3), 482–497. https://doi.org/10.1037/0022-3514.76.3.482

Smalley, S. L. (2009). Mindfulness and attention deficit hyperactivity disorder. Journal of Clinical Psychology. doi:10.1038/npg.els.0001482

Smallwood, J., & Schooler, J. W. (2007). The restless mind. PsycEXTRA Dataset. doi:10.1037/e633982013-080

Smyth, A., Werner, K. M., Milyavskaya, M., Holding, A. C., & Koestner, R. (2020). Do mindful people set better goals? investigating the relation between trait mindfulness, self-concordance, and Goal Progress. https://doi.org/10.31234/osf.io/vshgk

Sobol-Kwapinska, M. (2009). Forms of present time orientation and satisfaction with life in the context of attitudes toward past and future. Social Behavior and Personality: An International Journal, 37(4), 433-440. doi:10.2224/sbp.2009.37.4.433

Tarrasch, R., Margalit-Shalom, L., & Berger, R. (2017). Enhancing Visual Perception and Motor Accuracy among School Children through a Mindfulness and Compassion Program. Frontiers in Psychology, 8. doi:10.3389/fpsyg.2017.00281

Tate, R. (2010). Attention and Memory in Child Language Disorders. Emotional and Social Consequences of Memory Disorders. Handbook of Child Language Disorders, 219-233. doi:10.4324/9780203837764-16

Trousselard, M., Steiler, D., Raphel, C., Cian, C., Duymedjian, R., Claverie, D., & Canini, F. (2010). Validation of a French version of the Freiburg Mindfulness Inventory short version: Relationships between mindfulness and stress in an adult population. BioPsychoSocial Medicine, 4(1), 8. doi:10.1186/1751-0759-4-8

Wilding, E. (2004). Memory problems in brain-injured adults. The Handbook of Memory Disorders., 127(7), 1681-1682. doi:10.1093/brain/awh185

Wilson, T. D., Reinhard, D. A., Westgate, E. C., Gilbert, D. T., Ellerbeck, N., Hahn, C., . . . Shaked, A. (2014). Just think: The challenges of the disengaged mind. Science, 345(6192), 75-77. doi:10.1126/science.1250830

Zimbardo, P. G., & Boyd, J. N. (2014). Putting Time in Perspective: A Valid, Reliable Individual-Differences Metric. Time Perspective Theory; Review, Research and Application, 17–55. doi: 10.1007/978-3-319-07368-2_2

Zimbardo, P. G., & Boyd, J. (2010). The Time Paradox. London: Rider.

Zimbardo P, Boyd J. The time paradox: The new psychology of time that will change your life. Simon and Schuster, UK, 2008.

REGIONAL HEALTHCARE DISPARITIES: ANALYZING THE SERVICE RADII FOR SELECT PHYSICIAN SPECIALITIES ACROSS ARIZONA

004

Nandini Sodhi
Dr. Patricia Harrison - Monroe (Mentor)

Abstract:

There is a great disparity between healthcare access in urban and rural areas nationally. Arizona is no stranger to this crisis, with 7 out of 15 counties considered rural. One of the driving factors of this disparity is the large difference in the number of physicians practicing in rural versus urban counties. To quantify the crisis, this paper analyzes the numbers of primary care physicians (internists, family medicine specialists, and pediatricians), cardiologists, and psychiatrists practicing in each Arizona county. The more physicians serving in an area, the smaller their service radii. As the radius grows, accessibility declines due to factors like travel time. Given the scale of the crisis, there are current efforts to mitigate it. These include monetary incentives for practicing and training physicians and increased rural training programs. But these efforts need to be implemented at a larger scale if this growing crisis is going to be alleviated for future generations.

Introduction & Relevance:

The United States is one of the richest countries in the world and plays a major role in international politics. Its innovations, from Ford's affordable car to Facebook, have changed the world, but the country has its shortcomings as well. Nationalized health insurance has been a conversation for the last decade, and healthcare access is a mounting issue. The disparity between rural and urban centers is daunting. Some rural patients, for instance, have to travel hours for a short fifteen-minute appointment. While visits utilizing telemedicine have increased during the COVID-19 pandemic, the permanence of telemedicine access is unknown since initial legislation was temporary.

Arizona is not at the bottom of rural healthcare access nationally; however, healthcare access remains insufficient across the state. Arizona ranks 31st in the physician-to-people ratio across the U.S. (Koch, et al., 2019). Nationally, there are approximately 237 physicians for every 100,000 people (92 PCPs for every 100,000). Arizona reflects the national average with 236 physicians for every 100,000, but is significantly below the national average for PCPs, with only 78 physicians per 100,000 people. The PCP to population ratio is the highest in urban areas (80.1 PCPs/100,000 people) compared to rural areas (10.1 PCPs/100,000 people). This disparity is one of the driving factors why Arizona only meets 40% of its PCP needs (Vaidya, 2020). When reviewing the number of hospital beds per 1,000 people, Arizona is a whole bed below the national average (with ~2 beds/1,000 population versus the national average of ~3 beds/1,000 population) (Vaidya, 2020 & World Bank). Thus, Arizona ranks 21st nationally in this category (World Bank).

During my time as an undergraduate student, I have had a chance to further explore this issue and work with several mentors, including the director of the Arizona Telemedicine Program, Dr. Ronald S. Weinstein. The rural-urban divide in healthcare access is national, and federal and state governments are working in parallel to tackle this issue at both levels. In the last 30 years, Arizona has taken steps to alleviate this issue, for example, with the creation of the of the Arizona Telemedicine Program and broadband legislation, but there is still much work to be done. This paper attempts to quantify the current crisis by looking at the following specialties throughout Arizona and comparing their numbers between counties and cities:

(i) Cardiologists
(ii) Psychiatrists
(iii) Primary Care Providers (PCPs)
- Internists
- Family Medicine Specialists
- Pediatricians

Generalists Versus Specialists

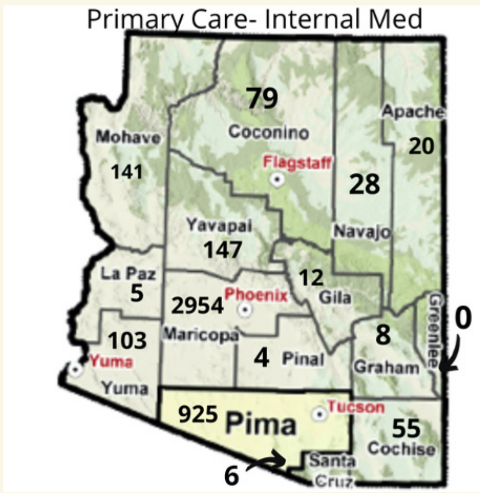
Primary Care Providers (PCPs), including internists, family medicine specialists, and pediatricians, are generalists. Generalists are crucial to triaging non-emergency patients and performing preventive care. If generalists are not easily accessible, many people forgo wellness and preventive care appointments. The Centers for Disease Control and Prevention (CDC) notes that over 100,000 lives a year could be saved if each person were to receive the recommended preventive clinical care (Clarke, 2017). Not only can treatable issues mount to something more, but skipping preventive and wellness care is more expensive. As of right now, 75% of the nation's healthcare spending is going toward preventive care, and it is saving $260 billion dollars a year (Beaton, 2017).

Preventive care includes screenings and tests for cholesterol, diabetes, cancers, age related diseases, immunizations, and more. Catching certain pathologies, like cancer, early is critical to survival. For example, women over 50 years old are screened for breast cancer annually at their wellness visit with their provider. Breast cancer is a very common type of cancer. In fact, it is the second most common type of cancer diagnosed in women after skin cancer. However, breast cancer has several subtypes resulting from different types of mutations. The degree of aggressivity is determined by the level of metastases, which is the migration of the breast tumor from the breast tissue. The migration is quantified via stages. Stage 1 breast cancer is localized in the primary tumor location (where the initial tumor was found). Stages 2-4 refer to the severity of metastasis. Stage 4 implies it has migrated to distant locations. Early detection of the primary breast tumor gives the patient more options for therapy (localized surgery, immunotherapy, radiation, chemotherapy, etc.) and better chances of survival. Without accessible healthcare, many older women forgo this critical screening, leading to more late-stage and lethal diagnoses.

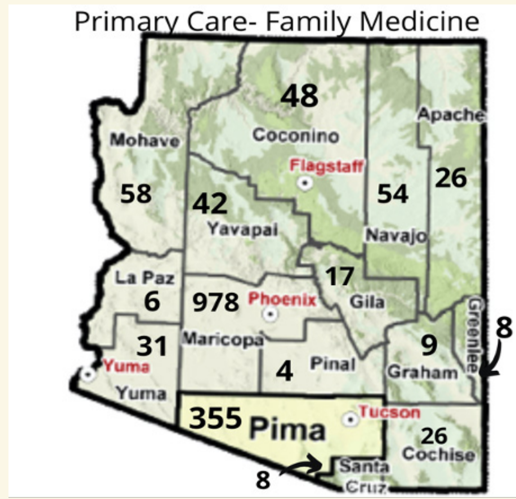
As the gatekeepers of preventive care, ideally PCPs should be accessible throughout the state. But when analyzing the numbers, there is an unequal distribution of internists, family medicine specialists, and pediatricians in Arizona (Figure 1):

Figure 1: *(Source: Arizona Medical Board, 2021)*

a)



b)



c)

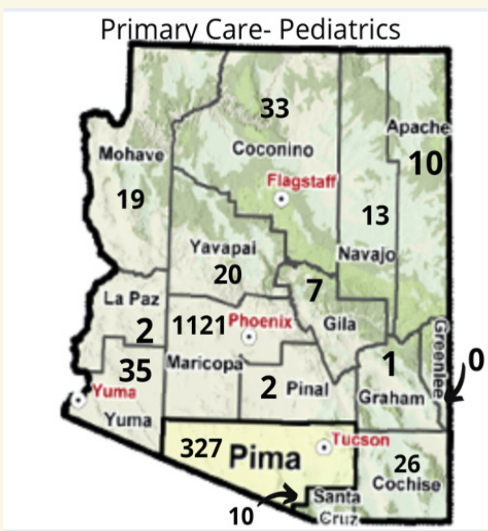


Figure 1 Maps:

a) Map visualizes the distribution of internists in Arizona per county

b) Map visualizes the distribution of family medicine specialists in Arizona per county

c) Map visualizes the distribution of pediatricians in Arizona per county

Each of these generalists plays a crucial role in providing preventative care in their area. The distribution of internists in Arizona is the highest in Maricopa County. Approximately 45% of internists in Maricopa County are in Phoenix. Pima county has the second largest concentration of internists, but 94% of Pima county internists are concentrated in Tucson.

Family Medicine specialists, similarly, are mostly concentrated in Phoenix (with almost 1,000 specialists). And again, Pima is second in the state. Focusing on concentrations in major cities, Yuma has 100% of its family medicine specialists of the county (Yuma) in its urban center of Yuma city (31 family medicine specialists). According to the data, Phoenix has the highest number of family medicine specialists, with Tucson coming in a close second. But when comparing the percentage of family medicine specialists available in the county versus the city, Tucson has a higher concentration. Phoenix only has 41% of the family medicine specialists in Maricopa County, while Tucson has 91% of the family medicine specialists in Pima county. This implies Maricopa County has a more even spread, leading to better access to family medicine specialists than in Pima County.

Similarly, pediatricians in Arizona are concentrated in Maricopa County by a large margin (1121), with Pima County coming in second, once again. In this case, over half the pediatricians (approximately 62%) are in Phoenix. In contrast, 89% of Yuma County's pediatricians are in Yuma, its urban center. In Tucson, as with other specialties reviewed above, a high percentage, 96%, in this case, are concentrated within the city.

The table below (Table 1) indicates the concentration of each type of generalist in major Arizona urban centers in comparison to the county they reside in.

Table 1: *(Source: Arizona Medical Board, 2021)*

	Internists	Family Medicine Specialist	Pediatrician
City/County Ratio			
Tucson/Pima	867/925 (94%)	324/355 (91%)	315/327 (96%)
Yuma/Yuma	102/103 (99%)	31/31 (100%)	31/35 (89%)
Phoenix/Maricopa	1339/2954 (45%)	399/978 (41%)	692/1121 (62%)
Flagstaff/ Coconino	69/79 (87%)	38/48 (79%)^77	31/33 (94%)

The columns are the calculated City/County ratio for the generalist specialty outlined in the top row. The ratios reflect the number of physicians of that generalist specialty concentrated in the city versus the whole county. While a physician present in the same area as the patient does not guarantee access to care due other socio-economic barriers, this analysis attempts outline probability using one determinant.

Although, it makes sense that physicians would be more concentrated in more populated areas where there is more demand, the analysis re-enforces the fact that many rural community members have to travel into urban centers to receive care, even basic care at times. As a state we should be striving for a more equitable, not equal, spread of physicians across population centers.

Specialists, on the other hand, focus on one organ or organ system. Out of the countless specialties and subspecialties present in medicine, I chose to examine cardiologists and psychiatrists and their concentrations throughout Arizona. One of my focuses is cardiology because cardiovascular disease (CVD) is one of the top causes of death in the U.S. For that reason, many Americans need to visit a cardiologist at least once in their lifetime.

I picked psychiatry as the other specialty because psychiatrists play an integral role in providing mental health services for all sectors of society. Nationally, mental health seems to be most accessible in urban areas, with minimal access in rural communities. I wanted to examine Arizona's trend more closely, especially since mental health issues are now becoming more commonly accepted as illnesses needing specialty care.

