Machine Learning in The Diagnosis of Mental Disorders

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Tyler Rasmussen

Summer, 2024

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Bryn E. Seabrook, Department of Engineering and Society

STS Research Paper

Introduction

According to the Center for Disease Control (CDC), suicide rates have risen by 37% from 2000 to 2018 and more than 1 in 5 U.S. adults will experience mental illness (CDC, 2023). Despite the ongoing mental health crisis, there exists a lack of mental health care for many individuals due to barriers to treatment and diagnosis that prevent treatment (Rivers et al. 2024 Introduction). It is clear that there exists a failure to address this mental health crisis. The use of machine learning (ML) to diagnose mental disorders, while a nascent area of research, shows promising results in meeting these needs. However, it is unclear the extent which machine learning in mental disorder diagnosis will affect society. Latour's (1992) actor-network theory will be used as a framework to examine the effects ML could have on society.

Methods

The primary methods of data collection consist of documentary research using library search engines such as Libra as well as data collected from government and professional websites. The data is then analyzed using both documentary analysis and network-analysis following actor-network theory. The data is then explained for understanding with various patterns and camps of thought discussed. Once technical understanding of the topic is established, actor-network theory is then used in order to draw conclusions about the effects of machine learning in diagnosing mental disorders.

Background of Machine learning and Mental Diagnosis

Machine Learning according to IBM (2024) is "a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to enable AI to imitate the way that humans learn, gradually improving its accuracy." There are many different types of algorithms used in order to train a machine learning model that can be classified as supervised or unsupervised. Supervised learning requires humans to label the data that is fed to the algorithm. Unsupervised learning does not.require the data to be labeled and allows the algorithm to classify and select patterns on its own. Both types of machine learning algorithms can be used in diagnosing psychiatric disorders. A variety of data collection methods such as questionnaires, electroencephalography (EEG), functional magnetic resonance brain scans (fMRI), and gut microbiome sampling can be used effectively with machine learning in healthcare.

Currently, psychiatric diagnoses are done by doctors using guidance from the Diagnostic and Statistical Manual of Mental Disorders (DSM) based on patient symptoms. This system leaves much to be desired in terms of accuracy with in one study investigating misdiagnosis, "rates reached 65.9% for major depressive disorder, 92.7% for bipolar disorder, 85.8% for panic disorder, 71.0% for generalized anxiety disorder, and 97.8% for social anxiety disorder" (Vermani et al., 2011, p. 2). This results in increased costs and poorer patient prognosis, not just for the patient. Insurance companies and hospitals often have to share the costs of misdiagnosis and pass them on as increased rates and billing. Thus the current methods of providing diagnosis of mental disorders are inadequate to contend with the mental health crisis..

As mentioned in the Introduction, it can be said we are currently experiencing a mental health crisis. Suicide rates are showing a staggering increase as shown in Figure 2. Furthermore, it is estimated that more than 1 in 5 U.S. adults will experience mental illness (CDC, 2023) and 6

in 100 people have experienced trauma that could cause Post Traumatic Stress Disorder (Wang et al. 2023 P. 1655). Despite the urgent need for mental healthcare, there exists many barriers to accessing such care as shown in figure 1. The most common of these barriers are lack of time and financial cost. It is imperative that solutions be found in order to provide the patients with the healthcare they need to alleviate the current mental health crisis.

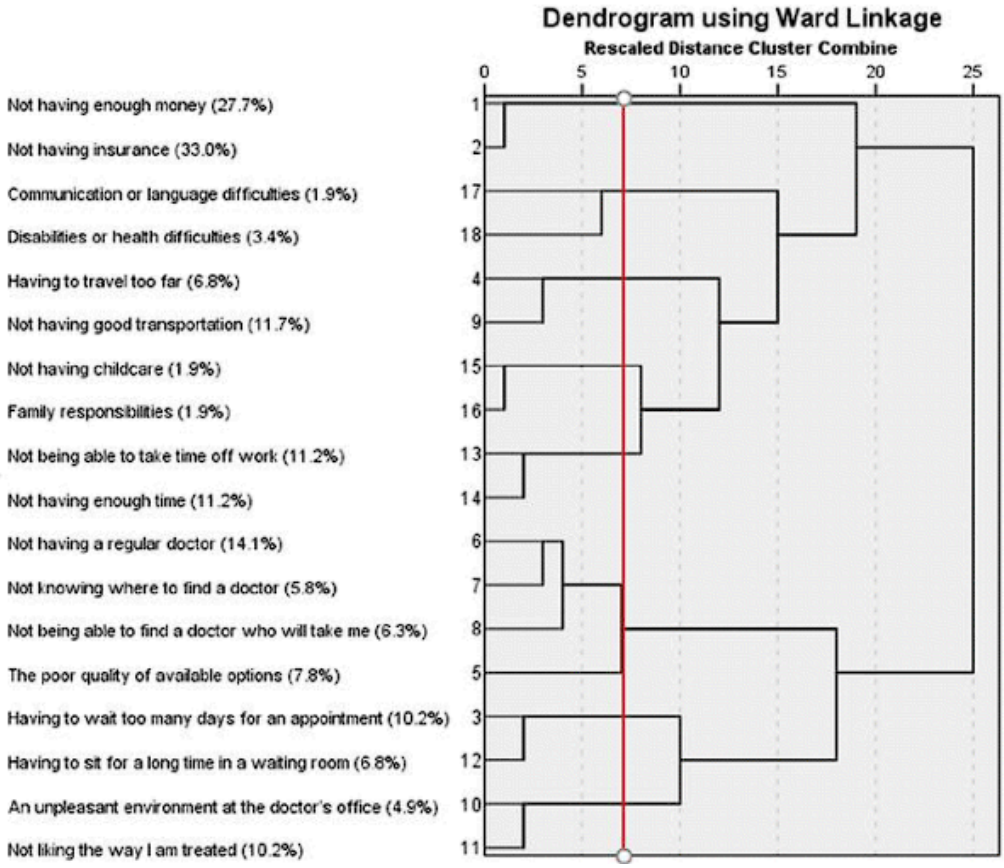

Figure 1. Dendrogram of Percentage of Barriers to Healthcare. Note that finances and time being the most commonly cited barriers. Adapted From Rivers et al, p. 5)

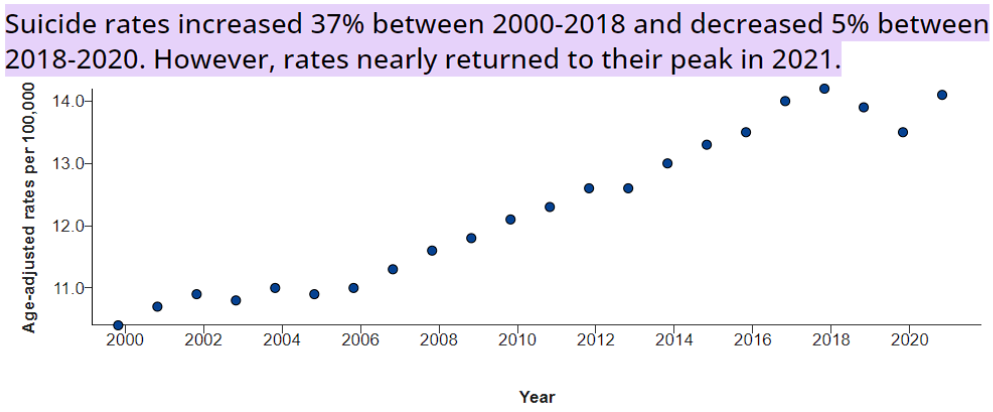

Figure 2. Graph of suicide rates from 2000-2018. Rates increased steadily from 2005 up until 2018. (CDC, 2023)

Actor-Network Theory and Machine Learning

Actor-network theory provides a useful and insightful method of analysis into how machine learning use in psychological diagnosis could affect society. Created by Bruno Latour (1992), it allows STS scholars to analyze different sociotechnical systems by looking at the relationships between different actors. These actors can be anything from human to nonhuman or from tangible to intangible. One of Latour's examples of the mechanical door closer includes the actors of human doorman, mechanical door closer, patrons entering the door, and the idea that the door should be closed and humans cannot be relied upon to close it. Through the analysis of how these actors interacted Latour was able to thoroughly examine how this new door closing technology affected society.