Looking at the distribution of cardiologists and psychiatrists across Arizona counties we can see the following trends:

Figure 2: *(Source: Arizona Medical Board, 2021)*

a)

b)

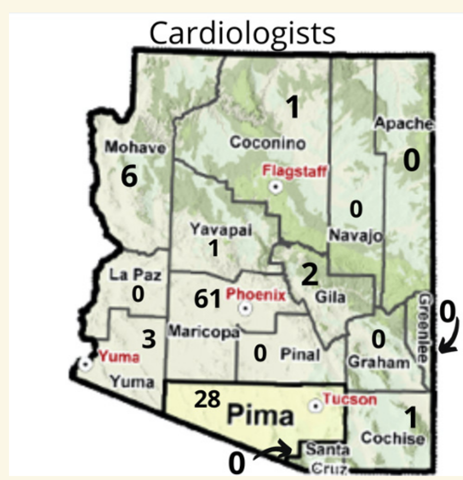


Figure 2 Maps:

a) Map visualizes the distribution of psychiatrists in Arizona per county
b) Map visualizes the distribution of cardiologists in Arizona per county

Psychiatrists in Arizona clearly outnumber cardiologists throughout the state looking at Figure 2. This was surprising because mental healthcare is notorious for being inaccessible. Even though there are still clear differences in accessibility between urban and rural centers, it is reassuring to see the high numbers of psychiatrists across the state. Similar to all the other trends analyzed in this paper, Maricopa County has the highest number of psychiatrists with 562, and Pima County in second with 228 (Figure 2a). In reference to psychiatrists, only La Paz and Graham have none practicing within the counties. Although psychiatrists outnumber cardiologists in the state, the fact that there are none practicing in two counties emphasizes the state's need for psychiatrists in rural areas.

According to Figure 2b, the distribution of cardiologists across the state, as expected, demonstrates a similar trend to that of primary care providers. Maricopa County has the highest concentration of cardiologists, followed by Pima County. Although there is a stark difference between the graphs in Figure 1 (Primary Care) above and the distribution of cardiologists, several counties throughout Arizona (Apache, Navajo La Paz, Santa Cruz, Graham, Greenlee, Pinal) have no cardiologists at all. Regarding primary care providers (Figure 1), only Greenlee had zero pediatricians or internal medicine specialists, although there are eight family medicine specialists. Since Greenlee is an extremely small county, oftentimes, hospitals in Graham County (west of Greenlee) are utilized by its citizens. Thankfully, when looking at internists, there are eight in Graham County and one pediatrician. But, when reviewing the distribution of cardiologists, Apache County, which is north of Greenlee County, has none; west of Graham and Pinal Counties there are also none; and Navajo County, north of Graham County, also has no cardiologists. Table 2 below details the distribution. The low numbers of cardiologists across the state means that people have to travel to major cities, such as Tucson and Phoenix, to visit a cardiologist. With the growing prevalence of cardiovascular disease and other cardiothoracic issues, this indicates a large disparity between urban and rural centers in terms of access to specialized cardiology care.

The exact distribution of the specialists (cardiologists and psychiatrists) analyzed throughout the state and major cities is detailed in Table 2 below.

Table 2: *(Source: Arizona Medical Board, 2021)*

	Cardiologists	*Psychiatrists*
City/County		
Tucson/Pima	27/28 (96%)	219/228 (91%)
Yuma/Yuma	3/3 (100%)	5/7 (100%)
Phoenix/Maricopa	24/61 (39%)	289/562 (51%)
Flagstaff/ Coconino	1/1 (100%)	20/20 (100%)

The columns are the calculated City/County ratio for the specialty outlined in the top row. The ratios reflect the number of the physicians of that specialty concentrated in the city versus the whole county.

The People

Given the population distribution, it makes sense that the physician specialties described above are concentrated in Maricopa and Pima Counties. Maricopa County has the highest population concentration, approximately 4.4 million, in the state. Over 60% of the population in Arizona resides in Maricopa. Pima County comes in second with a little over 1 million people. Although the state's population is concentrated in these two areas, it does not discount the fact that the rest of the population outside of these centers needs reliable, accessible medical care. Unfortunately, the more sparsely populated counties, such as Greenlee (almost 10,000 people), are representative of other counties that have a low number of physicians. Population density throughout the state follows a similar trend. Table 3 below summarizes these trends:

County	**Population**	**Population Density** *(People/Sq. Mile)*
Apache	71,887	11,218
Cochise	125,922	6,219
Coconino	143,476	18,661
Gila	54,018	4,795
Graham	38,837	4,641
Greenlee	9,498	1,848
La Paz	21,108	4,514
Maricopa	4,485,414	9,224
Mohave	212,181	13,461
Navajo	110,924	9,960
Pima	1,047,279	9,189
Pinal	462,789	5,374
Santa Cruz	46,498	1,238
Yavapai	235,099	8,128
Yuma	213,787	5,519

Table 3:

(Source: Census Bureau Quick Facts)

Focusing on the population concentration in urban areas compared to counties, Maricopa County had the highest number of people living outside of Phoenix — its urban center. Surrounding Phoenix, there are several suburban areas such as Chandler, Tempe, Scottsdale, etc. This aligns with the finding discussed above that many physicians in Maricopa County practice outside Phoenix in its surrounding areas. In contrast, when comparing Yuma County with the city of Yuma, there is no difference in population density. This suggests that the vast majority of Yuma County's residents live within Yuma city proper. In the case of Pima County, almost half a million people live outside its main urban center of Tucson. And, lastly, in Coconino County, approximately 68,438 people live outside the city of Flagstaff. Table 4 below outlines these numbers.

Table 4: *(Source: Census Bureau Quick Facts)*

Difference between:		
	Pima - Tucson Population	499,206 people
	Yuma County - Yuma Population	0 people
	Maricopa - Phoenix Population	2,804,422 people
	Coconino - Flagstaff Population	68,438 people

Population differences between County and its main Urban Center

The data above in Table 4 can be used to determine the need for more physicians outside urban centers. For example, the fact that 100% of family medicine physicians in Yuma County are in the city of Yuma may not be a problem since nearly all of the population of Yuma County resides within Yuma city. But the fact that 91% of family medicine physicians in Pima County are in Tucson is a hindrance to healthcare access since there are close to half a million people in Pima County living outside of Tucson who need accessible primary care physicians. Within Pima County the west most town is about two hours and fifteen minutes away. So patients may need to travel upwards of two hours to receive preventive care in Tucson within Pima County.

What Does This Mean?

The lack of physicians outside of urban centers goes hand in hand with the number of hospitals available. There are seven rural counties in Arizona including Apache, Gila, Graham, Greenlee, La Paz, and Navajo Counties. All of these rural counties have less than three hospitals each. In contrast, the urban centers of Phoenix and Tucson have 22 and 18 hospitals, respectively. The extremes across the counties in the state are zero hospitals in Greenlee County compared to 78 hospitals in Maricopa County. Table 5 outlines this information.

Table 5: *(Sources: Arizona Hospital Facility ID & List of Rural Counties And Designated Eligible Census Tracts in Metropolitan Counties)*

County	**Number of Hospitals**
Apache *(RURAL)*	2
Cochise	5
Coconino	4
Gila *(RURAL)*	2
Graham *(RURAL)*	1
Greenlee *(RURAL)*	0
La Paz *(RURAL)*	1
Maricopa	78
Mohave	6
Navajo *(RURAL)*	3
Pima	20
Pinal	5
Santa Cruz *(RURAL)*	1
Yavapai	7
Yuma	2
TOTAL	**137**

Number of Hospitals in each AZ County

Comparing the distribution of hospitals within counties, we see most hospitals are located within urban centers. This is possibly due to the cost of running a hospital. For example, 90% of Pima County's hospitals are in Tucson. 100% of the hospitals of Yuma County are in Yuma. And 75% of the hospitals of Coconino County are in Flagstaff. In contrast, 28% of the hospitals of Maricopa County are in Phoenix. This exception is most likely due to the fact that Phoenix is surrounded by densely populated suburban areas that are all interconnected.

Conclusion

While, thankfully, Arizona is not the worst state in the country regarding healthcare access, the current numbers are not reassuring, either. Due to a lack of resources and funding associated with the ongoing pandemic, many rural hospitals are closing, and healthcare workers are overwhelmed and suffering from burn out.

The COVID-19 pandemic has served as a reminder that it is vital that we come together to address this situation as a state. It is imperative that the state leadership takes the toll of the pandemic on the healthcare industry seriously.

It is critical Arizona invites and trains new physicians, while facilitating the construction of infrastructure so more physicians are able to practice across the state. The onus for this change is not only on physicians but also legislators, who need to realize this issue and promote concrete change. Arizona's rural counties and tribal nations need access to physicians and hospitals, generalist and specialty care, and they require this access at a reasonable distance, not hours or expensive helicopter rides away.

Looking Forward...

While telemedicine implementation may be a part of the answer, especially since some COVID-era telemedicine policy changes are likely here to stay, it has its limitations. Working actively to address those obstacles with increased technological access in clinics and broadband access is important, but the decreased healthcare workforce in the rural areas also needs to be addressed. Fortunately, the Arizona State Legislature has realized this issue and is working in conjunction with the University of Arizona to alleviate it; namely through the University of Arizona Primary Care Scholarship Program. This program awards a full tuition scholarship to current medical students who commit to serve as a primary care physician (which encompasses family medicine, general internal medicine, geriatric medicine, general pediatrics, general surgery, psychiatry, obstetrics and gynecology, and recently general surgery) in rural Arizona after residency (Arizona Primary Care Physician Scholarship).This incentivizes the next generation of physicians to serve in areas that have a disparity in access. A similar federal program also exists. These types of programs also ease the monetary consequences of choosing a less lucrative field, like primary care, since these students graduate with less debt. Some programs focus on recruiting practicing doctors (either residents or attendings) by offering student loan payments, monthly stipends, or student loan forgiveness if they practice in a medically underserved area for a certain period of time (Fenyanova, 2018). Additionally, rural physician salaries tend to be five to ten percent higher than urban physicians (Carey, 2020). Coupled with the lower cost of living in rural areas, this greatly increases the income of rural physicians compared to urban. There have also been efforts to create more training programs, whether that be medical school clerkships or residencies, in medically underserved areas. This allows for these areas to retain physicians early in their career. The University of Arizona, for example, has the Rural Health distinction track, which gives students the opportunity to rotate at a rural health center with a mentor during the summer after their first year. This provides rising second year medical students' with immersive clinical experience along with insight on where they would like to serve in the future.

These efforts have made a difference, but to make the change needed, all these programs need to be rolled out at a larger scale. It is critical we act in a timely manner so the gap between urban and rural healthcare doesn't continue to widen.

More schools should continue to waive tuition in exchange for service in rural areas and offer opportunities to experience rural medicine. More rural areas should incentivize physicians with loan forgiveness, stipends, higher income, etc. Finally, the government should provide more infrastructure to ensure schools and health centers are able to attract and retain physicians to work in underserved areas.

References:

Arizona Hospital Facility ID List. Arizona Department of Health Services. (2020, September). Retrieved February 17, 2021. https://www.azdhs.gov/documents/preparedness/public-healthstatistics/health-facility-cost-reporting/hospital/HOSP_fid.pdf.

Beaton, Thomas. (2017, August 29). How Preventive Healthcare Services Reduce Spending for Payers. HealthPayerIntelligence. Retrieved March 3, 2022. https://healthpayerintelligence.com/news/how-preventive-healthcare-services-reduce-spending-for-payers.

Carey, Liz. (2020, June 02). Pay is Higher for Rural Physicians, New England Journal Article Says. The Daily Yonder. Retrieved March 31, 2021. https://dailyyonder.com/pay-is-higher-for-rural-physicians-new-england-journal-article-says/2020/02/11/.

Clarke, Karen. Campbell County Health, & Messimer, F. (2017, November 1). Preventive care saves lives. Campbell County Health. Retrieved March 3, 2022. https://www.cchwyo.org/news/2017/november/preventive-care-saves-lives/.

Fedyanova, Yana. (2018). Incentivizing young doctors to practise in underserved areas. Canadian Medical Association Journal. 190(7): E203. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5828896/.

Find your MD/Resident/Temp License [Search Engine]. Arizona Medical Board. Retrieved February 17, 2021. https://www.azmd.gov/DoctorSearch/DoctorSearch.

Hospital beds (per 1,000 people). (n.d.). The World Bank. Retrieved February 17, 2021. https://data.worldbank.org/indicator/SH.MED.BEDS.ZS.

Koch, Bryna, Coates S, Campos-Outcalt D, Carter H, and Derksen D. (2019, October). Arizona Primary Care Physician Workforce Report. The University of Arizona Center for Rural Health. Retrieved February 17th, 2021. https://uahs.arizona.edu/sites/default/files/2019_az_primary_care_physician_workforce_report.pdf.

List of Rural Counties And Designated Eligible Census Tracts in Metropolitan Counties. (2018). Retrieved February 17, 2021, from https://www.hrsa.gov/sites/default/files/hrsa/ruralhealth/resources/forhpeligibleareas.pdf

Primary Care Physician Scholarship. (n.d.). The University of Arizona Health Services. Retrieved March 31, 2021. https://financial-aid.medicine.arizona.edu/pcp-scholarship.

Quick Facts.The United States Census Bureau. Retrieved February 17, 2021. https://www.census.gov/quickfacts/fact/table/US#.

Vaidya, A. (n.d.). States ranked by hospital beds per 1,000 population. Retrieved February 17, 2021, from https://www.beckershospitalreview.com/rankings-and-ratings/states-ranked-by-hospital-beds-per-1-000-population.html

TWO-DIMENSIONAL AERODYNAMIC ANAYSIS OF FLIGHT IN THE SMALLEST INSECTS

005

Hrithik Aghav, Laura A. Miller

Abstract:

Two-dimensional immersed boundary simulations were performed to determine how stroke plane angle and wing flexibility affect aerodynamic efficiency and energy efficiency for the smallest flying insects (here boundary refers to the boundary of the elastic structure immersed in the fluid). Extensive experimental data pertaining to small insect flight is unavailable due to the difficulties associated with directly observing the flight of the smallest insects and therefore, their flight mechanisms are still largely unknown. The immersed boundary method was used to solve the fully coupled fluid-structure interaction problem of a flexible wing immersed in a two-dimensional viscous fluid. We considered five different strokes: a horizontal stroke, three different hybrid strokes, and a vertical stroke. We also considered five different wing flexibilities ranging from rigid to highly flexible. Aerodynamic efficiency was defined as the ratio of the average vertical force coefficient to the average total force coefficient and energy efficiency was defined as the ratio of the average vertical force generated by a wing to the average power delivered by the wing to the surrounding fluid. The results indicate that at Reynolds numbers (Re) relevant to small insect flight (4 – 60), both aerodynamic efficiency and energy efficiency decrease with increasing stroke plane angle regardless of wing flexibility.

At Re pertinent to small insect flight, a rigid wing is aerodynamically as well as energetically more efficient than flexible wings at all stroke plane angles. This suggests that a rigid wing with a horizontal stroke could be aerodynamically as well as energetically the most efficient wing flexibility and stroke plane angle combination in the flight of the smallest insects.

Keywords: **Immersed Boundary Method, Computational Fluid Dynamics, Small Insect Flight**

Introduction

Insect flight has been a topic of interest for biologists and physicists for more than a century (Wang, 2005). There are several reasons why research pertaining to insect flight is considered important. For instance, insect flight research can aid in the understanding of the mechanisms that govern insect dispersal by flight. Dispersal is an important topic in entomology and its understanding can facilitate progress in the areas of population genetics, biogeography, pest management, and biodiversity among many others (Naranjo, 2019; Bodino et al., 2021; Cameron et al., 2013). The smallest flying insects are of significant agricultural importance and understanding their flight mechanisms may aid in the biological control of agricultural pests (Jones et al., 2015). For instance, the smallest flying insects include parasitoid wasps which are well-known biological control agents for arthropod pests in agricultural and forest ecosystems (Wang et al., 2019). A comprehensive understanding of the aerodynamics of parasitoid wasps could lead to better strategies for the control of arthropod pests. Thrips are another class of smallest insects that are known to infect cotton seedlings in the United States and are a huge challenge to deal with for growers (Cook et al., 2011). Insights into their flight mechanisms could assist in the development of techniques for limiting the damage they cause to agricultural crops. Engineers are also interested in insect flight research because of its utility in the design of micro-aerial vehicles (MAVs). In recent years, bio-inspired flapping-wing MAVs have been developed, thus, demonstrating the feasibility of flapping wing drones whose sizes are like those of insects (Liu et al., 2016). As can be deduced, research related to insect flight is of both agricultural and engineering importance.

List of Symbols

Re	Reynolds Number	k_S	Spring Stiffness
ρ	Fluid Density	R_L	Spring Resting Length
μ	Dynamic Viscosity	E_{spring}	Spring Elastic Potential Energy
c	Wing Chord Length	F_{spring}	Spring Deformation Energy
U	Wing Tip Velocity	k_{NIB}	Non-invariant beam bending stiffness
A_o	Stroke Amplitude	$\mathbf{F}_{\mathrm{beam}}$	Non-invariant beam deformation force
f	Stroke Frequency	$\delta(\mathbf{x})$	Two-Dimensional Delta Function
β	Stroke Plane Angle	$\mathbf{X}(n,t)$	Cartesian Coordinates of the Material Point
α	Angle of Attack	U_{max}	Maximum Velocity of the Wing Chord Center
α_o	Mean Angle of Attack	U_{rms}	RMS Velocity of the Wing Chord Center
B	Rotation Amplitude	F_V	Vertical Force
$\mathbf{x}$	Eulerian Position (x,y)	F_H	Horizontal Force
$\mathbf{u}(\mathbf{x},t)$	Fluid Velocity	C_V	Vertical Force Coefficient
$\mathrm{p}(\mathbf{x},t)$	Fluid Pressure	C_H	Horizontal Force Coefficient
$\mathbf{f}(\mathbf{x},t)$	Force per unit area	C_T	Total Force Coefficient
$\mathbf{F}(n,t)$	Force per unit length	$\overline{C_V}$	Average Vertical Force Coefficient
n	Lagrangian Position	$\overline{C_H}$	Average Horizontal Force Coefficient
t	Time	$\overline{C_T}$	Average Total Force Coefficient
$\mathbf{X}_{\mathrm{M}}$	Lagrangian Point Coordinate	P	Power delivered by the wing to the fluid
$\mathbf{X}_{\mathrm{M}}^{\mathrm{T}}$	Target Point Coordinate		
$\mathbf{X}_{\mathrm{SL}}$	Slave Node Coordinate		
k_T	Stiff Spring Stiffness		
$\mathbf{F}_{\mathrm{T}}$	Stiff Spring Deformation Force		
E_T	Stiff Spring Deformation Energy		

It has been known for a long time that the aerodynamics and flight dynamics of insect flight are significantly different from those of fixed-wing aircraft (Wang, 2004; Ellington, 1999). Fixed-wing airplanes generally execute steady, level flight at small angles of attack and therefore, primarily utilize aerodynamic lift to generate the weigh-supporting vertical force.

It has been observed, however, that insects pitch their wings at large angles of attack to remain airborne (Wang, 2004). This, coupled with the fact that insect flight is unsteady, requires us to move away from classical aircraft aerodynamics and develop novel techniques for understanding insect flight (Ellington, 1999). A variety of experimental studies (Jensen, 1956; Cloupeau et al., 1979; Wilkin and Williams, 1993; Hollick, 1940; Nachtigall, 1973) and theoretical analyses (Demoll, 1919; Weis-Fogh and Jensen, 1956; Von Holst and Kuchemann, 1942) have been carried out in the past century to try and unravel the complexities of insect flight. Most of the early theoretical work was based on quasi-steady analysis where the main assumption is that the instantaneous forces on a wing are determined by its current state of motion and are independent of its flight history.

Multiple experimental studies (Cloupeau et al., 1979; Wilkin and Williams, 1993; Dudley and Ellington, 1990; Vogel, 1967) have proved that this approach fails to explain the flight strategies employed by insects. More recent theoretical work has focused on important unsteady phenomena in flapping flight aerodynamics such as leading-edge vortex (LEV) to shed light on the aerodynamics of insects and birds (Nabawy and Crowther, 2017; Eldredge and Jones, 2019). Such theoretical work produces low-order analytical models that provide insights into the flow physics of flapping flight that are not available from higher-order computational methods (Nabawy and Crowther, 2017). Here order refers to the complexity of the model. A low order model is a simplification of a high-fidelity complex model. Along with more accurate theoretical models, many other methodologies have been developed and utilized to probe insect flight.

Computational fluid dynamics tools have been developed and employed to reveal the three-dimensional aerodynamics of insect flight. Their emergence was accompanied by the development of novel experimental techniques for probing insect flight (Ellington et al., 1996; Dickinson et al., 1999). All these developments have contributed immensely to our understanding of flight techniques employed by insects ranging in size from Drosophila to dragonflies. However, we haven't yet developed a comprehensive understanding of flight in the smallest insects. We are only now beginning to understand the flight mechanisms employed by the smallest insects (Williams and Murphy, 2020; Sarig and Ribak, 2021; Burrows and Dorosenko, 2019). It has been observed that the flight mechanisms employed by the smallest insects are significantly different from those employed by large insects (Miller and Peskin, 2004, 2005, 2009; Wang, 2000) and therefore, the smallest insects demand special attention. The flight of the smallest insects is different from that of large insects in that they fly at low Re (4 - 60) and flap their wings at very high frequencies - frequencies as high as 400 Hz (Weis Fogh, 1973). At such low Re, viscous effects are quite significant. Here "significant" means that the viscous forces cannot be neglected relative to the inertial forces. The Re is a measure of the ratio of inertial forces to viscous forces in fluid flows. In high Re flows, the inertial forces are significantly larger than the viscous forces and therefore, inviscid flow models and the Euler equations can be utilized to model the flow and obtain quantities like lift and moment coefficients. A low Re indicates that the inertial and viscous forces are of comparable magnitudes and therefore, it becomes essential to avoid using inviscid flow assumptions while modeling low Re flows. A notable example that illustrates this is Lighthill's 1973 paper where inviscid flow assumptions for the clap and fling mechanism at low Re resulted in lower forces than those obtained without making any simplifying assumptions about the viscous effects (Lighthill, 1973).

Experimental data pertaining to flight in the smallest insects is highly limited due to the small size of these insects and the difficulties associated with observing their flight in nature (Sane, 2003). The available experimental data for certain small insects indicates that tiny insects utilize the clap and fling mechanism during flight (Weis-Fogh, 1973; Ellington, 1984; Miller and Peskin, 2009). However, it is not clear if this is the only mechanism employed by the smallest insects.

The lack of extensive experimental data related to small insect flight has sparked debates and led to speculation about the flight strategies employed by these insects (Jones et al., 2015). Lift-based strokes in the horizontal plane and drag-based strokes in the vertical plane have been suggested as possible flight mechanisms. Here horizontal plane is perpendicular to the direction of gravity and vertical plane is parallel to the direction of gravity. Weis-Fogh's 1973 paper popularized lift-based strokes through its description of the clap and fling mechanism (Weis-Fogh, 1973). In this mechanism, the insect "claps" its wings at the end of the upstroke and "flings" them apart at the beginning of the downstroke. At the scale of tiny insects, this motion generates more vertical force coefficient than the wing kinematics utilized by large insects and birds. However, the main issue with this strategy is that to fling the wings apart, very large forces are required at Re relevant to flight in the smallest insects (Miller and Peskin, 2009). The reason for this is as follows: during the fling, the wings are moving rapidly at a high angle of attack, and this leads to the generation of high pressure on the windward side of the wings. Also, the fling leads to the generation of large leading-edge vortices around each wing that are positioned in the vicinity of the leeward side of the wings, and this results in a low-pressure region on the leeward side of the wings. The large pressure difference between the windward and leeward sides of the wings results in a large pressure force acting on the wings which, in turn, leads to a large pressure drag on the wings. Also significant is the large viscous drag acting on the wing due to the low Re. The large pressure and viscous drag together make flinging the wings apart a challenging task. Horridge suggested that the smallest insects might employ drag-based vertical strokes to generate vertical force (Horridge, 1956). Subsequent work emphasized the importance of drag-based strokes in insect flight with Wang's work demonstrating that an idealized dragonfly wing motion supports three-quarters of the insect's weight with drag (Wang, 2004). Similar other studies focused on insect flight have considered high Re and therefore, the low Re regime that is relevant to flight in the smallest insects hasn't been investigated rigorously. A recent study conducted by Jones et al. was focused extensively on the performance of lift and drag-based strokes at low Re with the primary focus being vertical force production (Jones et al., 2015). They had performed their investigations for a rigid and a flexible wing.

A specific flight strategy is a result of compromise among many factors including vertical force generation and therefore, it is important to rigorously investigate other flight factors like stability, aerodynamic efficiency, control, etc. as well (Jones et al., 2015). To further understand flight in the smallest insects, computational fluid dynamics is a convenient tool as it allows for the exploration of a variety of wing kinematics and different combinations of flight parameters which otherwise might be difficult or even impossible to measure directly with experiments.

For our study, we considered four different flexible wings and a rigid wing. We were primarily interested in aerodynamic efficiency and energy considerations. We defined a new dimensional quantity and termed it "pseudo-efficiency". This quantity is equal to the ratio of the average vertical force generated by a wing to the average power delivered by the wing to the surrounding fluid and was used as a measure of energy efficiency. Aerodynamic efficiency was defined as the ratio of the time-averaged vertical force coefficient to the time-averaged net force coefficient. The immersed boundary method was utilized to solve the fully coupled fluid structure interaction problem of a flexible wing immersed in a two-dimensional viscous fluid. For each wing, we conducted a parametric study: Five different stroke plane angles were considered: 0° (horizontal stroke), 22.5° (hybrid stroke), 45° (hybrid stroke), 77.5° (hybrid stroke), 90° (vertical stroke). For each stroke plane angle, we varied the Re and for each Re, we calculated the dimensionless forces as functions of dimensionless time (time was normalized using the stroke period). Further, we calculated the aerodynamic efficiency and the pseudo-efficiency as functions of Re for each stroke plane angle. The main objective of this study was to determine how wing flexibility and stroke plane angle affect aerodynamic efficiency and pseudo-efficiency in the flight of the smallest insects.

Methods

Prescribed Wing Kinematics:

The wing kinematics utilized in this study are identical to those used in similar previous studies (Wang, 2004; Jones et al., 2015). This was a two-dimensional study and therefore, it would be more appropriate to call the kinematics utilized here as the kinematics of the wing chord. The kinematics of the wing chord were defined as follows: -

$$[x(t), y(t)] = \frac{A_0}{2}(1 + \cos(2\pi ft))\,[\cos(\beta), \sin(\beta)] + C_o$$

$$\alpha(t) = \alpha_o + B \sin(2\pi ft)$$

in which [x(t), y(t)] is the position of the center of the wing chord at time t, f is the flapping frequency, β is the stroke plane angle, A_o is the stroke amplitude, a (t) is the angle of attack of the wing, a_o is the mean angle of attack, B is the rotation amplitude, and C_o was used to adjust the initial position of the chord. Fig. 1 illustrates a wing stroke that is governed by thiskinematics. For all parameter combinations, the wing chord length c = 0.1 m, A_o = 0.25 m, f =1 Hz, B = π/4, a_o = π/2, and C_o = 0.2 m for x(t) and 0.3 m for y(t). The stroke plane angles (β) considered were 0°, 22.5°, 45°, 77.5°, and 90°.

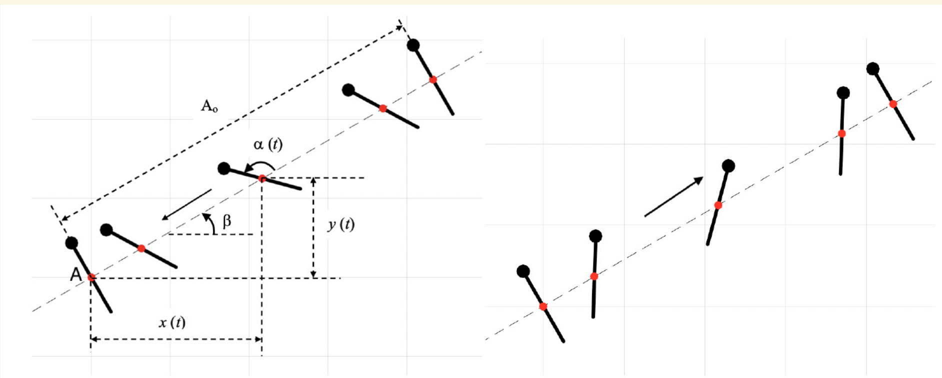


Figure 1. A wing stroke generated by the kinematics described in the methods section. The left part of the figure illustrates the downstroke while the right part illustrates the upstroke. The solid arrows represent the direction of motion of the wing center.

Numerical Method:

We used an open-source two-dimensional computational fluid dynamics code, IB2d, (Battista et al., 2015, 2017, 2018) based on the immersed boundary method developed by Charles Peskin (Peskin, 1972, 1977, 2002) to conduct our investigations. The immersed boundary method has been used to model and study a variety of fluid phenomena such as dynamics of natural and prosthetic heart valves (Griffith et al., 2009), insect flight (Miller and Peskin, 2004, 2005, 2009), jellyfish swimming (Taheri, 2018), and bacterial flagella (Maniyeri et al., 2012). The Navier Stokes equations (in Eulerian form) that govern the fluid behavior are as follows:

$$(1) \qquad \rho\left(\frac{\partial u}{\partial t} + (\mathbf{u}(\mathbf{x}, t) \bullet \nabla)\mathbf{u}(\mathbf{x}, t)\right) = -\nabla p(\mathbf{x}, t) + \mu\Delta\mathbf{u}(\mathbf{x}, t) + \mathbf{f}(\mathbf{x}, t),$$

$$(2) \qquad \nabla \bullet \mathbf{u}(\mathbf{x}, t) = 0$$

Here, u(x, t) is the fluid velocity at position x = [x, y] and time t, p(x, t) is the fluid pressure, f(x, t) is the force per unit area applied to the fluid by the immersed boundary. Eq. (2) is the mass conservation equation for incompressible fluid flow. The equations that govern the interaction between the fluid and the immersed boundary are as follows:

$$(3) \qquad \mathbf{f}(\mathbf{x}, t) = \int \mathbf{F}(n, t)d(\mathbf{x} - \mathbf{X}(n, t))\, dn,$$

$$\frac{\partial X(n,t)}{\partial t} = \mathbf{U}(\mathbf{X}(n, t)) = \int u(x,\ t)\delta(x\ -\ X(n,\ t))\, d\mathbf{x}$$

(4)

Here F(n, t) is the force per unit length applied by the immersed boundary to the fluid as a function of Lagrangian position n and time t. δ (x) is a two-dimensional delta function and X(n, t) gives the Cartesian coordinates at time t of the material point labeled by the Lagrangian parameter n. Eq. (3) applies force from the Lagrangian boundary to the fluid grid, and Eq. (4) evaluates the fluid velocity at the boundary. The boundary is then moved at the local fluid velocity, and this enforces the no-slip boundary condition. Eq. (3) and Eq. (4) couple the Eulerian and Lagrangian variables. In our study, the two-dimensional wing was the immersed boundary.

The forces exerted on the wing by the fluid were non-dimensionalized by $0.5\rho U^2_{rms}c$ where ρ is the fluid density, U_{rms} is the root mean square speed of the center of the wing chord, and c is the chord length. The vertical and horizontal force coefficients are denoted by C_V and C_H, respectively, and the total force coefficient is denoted by C_T. The vertical force generated by the wing is denoted by F_V and the instantaneous power delivered by the wing to the fluid is denoted by P. A horizontal bar above these quantities represents their time-averaged value.

Table 1: Numerical Parameters

Parameter	Value
Grid Resolution	256 × 256
Length of grid in x direction	6c
Length of grid in y direction	6c
Time step (dt)	1.25e-5t
Spatial step size (dx)	2.35e-2c
Final simulation time	4t
Wing chord length	c
Length of Stroke Plane	2.5c
Fluid density	3047.62618 kg/m^2

We used a chord based Re for our simulations. The Re was defined as follows:

$$Re = \frac{\rho U_{max} c}{\mu} = \frac{\rho \pi f A_o c}{\mu}$$

Here, U_{max} is the maximum speed of the center of the wing chord during one wing stroke, and c is the chord length. The Reynolds numbers considered in our study are 1, 4, 16, 32, 64, and 128. The fluid viscosity was varied to alter the Re. The density (ρ) was kept constant at 3047.63 kg/m^2 . We used an open-source software called VisIt, developed, and maintained by Lawrence Livermore National Laboratory for flow visualization and used the data analysis software within IB2d for performing data analysis on the simulation data (Childs, 2012; Battista et al., 2017). Table 1 summarizes the numerical parameters that were used for the simulations in our study (c is the chord length and t is the stroke period).

Grid Convergence Study:

Table 2: Average percent difference in force coefficients over the last two wing beats between different grid resolutions at $Re = 128$ and $\beta = 0°, 45°, 90°$ for the rigid wing

β	Grids	Average % Difference (C_V)	Average % Difference (C_H)
0°	128 × 128 vs. 256 × 256	22.39	0.2274
	256 × 256 vs. 512 × 512	4.3687	3.1439
45°	128 × 128 vs. 256 × 256	26.8348	37.6068
	256 × 256 vs. 512 × 512	3.4384	0.2983
90°	128 × 128 vs. 256 × 256	1.8934	0.7861
	256 × 256 vs. 512 × 512	0.4544	2.0455

To test for spatial convergence, we compared three different grid resolutions: a uniform 128 × 128 discretization, a uniform 256 × 256 discretization, and a uniform 512 × 512 discretization. We considered the following three stroke plane angles for our grid convergence study: 0°, 45°, and 90°. We performed our simulations for the rigid wing at Re =128. This choice was motivated by the difficulty of resolving thin boundary layers and vortex separation for the rigid wing at Re = 128. The force coefficients for the three grid resolutions showed overall good agreement for β = 0° and β = 90° (Fig. 2). The three grid resolutions were compared by calculating the average percent difference in force coefficients between the selected grid and the next finest grid over the last two wing beats (Table 2). The 256 × 256 grid was used for all simulations in our study because the average percent difference in force coefficients between this grid and the next finest grid (512 x 512) was < 5 % for all the three stroke plane angles. The grid convergence study was performed using a custom MATLAB code.

Wing Geometry: Rigid Wing:

To model the rigid wing, we used 75 Lagrangian points. The only fiber model used for the rigid wing simulations was target points. They were used to prescribe the motion of all the Lagrangian points. Each Lagrangian point was associated with a target point and was connected to its target point using a stiff spring, i.e., a spring with zero resting length. The deformation energy of the stiff spring is as follows:

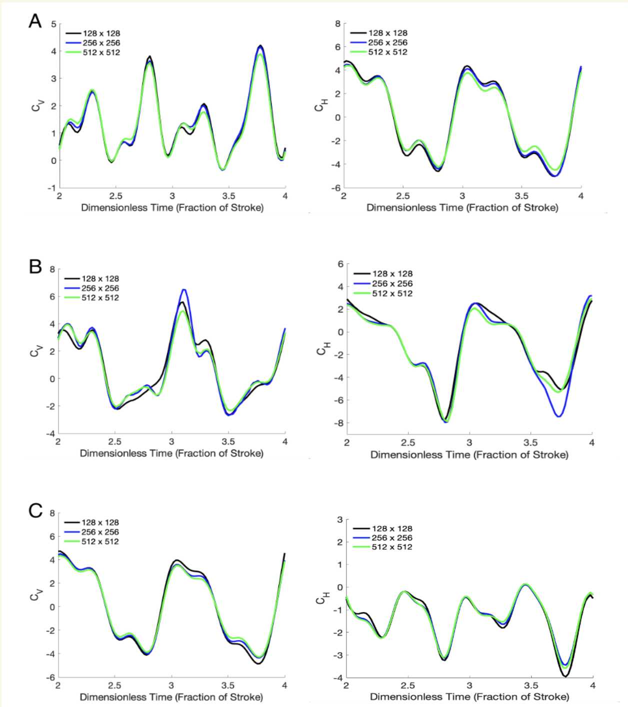


Figure 2. C_V and C_H as functions of dimensionless time for the rigid wing at Re = 128 and (A) $\beta = 0°$, (B) $\beta = 45°$, and (C) $\beta = 90°$. The coefficients were calculated for the following three grid resolutions: 128 x 128, 256 x 256, and 512 x 512.

$$E_T(\mathbf{X}_M) = \frac{1}{2} k_T \, || X_M - X_M^T ||^2,$$

where k_T is the stiffness value of the stiff spring and X_M and X^T_M are the coordinates of the Lagrangian point and the target point, respectively. The corresponding deformation forces are as follows:

$$\mathbf{F}_T = - \frac{\partial E_T}{\partial X_M} = < - k_T(x_M - x_M^T), - k_T(y_M - y_M^T) >$$

It is standard for the stiffness k_T to be very large to move the Lagrangian points in a prescribed manner. The non-dimensional stiffness value k_T' for the rigid wing was 9.6×10^4. k_T was non- dimensionalized in the following manner:

$$k_T^{'} = \frac{k_T}{\rho U_{rms}^2 c}$$

Flexible Wing:

We modeled four different flexible wings for our study. Three different fiber models were used for all the flexible wings - target points, non-invariant beams, and springs. 75 Lagrangian points were used to model each wing and only the first 25 points starting from the leading edge were tethered to target points. The remainder of the wing was allowed to bend as it interacted with the fluid. The non-dimensional stiffness value of the spring associated with the target points was equal to 9.6×10^4 for all the flexible wings. This value was chosen arbitrarily, and it was large enough to ensure that the rigid portion of the flexible wings was moving as per the prescribed kinematics.

Two new fiber models that were used here are springs and non-invariant beams. Springs were used to model the resistance to stretching between successive Lagrangian points. We used Hookean springs to model the connections between successive Lagrangian points. Hookean springs were used because they are relatively easier to implement than other non-linear springs for the purpose of preventing the stretching of the wing during its motion. The Hookean springs implemented in this study prevented the stretching of the wing during its motion to a satisfactory extent. The elastic potential energy of a Hookean spring is as follows:

$$E_{spring} = \frac{1}{2} k_S (|| X_{SL} - X_M|| - R_L)^2,$$

where k_s is the spring stiffness, R_L is the resting length of the spring, and X_M and X_{SL} are the master and slave node coordinates, respectively.

The corresponding deformation force is given by a derivative of the elastic potential energy:

$$F_{spring} = k_S(1 - \frac{R_L}{|| X_{SL} - X_M||}) < x_{SL} - x_M, y_{SL} - y_M >$$

The non-dimensional spring stiffness k_s' was equal to 9.6×10^8 for all the flexible wings and the resting length R_L was equal to 0.0135 m which was the distance between successive Lagrangian points. This large non-dimensional value of spring stiffness was chosen to ensure minimal stretching of the wings during the simulations. k_s was non-dimensionalized in the following manner:

$$k_S^{'} = \frac{k_S}{\rho U_{rms}^2 c}$$

Non-invariant beams were used to model the resistance to bending between 3 successive Lagrangian points. For a set of 3 successive Lagrangian points, say X_L, X_M, and X_R, a non-invariant beam was used to connect the three points. This model assumes a prescribed curvature in both x and y components between the 3 Lagrangian points with corresponding bending stiffness k_{NIB} (Battista et al., 2017). The bending deformation forces are as follows:

$$F_{beam} = k_{NIB} \frac{\partial^4}{\partial n^4}(\mathbf{X}(n, t) - \mathbf{X}_b(n, t))$$

where X(n, t) is the current Lagrangian configuration at time t, $X_b(n, t)$ is the preferred configuration of the non-invariant beam at time t, and k_{NIB} is the bending stiffness of the non-invariant beam. Non-invariant beams were used to connect successive triplets of Lagrangian points over the entire wing for all the flexible wings. The non-dimensional bending stiffness of all non-invariant beams was the same for a specific flexible wing. As mentioned earlier, we modeled four different flexible wings. The non-dimensional bending stiffness values (k_{NIB}') for the four wings are as follows: 3.2×10^{10} (most flexible), 3.4×10^{10}, 6.8×10^{10}, 9.6×10^{10} (least flexible). Wing flexibility is directly related to the non-dimensional bending stiffness value. Wing flexibility increases as k_{NIB}' decreases and it decreases as k_{NIB}' increases. The prescribed curvature in both x and y directions was set to 0 for all the flexible wings. k_{NIB} was non-dimensionlized in the following manner:

$$k'_{NIB} = \frac{k_{NIB}}{\rho U_{rms}^2 c^3}$$

The following are the equations that were used to calculate the aerodynamic performance parameters for our study:

$$C_V = \frac{F_V}{\rho U_{rms}^2 c}$$

$$C_H = \frac{F_H}{\rho U_{rms}^2 c}$$

$$C_T = \frac{F_T}{\rho U_{rms}^2 c}$$

$$C_T = \frac{F_T}{\rho U_{rms}^2 c}$$

$$F_T = \sqrt{F_H^2 + F_V^2}$$

$$\text{Aerodynamic Efficiency} = \frac{\overline{C_V}}{\overline{C_T}}$$

$$\text{Pseudo-efficiency} = \frac{\overline{F_V}}{P}$$

Results

Rigid Wing:

We performed simulations for the rigid wing using the prescribed kinematics presented in the methods section. Each simulation consisted of four stroke cycles. Calculation of aerodynamic forces, aerodynamic efficiency, and pseudo-efficiency was done using simulation data for only the third and fourth stroke cycles that exhibited periodic variation of aerodynamic forces with non-dimensional time. Time was normalized using the stroke period. C_V and C_H are shown as functions of dimensionless time during the third and fourth stroke cycles for all Re considered in this study for $\beta = 0°$ (Fig. 3), the hybrid strokes (Fig. 4), and $\beta = 90°$ (Fig. 5). For the horizontal stroke ($\beta = 0°$), for both the half-strokes, the magnitude of the instantaneous horizontal force coefficient decreased as the Re increased. The magnitude of the instantaneous vertical force coefficient decreased with increasing Re until Re = 64, and then increased with increasing Re. This was true for wing rotation at the beginning of half-strokes up until mid- translation. For the majority of the remaining stroke, the magnitude of instantaneous C_V decreased with increasing Re until Re = 16, and then increased with increasing Re. For the horizontal stroke, at higher Re, C_V was positive for almost the entire stroke while at lower Re, C_V was positive during wing rotation at the beginning of half-strokes and during majority of the translation phase and was negative during wing rotation at the end of half-strokes.

For $\beta = 22.5°$, for the downstroke, the magnitude of instantaneous C_H decreased with increasing Re. For the upstroke, the magnitude of instantaneous C_H decreased with increasing Re for wing rotation at the beginning up until mid-translation. For most of the upstroke beyond mid-translation, the magnitude of instantaneous C_H decreased with increasing Re until Re = 64, and then increased with increasing Re. The magnitude of instantaneous C_V decreased with increasing Re for the downstroke. For the upstroke, the magnitude of instantaneous C_V decreased with increasing Re for wing rotation at the beginning up until mid-translation. For most of the remaining upstroke, the magnitude of instantaneous C_V decreased with increasing Re until Re = 32, and then increased with increasing Re. At both low and high Re, C_V was positive for most of the downstroke. For the upstroke, at both low and high Re, C_V was mostly negative.

For β = 45°, the magnitude of both the instantaneous horizontal force coefficient and the instantaneous vertical force coefficient decreased as the Re increased. C_V was positive during the downstroke and negative during the upstroke. For β = 77.5°, for the downstroke, the magnitude of instantaneous C_H decreased with increasing Re until Re = 16, and then increased

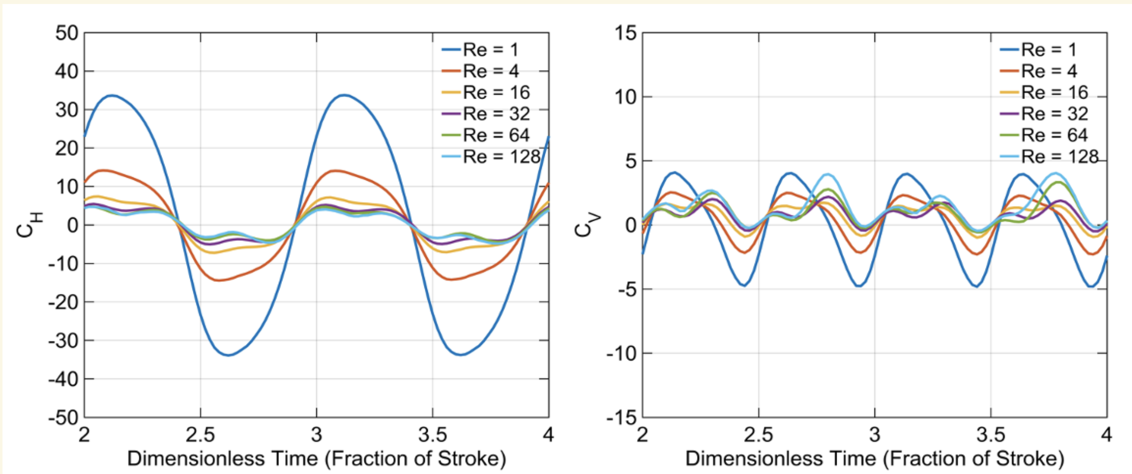


Figure 3. C_H and C_V as functions of dimensionless time during the third and fourth stroke cycles for the horizontal stroke. (Rigid Wing)

with increasing Re. For the upstroke, the magnitude of instantaneous C_H decreased with increasing Re. This was valid from the beginning of the upstroke up until mid-translation. For most of the remainder of the translation phase, the magnitude of C_H increased with increasing Re, and it decreased with increasing Re for wing rotation at the end of the upstroke. As far as C_V is concerned, the magnitude of instantaneous C_V decreased with increasing Re for both the upstroke and the downstroke. C_V was primarily positive for the downstroke and negative for the upstroke.

For the vertical stroke (β = 90°), the variation in the magnitude of instantaneous C_H with increasing Re was almost identical to the variation of the magnitude of instantaneous C_V with increasing Re for the horizontal stroke. The magnitude of instantaneous C_V decreased with increasing Re for the vertical stroke. Like β = 45°, C_V was positive during the downstroke and negative during the upstrokes.

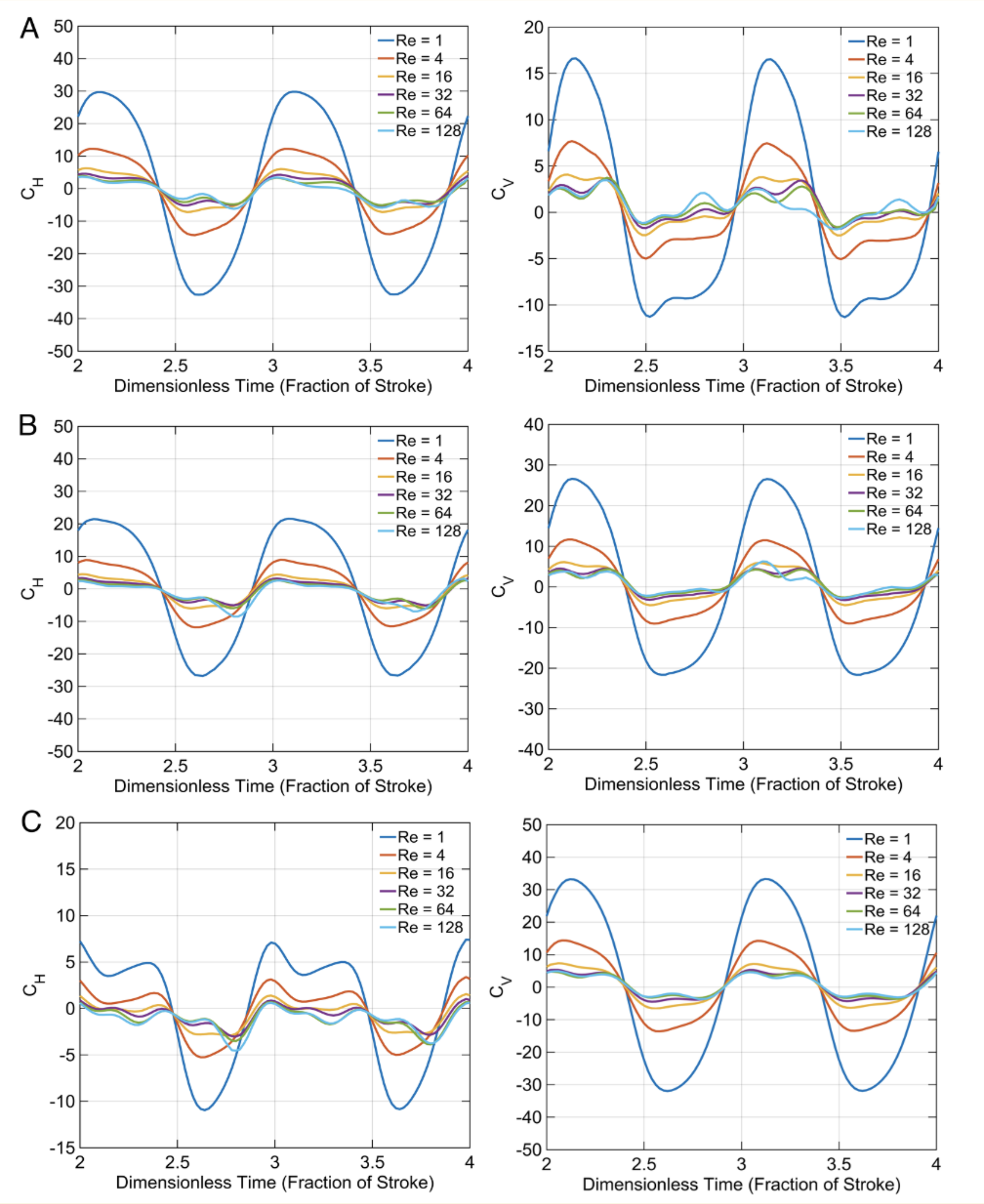


Figure 4. C_H and C_V as functions of dimensionless time during the third and fourth stroke cycles for (A) $\beta = 22.5°$, (B) $\beta = 45°$, and (C) $\beta = 77.5°$. (Rigid Wing)

As far as maximum C_v value is concerned, for the horizontal stroke, at higher Re, C_v achieved its maximum value during wing rotation at the end of the upstroke while at lower Re, C_v achieved its maximum value during wing rotation at the beginning of the downstroke.

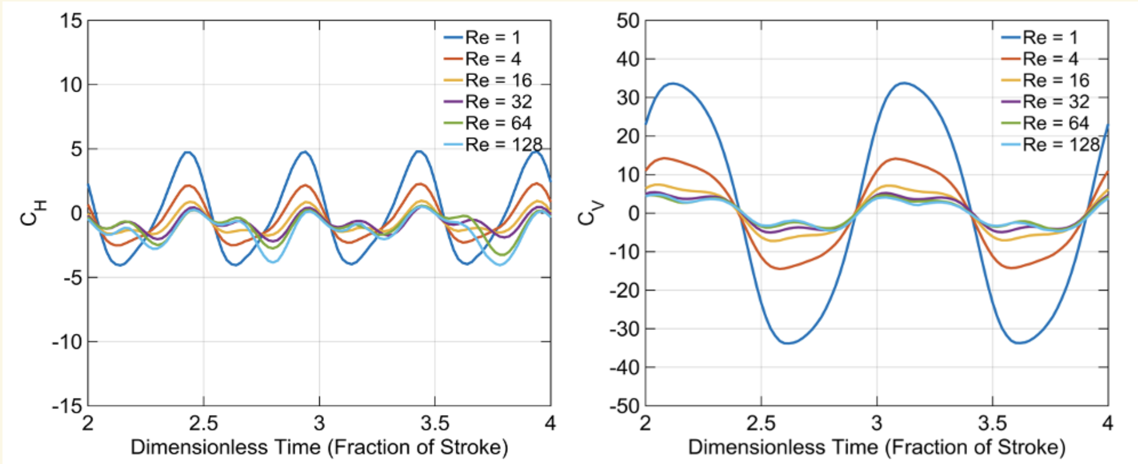


Figure 5. C_H and C_V as functions of dimensionless time during the third and fourth stroke cycles for the vertical stroke. (Rigid Wing)

For β = 22.5°, at lower Re, C_V achieved its maximum value during wing rotation at the beginning of the downstroke while for higher Re, C_V achieved its maximum value during wing rotation at the end of the downstroke. For β = 45°, β = 77.5°, and β = 90°, C_V achieved its maximum value during wing rotation at the beginning of the downstroke for both higher and lower Re. For β = 0° and β = 22.5°, maximum C_H values were higher than maximum C_V values while for β = 45°, β = 77.5°, and β = 90°, maximum C_V values were higher than maximum C_H values. Maximum C_V value decreased as Re increased for β = 22.5°, β = 45°, and β = 90°. For the horizontal stroke, maximum C_V value decreased as Re increased until Re = 16, and then increased as Re increased. For β = 77.5°, maximum C_V value decreased as Re increased until Re = 64, and then increased as Re increased. The basis for some of these results can be understood by looking at the vorticity plots of the fluid around the wing chord at Re = 4 and Re = 128 during the fourth stroke cycle (Fig. 6). Since the wing kinematics were the same for all the five stroke plane angles, the vorticity plots were quantitatively the same for all stroke plane angles at a specific Re. Vorticity plot for β = 22.5°, β = 45°, β = 77.5°, and β = 90°, are equivalent to the vorticity plots for the horizontal stroke rotated in the counterclockwise sense by 22.5°, 45°, 77.5°, and 90°, respectively. Therefore, vorticity plots for only the horizontal stroke are included in this paper.

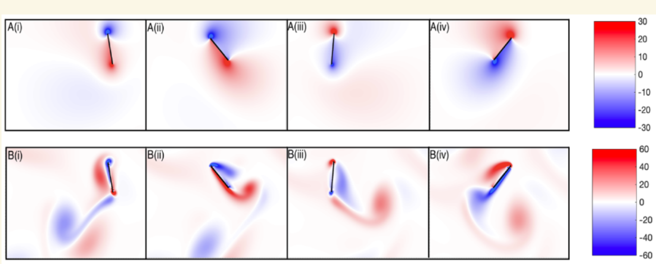


Figure 6. Vorticity plots of the fluid during the fourth stroke cycle at (A) *Re* = 4 and (B) *Re* = 128 shown at (i) the beginning of downstroke, (ii) midway through downstroke, (iii) beginning of downstroke, (iv) midway through upstroke. The color map displays the fluid vorticity (Red - counterclockwise rotation, Blue - clockwise rotation) (Rigid Wing).

Fig. 6 shows the vorticity plots of the fluid around the rigid wing during the fourth stroke cycle at Re = 4 and Re = 128. For Re = 4, leading edge (LEV) and trailing edge (TEV) vortices were formed during wing rotation at the beginning of the downstroke, and they remained attached to the wing during the entire translation phase of the downstroke. These vortices were shed during wing rotation at the end of the downstroke and new vortices were formed during wing rotation at the beginning of the upstroke. They remained attached to the wing during the entire translation phase of the upstroke and were shed during wing rotation at the end of the upstroke. For Re = 128, the vorticity plots were significantly different. A LEV and a TEV were formed during wing rotation at the beginning of the downstroke. The LEV remained attached to the wing during the entire translation phase of the downstroke while the TEV was shed midway through the downstroke. The LEV was shed during wing rotation at the end of the downstroke. New vortices were formed during wing rotation at the beginning of the upstroke. Like the downstroke, the LEV remained attached to the wing during the entire translation phase of the upstroke while the TEV was shed midway through the upstroke. As can be concluded, vortex shedding for Re = 4 was symmetric while that for Re = 128 was asymmetric.

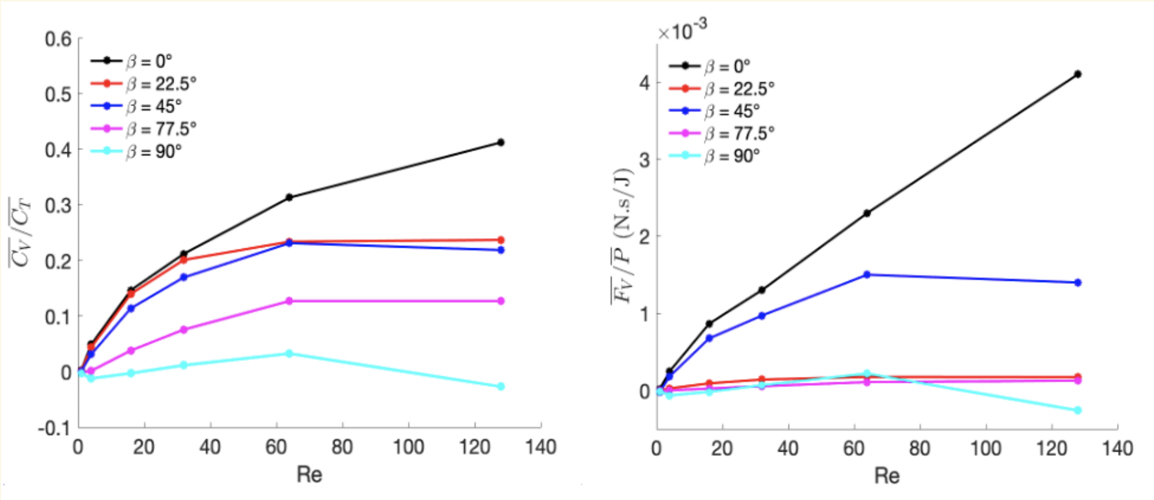


Figure 7. $\overline{C_V}/\overline{C_T}$ as a function of *Re* for all stroke plane angles considered in this study (to the left of the figure) and $\overline{F_V}/\overline{P}$ as a function of *Re* for all stroke plane angles considered in this study (to the right of the figure) (Rigid Wing).

Aerodynamic efficiency (C_V/C_T) and pseudo-efficiency (F_V/P) were calculated as functions of Re for the rigid for all stroke plane angles considered in this study. C_V and C_T were averaged over the third and fourth stroke cycles and similarly, pseudo-efficiency was calculated using the simulation data for only the third and fourth stroke cycles.

Fig. 7 shows the variation of these quantities with Re for the rigid wing for all stroke plane angles considered in this study. For both $\beta = 0°$ and $\beta = 77.5°$, C_V/C_T and F_V/P increased monotonically with increasing Re. For $\beta = 22.5°$, aerodynamic efficiency increased monotonically with increasing Re while pseudo-efficiency increased with increasing Re until Re = 64 and then decreased with increasing Re. For $\beta = 45°$ and $\beta = 90°$, both C_V/C_T and F_V/P increased with increasing Re until Re = 64 and then decreased with increasing Re. The horizontal stroke generated more C_V/C_T than the hybrid and the vertical strokes at all Re considered in this study. Same was the case for the pseudo-efficiency (F_V/P).

Flexible Wings:

As mentioned earlier, we modeled four different flexible wings for our study. The focus of this section is on the variation of aerodynamic efficiency and pseudo-efficiency with wing flexibility at Re > 100 and Re relevant to small insect flight for all stroke plane angles considered in this study:

Variation of Aerodynamic Efficiency with wing flexibility:

Fig. 8 shows aerodynamic efficiency (C_V/C_T) as a function of Re for all the wings at all stroke plane angles considered in this study. At Re > 100, for $\beta = 0°$ and $\beta = 45°$, C_V/C_T increased with increasing wing flexibility until $k_{NIB}' = 6.8 \times 10^{10}$, and then decreased with increasing wing flexibility. For $\beta = 22.5°$, C_V/C_T increased with increasing wing flexibility until $k_{NIB}' = 3.4 \times 10^{10}$, and then showed a decrease for the most flexible wing. For $\beta = 77.5°$ and $\beta = 90°$, C_V/C_T initially decreased with increasing flexibility, then increased with increasing wing flexibility until $k_{NIB}' = 3.4 \times 10^{10}$, and then showed a decrease for the most flexible wing.

At Re pertinent to small insect flight, for $\beta = 0°$ and $\beta = 45°$, C_V/C_T initially decreased with increasing wing flexibility, then increased with increasing wing flexibility until $k_{NIB}' = 6.8 \times 10^{10}$, and then decreased with increasing flexibility. For $\beta = 22.5°$ and $\beta = 77.5°$, C_V/C_T decreased monotonically with increasing wing flexibility. For $\beta = 90°$, C_V/C_T decreased with increasing wing flexibility until $k_{NIB}' = 3.4 \times 10^{10}$, and then showed an increase for the most flexible wing.

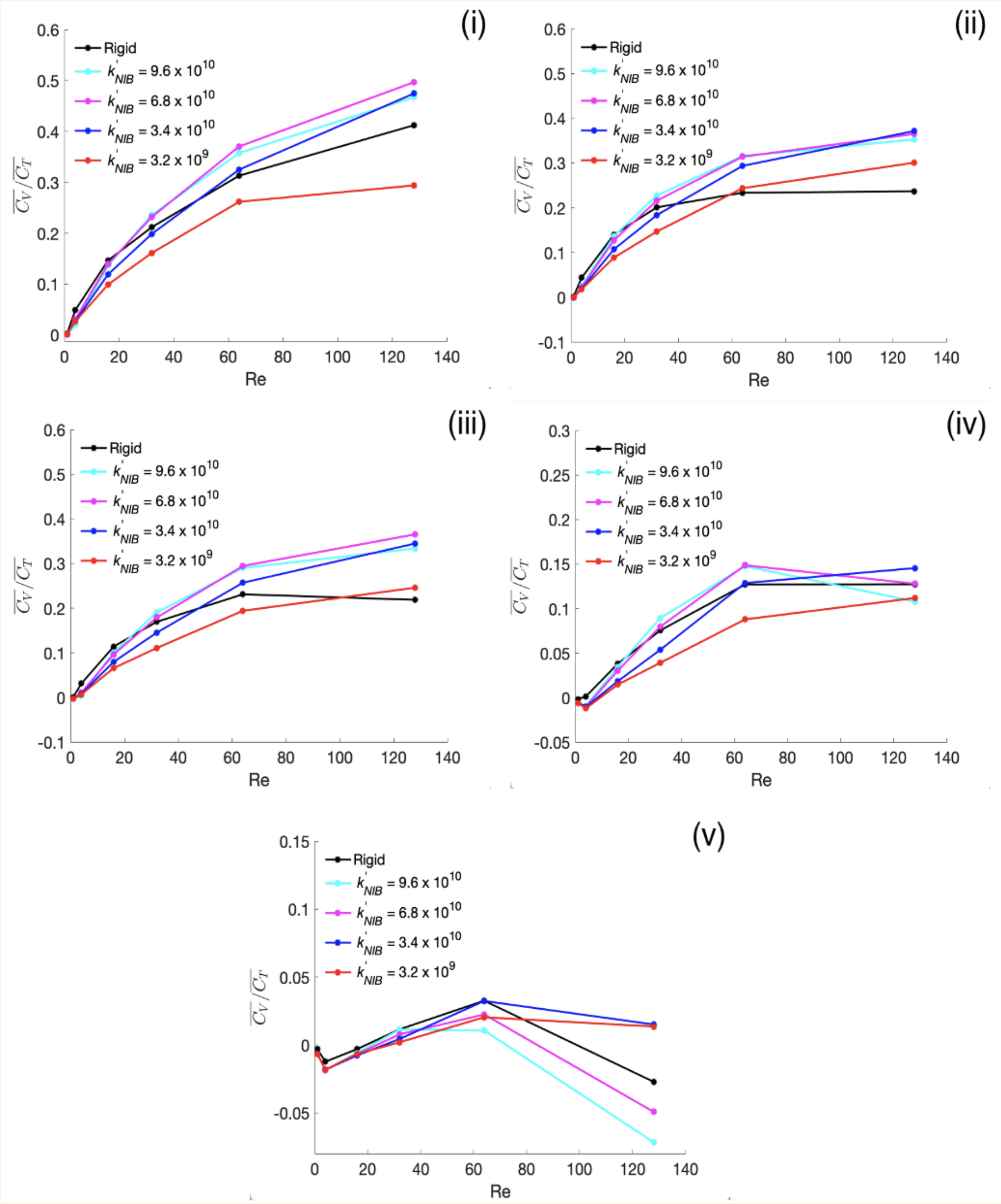


Figure 8. $\overline{C_V}/\overline{C_T}$ as a function of *Re* for all wings considered in this study for $(i) : \beta = 0°, (ii) : \beta = 22.5°, (iii) : \beta = 45°, (iv) : \beta = 77.5°, (v) : \beta = 90°$.

Variation of Pseudo-efficiency with wing flexibility:

Fig. 9 shows pseudo-efficiency (F_v/P) as a function of Re for all the wings at all stroke plane angles considered in this study. At Re > 100, for β = 0° and β = 45°, F_v/P initially decreased rapidly with increasing flexibility, and then decreased monotonically with increasing flexibility at a relatively smaller rate.

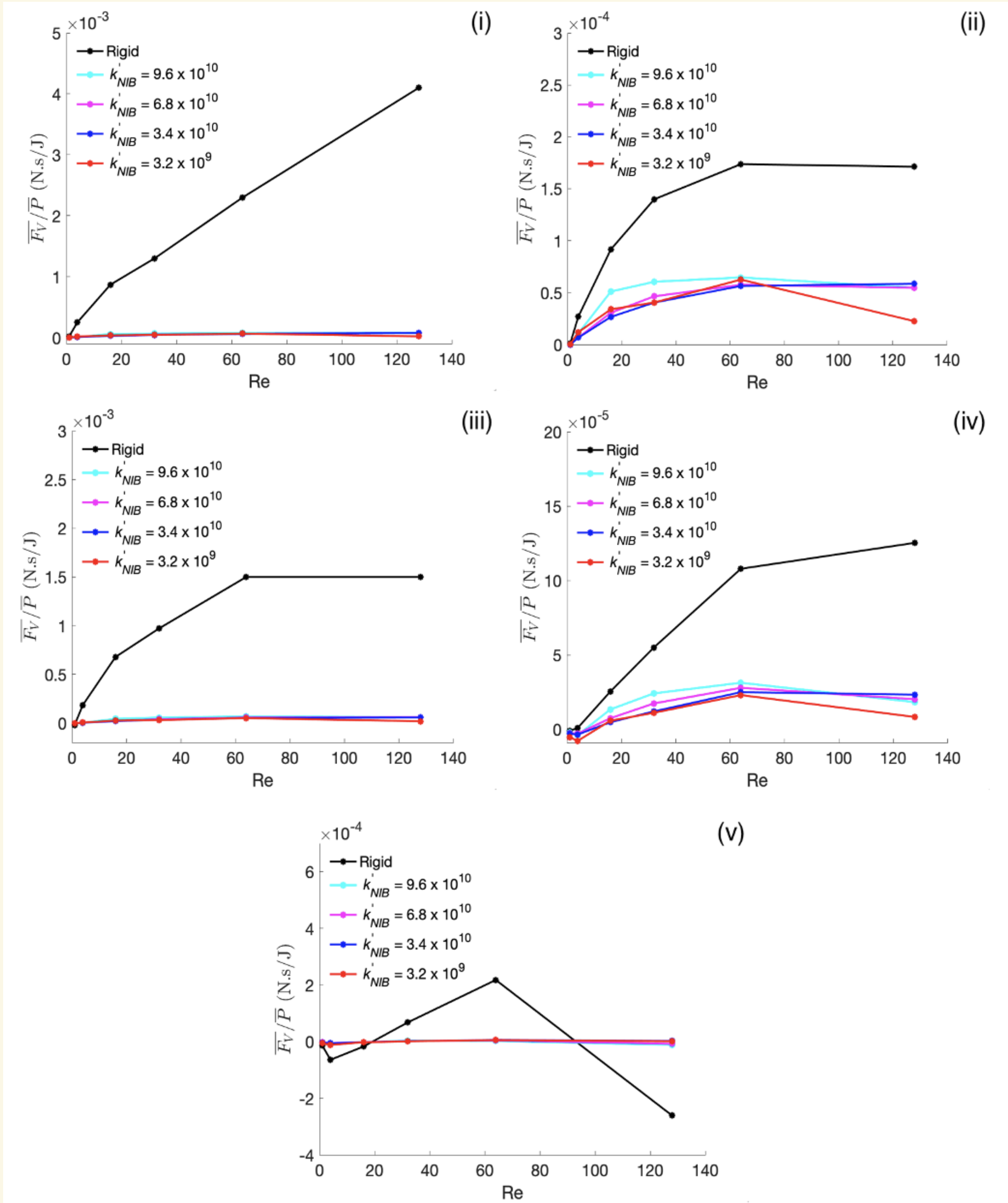


Figure 9. $\overline{F_V}/\overline{P}$ as a function of *Re* for all wings considered in this study for $(i) : \beta = 0°, (ii) : \beta = 22.5°, (iii) : \beta = 45°, (iv) : \beta = 77.5°, (v) : \beta = 90°$.

For β = 22.5°, F_v/P decreased with increasing wing flexibility until $k_{NIB}' = 6.8 \times 10^{10}$, then showed an increase for $k_{NIB}' = 3.4 \times 10^{10}$, and then decreased with further increase in flexibility. For β = 77.5°, F_v/P initially decreased with increasing wing flexibility, then increased with increasing flexibility until $k_{NIB}' = 3.4 \times 10^{10}$, and then showed a decrease for the most flexible wing. For β = 90°, F_v/P increased with increasing wing flexibility until $k_{NIB}' = 3.4 \times 10^{10}$, and then showed a slight decrease for the most flexible wing.

At Re relevant to the flight of the smallest insects, for $\beta = 0°$, $\beta = 22.5°$, and $\beta = 45°$, F_v/P decreased with increasing wing flexibility until $k_{NIB}' = 3.4 \times 10^{10}$, and then increased with increasing wing flexibility. For $\beta = 77.5°$, F_v/P decreased monotonically with increasing wing flexibility. For $\beta = 90°$, F_v/P increased with increasing wing flexibility until $k_{NIB}' = 6.8 \times 10^{10}$ and then decreased with increasing flexibility.

Discussion

The results of this study imply that: (1) at Re relevant to small insect flight, aerodynamic efficiency decreases monotonically with increasing stroke plane angle regardless of wing flexibility; (2) at Re pertinent to the flight of the smallest insects, adding flexibility to a rigid wing depreciates aerodynamic efficiency regardless of the stroke plane angle - a rigid wing is aerodynamically more efficient than flexible wings at all stroke plane angles. Implication (1) agrees with the paper of Jones et al. where they suggested that for the smallest flying insects, a lift-based mechanism generates more C_v/C_T than a drag-based mechanism (vertical stroke) (Jones et al., 2015).

Our study also took energy considerations into account. As stated in the introduction section, we defined a new dimensional quantity, pseudo-efficiency, which is the ratio of the average vertical force generated by a wing to the average power delivered by the wing to the surrounding fluid, to establish a measure of energy efficiency. The results of our study imply that: (1) at Re relevant to the flight of the smallest insects, pseudo-efficiency decreases monotonically with increasing stroke plane angle regardless of wing flexibility, (2) at Re relevant to small insect flight, adding flexibility to a rigid wing deteriorates pseudo-efficiency regardless of the stroke plane angle - a rigid wing is energetically more efficient than flexible wings at all stroke plane angles.

Flexibility has been shown to play an important role in drag reduction in biological structures (Miller and Peskin, 2004, 2009; Alben et al., 2002). Our results indicate that it plays an important role in augmenting aerodynamic efficiency in insect flight at high Re. However, our study has also shown that it negatively impacts energy efficiency in insect flight. In general, a rigid wing outperforms flexible wings in terms of energy efficiency in insect flight.

Conclusions

Our study indicates that increasing the stroke plane angle negatively impacts aerodynamic efficiency and pseudo-efficiency for all wing flexibilities and adding flexibility to a rigid wing deteriorates both its aerodynamic efficiency and pseudo-efficiency at all strokes a rigid wing deteriorates both its aerodynamic efficiency and pseudo-efficiency at all stroke plane angles in the flight of the smallest insects. A drag-based vertical stroke is inferior to both horizontal and hybrid strokes in terms of pseudo-efficiency and aerodynamic efficiency at almost all Re regardless of the wing flexibility, and this could explain why we are not aware of any small insects that use drag-based strokes at very low Re. A frequently observed trend in our results was that both aerodynamic efficiency and pseudo-efficiency increased with increasing Re. The smallest insects, however, restrict their flight to low Re. This was a simplified study that utilized 2D kinematics and symmetrical strokes. Further insight could be obtained into the flight of the smallest insects by considering additional flight strategies like bristled wings and wing-wing interactions.

Our study only analyzed aerodynamic efficiency and pseudo-efficiency as the driving factors behind the choice of flight mechanism in the smallest insects. However, there are several other factors that influence the choice of flight strategy in the smallest insects. Stability and control, maneuverability, and wing strength are some of the many factors that influence the type of flight mechanisms that the smallest insects employ. It is important to consider these factors while investigating the flight of the smallest insects in the future to gain a more complete understanding of small insect flight. The overall flight strategy is a result of compromise among many factors and not just aerodynamic efficiency and energy efficiency. Further, the wing strokes need not always be symmetrical like they were in this study. Many insects have been observed to use asymmetrical strokes (Jones et al., 2015), i.e., strokes where the upstroke and the downstroke are asymmetrical. The tiny insect Encarsia formosa, for instance, has been observed to have a faster upstroke than downstroke (Weis-Fogh, 1973). Future two-dimensional studies should implement asymmetrical strokes to arrive at more practical results. The asymmetry of strokes coupled with the vortical asymmetry presented in fig. 6 could provide more insight.

This study was primarily two-dimensional in nature, and it would certainly differ from a three-dimensional study that more closely mimics real life insect flight. A two-dimensional study fails to capture the spanwise flow effects observed in three dimensions.

These effects may lead to vortex shedding patterns and crossflow that are different from those observed in two-dimensions which, in turn, could result in the aerodynamic performance of 3-D wings being different from that of 2-D wings. Also, since our study was two-dimensional in nature, the wing flexibility was constant along the span. In future three-dimensional studies, the degree of flexibility along the wingspan could be varied to simulate real life insect wings more closely and thus, arrive at more accurate results. These studies could also consider wing-wing interactions like the clap and fling mechanism that is believed to be employed by the smallest insects (Weis-Fogh, 1973). They could also consider the impact of the presence of the insect body between the two wings on the overall flow behavior. We utilized a linear stroke plane in our study. However, the wing tips of many insects have been found to trace out patterns shaped like ovals, parabolas, or a simple figure of eight. Future work could employ these patterns to investigate flight in the smallest insects more rigorously. Extensive efforts are required in the future to completely unravel the complexities of flight in the smallest insects.

References

Wang, Z.J. Dissecting Insect Flight. Annu. Rev. Fluid Mech. 2005, 37, 183-210.

Naranjo, S.E. Assessing insect flight behavior in the laboratory: a primer on flight mill methodology and what can be learned. Annals of the Entomological Society of America 2019, 112, 182-199.

Bodino, N.; Cavalieri, V.; Dongiovanni, C.; Simonetto, A.; Saladini, M.A.; Plazio, E.; Gilioli, G.; Molinatto, G.; Saponari, M.; Bosco, D. Diseprsal of Philaneus spumarius (Hemiptera: Aphrophoridae), a vector of Xylella fastidiosa, in olive grove and meadow agroecosys tems. Environmental entomology 2021, 50, 267-279.

Cameron, P.; Wigley, P.; Charuchinda, B.; Walker, G.; Wallace, A. Farm-scale dispersal of Bacte ricera cockerelli in potato crops measured using Bt mark-capture techniques. Entomolo gia Experimentalis et Applicata 2013, 148, 161-171.

Jones, S.; Laurenza, R.; Hedrick, T.L.; Griffith, B.E.; Miller, L.A. Lift vs. drag based mechanisms for vertical force production in the smallest flying insects. Journal of Theoretical Biology 2015, 384, 105-120.

WANG, Z.z.; LIU, Y.q.; Min, S.; HUANG, J.h.; CHEN, X.x. Parasitoid wasps as effective biologi -cal control agents. Journal of Integrative Agriculture 2019, 18, 705-715.

Cook, D.; Herbert, A.; Akin, D.S.; Reed, J. Biology, crop injury, and management of thrips (Thysanoptera: Thripidae) infesting cotton seedlings in the United States. Journal of Integrated Pest Management 2011, 2, B1-B9.

Ellington, C.P. The novel aerodynamics of insect flight: applications to micro-air vehicles. Jour nal of Experimental Biology 1999, 202, 3439-3448.

Liu, H.; Ravi, S.; Kolomesnkiy, D.; Tanaka, H. Biomechanics and biomimetics in insect inspired flight systems. Philosophical Transactions of the Royal Society B: Biological Sciences 2016, 371, 201503090.

Wang, Z.J.; The role of drag in insect hovering. Journal of Experimental Biology 2004, 207, 4147-4155.

Jensen, M. Biology and physics of locust flight. III. The aerodynamics of locust flight. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 1956, 239, 511-552.

Cloupeau, M.; Devillers, J.; Devezeaux, D. Direct measurements of instantaneous lift in desert locust, comparison with Jensen's experiments on detached wings. Journal of Experimental Biology 1979, 80, 1-15.

Wilkin, P.J.; Williams, M.H. Comparison of the aerodynamic forces on a flying sphingid moth with those predicted by quasi-steady theory. Physiological zoology 1993, 66, 1015-1044.

Hollick, F. The flight of the dipterous fly Muscina stabulans Fallen. Philosophical Transac tions of the Royal Society of London. Series B, Biological Sciences 1940, 230, 357-390.

Nachtigall, W. Insects in Flight:A Glimpse Behind the Scenes in Biophysical Research; Mc Graw Hill: New York, 1973.

Demoll, R.; others. Der Flug der Insekten und der Vögel. Naturwissenschaften 1919, 7, 480-482.

Weis-Fogh, T.; Jensen, M. Biology and physics of locust flight. I. Basic principles in insect flight. A critical review. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 1956, 239, 415-458.

Von Holst, E.; Kuchemann, D. Biological and aerodynamical problems of animal flight. The Aeronautical Journal 1942, 46, 39-56.

Dudley, R.; Ellington, C. Mechanics of forward flight in bumblebees: II. Quasi-steady lift and power requirements. Journal of Experimental Biology 1990, 148, 53-88.

Vogel, S. Flight in drosophila: III. Aerodynamic characteristics of fly wing sand wing mod els. Journal of Experimental Biology 1967, 46, 431-443.

Nabawy, M.R.; Crowther, W.J. The role of the leading edge vortex in lift augmentation of steadily revolving wings: a change in perspective. Journal of the Royal Society Inter face 2017, 14, 20170159.

Eldredge, J.D.; Jones, A.R. Leading-edge vortices: mechanics and modeling. Annual Re view of Fluid Mechanics 2019.

Ellington, C.P.; van den Berg, C.; Willmott, A.P.; Thomas, A.L.R. Leading-edge vortices in insect flight. Nature 1996, 384, 626-630.

Dickinson, M.; Lehmann, F.; Sane, S. Wing Rotation and the Aerodynamic Basis of Insect Flight. Science 1999, 284, 1954-1960.

Williams, E.; Murphy, D. Wing Flapping by a Tiny Parasitoid Wasp. Bulletin of the American Physical Society 2020.

Sarig, A.; Ribak, G. To what extent can the tiny parasitoid wasps, Erectmocerus mundus, fly upwing? Journal of Applied Entomology 2021.

Burrows, M.; Dorosenko, M. Take-off mechanisms in parasitoid wasps. Journal of Experimental Biology 2017, 220, 3812-3825.

Miller, L.A.; Peskin, C.S. When vortices stick: an aerodynamic transition in tiny insect flight. J Exp Biol 2004, 207, 3073-88.

Miller, L.A., Peskin, C.S. A computational fluid dynamics study of 'clap and fling' in the smallest insects. J Expl Biol 2005, 208, 195-212.

Miller, L.A., Peskin, C.S. Flexible clap and fling in tiny insect flight. J Expl Biol 2009, 212, 3076-90.

Wang, Z.J. Two dimensional mechanism for insect hovering. Physical review letters 2000, 85, 2216.

Weis-Fogh, T. Quick estimates of flight fitness in hovering animals, including novel mechanisms for lift production. J Expl Biol 1973, 59, 169-230.

Sane, S.P. The aerodynamics of insect flight. J Expl Biol 2003, 206, 4191-208.

Weis-Fogh, T. Unusual mechanisms for the generation of lift in flying animals. Scientific Ameri can 1975, 233, 80-87.

Ellington, C.P. The aerodynamics of hovering insect flight, III. Kinematics. Philos. Trans. R. Soc. Lond. B Biol. Sci. 1984, 305, 41-78.

Horridge, G.A.; others. The flight of very small insects 1956.

Battista, N.A.; Baird, A.J.; Miller, L.A. A mathematical model and MATLAB code for mus cle-fluid-structure simulations. Integrative and comparative biology 2015, 55, 901-911.

Battista, N.A.; Strickland, W.C.; Barrett, A.; Miller, L.A. IB2d Reloaded: a more powerful Py thon and MATLAB implementation of the immersed boundary method. Math. Method. Appl. Sci 2017 https://doi.org/10.1002/mma.4708 , 1-26.

Battista, N.A.; Strickland, W.C.; Barrett, A.; Miller, L.A. IB2d Reloaded: A more powerful Python and MATLAB implementation of the immersed boundary method. Mathemati cal Methods in the Applied Sciences 2018, 41, 8455-8480.

Peskin, C.S. Flow patterns around heart valves: a numerical method. Journal of computational physics 1972, 10, 252-271.

Peskin, C.S. Numerical analysis of blood flow in the heart. Journal of computational physics 1977, 25, 220-252.

Peskin, C.S. The immersed boundary method. Acta Numerica 2002, 11, 479-517.

Griffith, B.E.; Luo, X.; McQueen, D.M.; Peskin, C.S. Simulating the fluid dynamics of natural and prosthetic heart values using the immersed boundary method. Interna tional Journal of Applied Mechanics 2009, 1, 137-177.

Taheri, A. Lagrangian coherent structure analysis of jellyfish swimming using immersed bound ary fsi simulations. J. of Mech. and Civil Engineering 2018, 15, 69-74.

Maniyeri, R.; Suh, Y.K.; Kang, S.; Kim, M.J.; Numerical study on the propulsion of a bacterial flagellum in a viscous fluid using an immersed boundary method. Computers & fluids 2012, 62, 13-24.

Childs, H. VisIt: An end-user tool for visualizing and analyzing very large data 2012.

Alben, S.; Shelley, M.; Zhang, J. Drag reduction through self-similar bending of a flexible body. Na ture 2002, 420, 479-481.

Lighthill, M.J.; On the Weis-Fogh mechanism of lift generation. Journal of Fluid Mechanics 1973, 60, 1-17.

HOW UNDERSTANDING FAMILY SYSTEMS PROMOTES THE RESPECTFUL CARE OF CHILDREN

006

Madzy LaMonica

Magda Gerber's approach to respectful care of infants and children begins with the idea of seeing infants, from the moment they are born, as whole, competent people. This approach named the RIE philosophy, or Resources for Infant Educarers has a simple core belief: respect infants by seeing them as capable participants in their world (Gerber, 2013, p. xv). The guidelines of this theory seem intuitive but in practice are often countercultural and are occasionally difficult to put into practice without significant perspective shifts. This approach has now been established as a non-profit organization with the primary goal of teaching families respectful parenting. With this approach, Magda Gerber founded the Program for Infants and Toddlers (PITC), a program for low-income and at-risk children. The PITC program inspired the framework of many future infant and toddler childcare programs. Learning technical principles and research from the Family Systems Theory related to attachment theory, conflict, and power in family dynamics might allow parents and educators to better understand and put into practice the principles of RIE. This understanding will ultimately promote more respectful care of young children.

Educaring and Attachment

Magda Gerber, an ordinary Hungarian parent, realized she was embarking on motherhood with little knowledge of child-rearing. She soon turned to Dr. Emmi Pikler, a Hungarian pediatrician who was developing groundbreaking clinical research on infants' gross motor development in orphanages. Gerber began working with Dr. Pikler, turning her attention to research, where she recognized infants' intrinsic motivation to learn motor processes and understand their world. After being forced to flee Hungary as a refugee in 1957, Gerber continued her work in Los Angeles where she adapted Dr. Pikler's research to support infants in family settings, rather than the clinical orphanages where Dr. Pikler worked. After collaborating with other early childhood professionals, she established the RIE center, where she demonstrated her approach to others.

Gerber's teachings at the RIE center outline an unconventional pathway to attachment, so exploring attachment theory offers theoretical reasoning behind RIE's approach. The key principles of attachment theory will first be discussed. Attachment theory defines the types of bonds that form between an infant and a caregiver (Warren, 2021). The ideal bond is known as secure attachment, when an infant anticipates their needs will be met from a consistent caregiver. The caregiver can achieve this attachment by repeatedly and warmly responding to the infant's cries and other communicated needs (Warren, 2021). This bond results in children who feel confident to explore their environment and grow apart from their caregiver while viewing them as an available and secure base. Children who are securely attached have better outcomes than those who have other forms of attachment (Warren, 2021). These children feel confident in the ability to separate and reunite with their parents within the safety of their parents' consistent care. This comfort children feel from consistent responsiveness may take many forms- the RIE method offers one approach.

RIE forges attachment through the practices of "educaring,", independent play, and handling conflicts between children. Educaring involves maintaining complete intimacy and attention during caregiving tasks. Examples of caregiving tasks include changing an infant's diaper, feeding, bathing, or dressing an infant (Hammond, 2021). Traditional parenting might view diaper changes as an unpleasant time in which infants are often uncomfortable or upset. To diminish children's discomfort, parents might distract children with a song or appealing toy. The parents may also rush through the process to quiet the child's discomfort as quickly as possible and to console them. A RIE parent might approach an uncomfortable baby needing to be changed in a way that leans into discomfort as a means of connection and respect; this is countercultural.

This parent would not distract, console, or rush, but instead, acknowledge the child's discomfort and join them with empathy. If the child communicates that they would rather be doing something else, such as looking at something in another room, the parent would observe the child and their interest. Then, the parent may recognize the child's interest verbally before returning an infant's attention back to the diaper changing process. Perhaps most importantly to Gerber's followers, children are not distracted by or rushed through diaper changes because they believe that it is not respecting an infant's body. It is seen as a predecessor to understanding consent to involve the child in the process of their body being touched so intimately, rather than being urged to not pay attention to what is done to their body. To facilitate this, these parents ask their baby to collaborate with them on the task of changing diapers; the parent asks the child to assist by lifting their legs or pulling open the diaper tabs. RIE parents believe this process of slowing down allows babies to feel seen, taken care of, and respected (Gerber, 2013, p.33). Secure attachment figures are responsive (Warren, 2021); therefore, when caregivers allow the baby to act as a participant in diaper changes, this responsive care fosters secure attachment. After these caregiving tasks, RIE parents are instructed to give infants ample time to play independently and entertain themselves without unnecessary adult intervention (Gonzalez-Mena, 2013, p.45). This is where parents may demonstrate to their children that they are a secure base for attachment, as the children are allowed time to explore from birth. When the child needs their caregiver again, the infant is given complete attention by the caregiver until the baby's needs are met. Understanding the complete theory behind Gerber's ideas about caregiving times and independent play may encourage parents and practitioners to apply these principles when it feels uncomfortable to face a child's discomfort during a diaper change without rushing or consoling to allow for additional time for infants to participate in caregiving. If these ideas promote more securely attached babies, and securely attached babies have better outcomes, sharing these ideals may improve the livelihoods of some children.

Infant Cries and Conflict

The conflict and communication in conflict styles, established by the Family Systems Theory, also supports RIE's philosophy surrounding tension and conflict when caring for children. The earliest conflicts with infants might begin from babies' cries. Gerber believed that babies should be allowed and encouraged to express negative feelings (Gerber, 2013, p. 7). This may feel counter to the perception many adults have of an infant's cry, as many feel it is combative, stressful, something to fear, and something that is the adult's ultimate responsibility to fix or stop.

This perspective can be shifted by understanding the theory of pluralistic families, which are those who are able to deal with conflict preemptively, use direct communication, and do not shy away from conflict with other family members (Segrin & Flora, 2019). Accepting and not feeling alarmed by a baby's cry is part of modeling these pluralistic conflict patterns. The RIE method attempts to not only accept cries, but also to be proactive in reducing incidents of a child's discomfort by providing a consistent and predictable routine (Gonzalez-Mena, 2013, p.44). This conveys the pluralistic conflict pattern of preemptively coping with conflict. Viewing an infant's cry as a potential miscommunication from caregivers, such as missing a step in a routine, is also an example of communal coping. For example, RIE parents are encouraged to tell a baby that they will be picking them up before they do so, especially if the baby is currently focused on another activity, like a toy. If a parent misses this step in routine and picks the child up without forewarning, the child may cry. Then, rather than immediately seeking to pacify the child, the caregiver might provide empathy for the child by not avoiding this moment of conflict and seeing the cries as communication. The parent might then apologize to the child for missing the step in the routine, fulfilling the direct communication requirement of the pluralistic conflict model.

Communal coping, another model of familial coping in the Family Systems Theory, is a highly effective coping method in which families work to solve problems as a group (Segrin & Flora, 2019). Utilizing and modeling communal coping is advantageous, as strong family coping methods are seen as resources in the ABC-X model of family stress (Segrin & Flora, 2019). This model measures the number of resources a family has to cope with stress. Some of the most important resources a family can have in this model are effective and appropriate family conflict styles and communication styles. The more effective resources a family has under this model, the more likely they are to effectively cope with a stressor (Segrin & Flora, 2019). Communal coping may help parents shift their perspectives on how to handle a crying infant. In the example of the crying infant that was picked up unexpectedly, the caregiver would work with their infant to determine the source of the cry, reflecting on the baby's state of needs as well as their own actions. This communal approach reflects how the caregiver sees an infant's cries as reflective of how the environment might affect them, rather than only the infant's internal state of need. Teaching families how to use effective conflict strategies, such as communal coping and the pluralistic conflict style, and communicating how these approaches are stronger coping methods, might allow families to more effectively move through conflict and stress. The RIE method provides clear strategies for communal coping and pluralistic conflict styles. As the RIE method is currently not widely understood or known, teaching families about these conflict and communication styles from the Family Systems Theory might help them to better understand the intricacies of the RIE approach. In turn, this may help families to implement the respectful parenting strategies of RIE into their own homes.

Power in Work with Children

RIE discourages the use of conventional dynamics of power with children, such as using rewards and punishments. It may be helpful for families to understand the theories surrounding the dynamics of power in the Family Systems Theory to support RIE's alternative disciplinary measures. Power dynamics expressed in a child and caregiver relationship are an important part of secure attachments, as power consists of the capability one has to change another's behavior and its strength is determined by how dependent one is on the other (Segrin & Flora, 2019). With the RIE method, parents are encouraged to influence their children's behavior without offering rewards or praise. Instead, a parent might state, in a neutral tone, that the child's behavior is appreciated (Gerber, 2013, p. 11). This may be because reward power, or trying to change another's actions by providing a reward, is only effective when the reward is actively wanted (Segrin & Flora, 2019). For reward power to work, the recipient, the child, must want the reward each time, which is challenging to predict and may be unreliable. Initial rewards become tiring and no longer wanted. Then, rewards need to be continually changed or increased to be motivating. RIE parents also do not use coercive power, which consists of trying to change one's behavior by threatening punishments (Segrin & Flora, 2019). Gerber believed when parents become angry or beg their children to comply, children feel that they are too powerful, which creates fear in them that they are not safe and secure, thus damaging attachment (Gerber, 2013, p.11). For this reason, RIE practitioners use referent power, which asserts that one has power based on the role one plays within that dynamic (Segrin & Flora). This type of power may be demonstrated by telling a child that they will be stopped from playing with something unsafe and explaining why they should not. In this way, the child understands over time that it is the caregiver's role to stop them from doing things that are unsafe or that may hurt others. They also recognize that they are not responsible for maintaining their own safety, and they are confident in continuing to view their parents as their secure base. It may occasionally be difficult to not rely on coercive and reward power as they may seem easier than referent power in moments where caring for children is difficult. It is often quicker to offer a reward, such as candy, to influence a child to do something than to face conflict in the pluralistic way and stop a child from doing something with referent power. As such, understanding power dynamics is important to explain Gerber's disagreement with rewards and punishments and may help caregivers to rely on referent power when children need boundaries.

Positive Parenting and Outcomes for Children

Applying a positive parenting approach like the RIE method may help families to break cycles of generational parenting patterns and to improve childhood outcomes. Poverty may be a factor that increases the likelihood of childhood mistreatment according to the Family Stress Model, or FSM (Masarik & Conger, 2017). This is understandable when contextualizing that poverty causes increased stress on parents. Families living in poverty may have to leave their children unsupervised more often because they must work outside of the home, and parents may not have experience and training in child development or child rearing. Despite this risk factor for negative parenting, Masarik and Conger (2017) outlined the following about a study of families in poverty:

There is also support in the FSM literature that positive or adaptive parenting practices are linked to child wellbeing, even in the face of economic stress. For instance, parents who engaged in warm and supportive behaviors toward their child were more likely to have preschoolers who were securely attached and engaged in self-regulatory behaviors, which in turn, predicted better cognitive outcomes in first grade. In brief, several researchers noted that positive parenting behaviors were associated with child and adolescent increases in mastery, prosocial behavior, optimism, and healthy eating behaviors as well as reductions in internalizing symptoms, delinquency, and risky health behaviors. (pp. 3-4)

When parents can engage in positive, warm, and supportive relationships with their children, despite the extreme familial stress of poverty, children experience better outcomes. Perhaps establishing training in positive parenting methods for parents who may have very high stress measures in the FSM may reduce the effect of poverty on children's outcomes. Families in poverty are often stuck in multi-generational cycles due to systemic societal factors (Cherlin, 2018). The multi-generational nature of this family stress may result in several generations of challenging parent-child relationships. While more research is needed on both the effects of positive parenting on these risk factors and on potential interventions to encourage positive parenting, it seems worthwhile to discuss the potential of positive parenting to reduce the negative effects of family stress in childhood.

The RIE method teaches parents to respond to their children in ways that are sensitive and work to understand their point of view. For a family without stress, these approaches promote more respectful relationships. Masarik and Conger's (2017) research might indicate that sharing the RIE method may also help families with stress mitigate the effects stress may have on their future development. The research may also indicate the impact positive parenting may have on families with less stress on improving parent-child relationships, family dynamics, and childhood outcomes.

Conclusion

Studying family dynamics and relational factors such as attachment theory, conflict, and power offer tools to describe to parents, families, and practitioners why the RIE principles are important and worthwhile parenting approaches. If more families can be effectively educated in the RIE method, it may help children to feel secure in themselves through secure attachment, eliminate inappropriate power dynamics with referent power, improve a family's process of coping with conflict, and might have the potential to reduce the effect of adverse experiences. It may be worth creating accessible training resources that teach the RIE method that contextualize RIE through theory, such as Family Systems Theory, because of its emphasis on respectful care. When children are cared for respectfully, their caregivers are modeling a code of behavior that children may repeat with their peers, other authority figures, and with themselves. RIE places particular attention to considering the perspective of the child and allowing the child to be present with their emotions rather than pacifying them. This issue of pacification seems particularly relevant in a world that discusses the way in which individuals often cope through distraction with things like technology or food, rather than facing conflict and negative emotions. RIE 's promotion of a child's bodily autonomy and involvement of children in their own care when bodily consent cannot be maintained teaches the child that they have agency over their own body. These early teachings of consent seem relevant in society that grapples with the nature of consent, how to conceptualize it for children, and the consequences of not properly communicating its nature. Finally, RIE 's disciplinary approach of facing conflict rather than ending them quickly with rewards or incentives may introduce an alternative to instant gratification, which has been discussed in popular culture as a problematic and limited motivational force. While these factors require empirical inquiry as they relate to RIE, and cannot be related causally, it seems notable to begin approaching each of these issues proactively at each stage of education and childcare. The foundation children receive from their parents in early childhood, of course, has a lasting effect on their life beyond their parent's care. If shifting a parent's approach to parenting may have a more positive and respectful impact on their children, it may be worth establishing research, training, and intervention programs to further promote such positive parenting models.

References

Cherlin, A. J. (2018). How Inequality Drives Family Formation. Unequal Family Lives, 69–82. https://doi.org/10.1017/9781108235525.004

Gerber, M. (2013). The RIE manual: for parents and professionals (2nd ed.). Resources for Inant Educarers

Gonzalez-Mena, J. (2013). Ten easy steps to the demonstration infant program (DIP) philosphy. In M. Gerber (Ed.), The RIE Manual: For parents and professionals (2nd ed., pp.43-45) Resources for Infant Educarers.

Hammond, R. A. (2021). Seeing infants differently: Magda Gerber's contributions to the early care and education field and their continuing relevance. Early Child Development and Care, 1–14. https://doi.org/10.1080/03004430.2020.1865942

Masarik, A. S., & Conger, R. D. (2017). Stress and child development: a review of the Faily Stress Model. Current Opinion in Psychology, 13, 85–90. https://doi.org/10.1016/j.copsyc.2016.05.008

Segrin, Chris, and Jeanne Flora. Family Communication, Taylor & Francis Group, 2018. ProQuest Ebook Central, http://ebookcentral.proquest.com/lib/uaz/detail.action?docID=5614766.

Warren, S. (2021). Attachment theory.