The effectiveness of this approach makes it ideal for analyzing the introduction of ML in the diagnosis of mental disorders. However, current literature analyzing machine learning in the medical field using ANT is lacking which is not surprising due to the nascent nature of ML in

medicine. However, there does exist literature analyzing artificial intelligence and machine learning using ANT. Notably Tommaso Venturini, a former student of Bruno Latour, the creator of ANT, writes a compelling article on AI under the lens of ANT (Venturini, 2024). There also exists other areas of STS research that use actor network theory to analyze ML in application of specific areas such as in education (Sperling, 2022). Overall, there appears to be a distinct lack of literature using ANT to evaluate the introduction of AI and ML into different parts of society.

Venturini (2024) draws from a multitude of different sources that the late Bruno Latour authored on AI and its effects on society using an ANT approach. His own experiences as a student of Latour as well as his expertise in STS are used to show what effects ANT could predict AI would have on society. He notes that Latour had a "relative indifference" to AI as he viewed it as more of a distraction from more important issues like the ecological crisis. Using an actor-network approach, Venturini analyzes the potential positive and negative effects of AI as well as the dangers of how society could react to AI as well as some of the benefits society could see. Whilst the article uses a Latourian stream of consciousness style of thought when discussing AI and society which negatively impacts the brevity and conciseness of the article, it does offer an excellent perspective of AI on society. Venturini discusses how generative AI can be a useful tool for writing as a "less obedient typewriter" and that AI as a whole should not be viewed as extraordinary but just another technology. Actor-network theory shows that AI requires a vast network of actors from users to researchers, artists, writers, and programmers to provide data for the AI model to use and then the variety of tasks that users may find for the model (Venturini, 2024).

One study looking at machine learning in Swedish Primary Education using actor-network theory to examine how the introduction of an AI engine into the classroom to help

teach mathematics created a network of actors and what could be learned from it (Sperling et al., 2024). Surprisingly, it was found that the AI did not replace any human labor in teaching but rather required students and teachers to adapt to this new AI actor. The AI was sometimes wrong in its answers and calculations so teachers and students would have to identify this and conduct small class discussions on why the AI answer was a mistake. Whilst many believe AI is predicted to have a transformative role in education, currently the technology does not have the effect of bypassing teacher capacity and reducing costs as expected. Nonetheless, a useful insight into the various actors that surround AI in education and how they must adapt in order for the successful integration of AI into education could be acquired.

It seems that STS scholars believe that AI should not be the cause for any extraordinary alarm. Viewed under ANT, it appears very much to be just another technology, different as all technologies are but not so extraordinary to warrant special attention. Scholars believe that it is essential that a focus on critical thinking be achieved in order to fully take advantage of AI and function in a world with generative AI. These same ideas remain relevant in the introduction of machine learning into mental disorder diagnosis.

Results and Discussion

Machine learning in psychological diagnosis is currently in its infancy and yet to be implemented in clinical practice. However, a multitude of studies show ML has immense potential to improve healthcare outcomes for those suffering from mental illness. Whilst the effects on patients are most notable, ANT analysis shows that other actors such as hospitals, insurance companies, research institutions, and patient's communities also benefit. It is clear that currently ML has the potential to improve society. However, society must be careful to avoid or

overcome several barriers that would prevent the effective use of ML in psychiatry and psychology. Successfully integrating ML into mental health diagnosis will help solve the mental health crisis.

Effectiveness of machine learning in psychology

There are many studies on the efficacy of using machine learning models in order to diagnose patients, our actor of foremost concern. Most, if not all of these studies show promising results. Computers, an essential technical actor, are able to find patterns that humans have trouble spotting such as those in fMRI brain scans (Huang et al., 2024, p. 1) or readings from EEG sensors (Othmani et al., 2024, p. 24135), or even the gut microbiome (Stiernborg et al., 2024, p. 1). Moreover, ML algorithms can also assist with standard diagnostic questionnaires and classify mental disorders based on this data much like doctors due currently under the DSM. There exists a wide variety of data streams and applications for ML to assist doctors in diagnosing mental disorders.

Machine learning techniques used in psychology are varied from simple regression models to more complex neural networks (Othmani et al., 2023, p. 24135). There is no such thing as a single machine learning algorithm that can fit every problem in psychology (Hawes et al., 2023, p. 3). More research is needed to determine the best algorithms for every diagnosis but among the more promising models are random forest trees, recursive ensemble feature selection, and Gaussian process regression (Peralta-Marzal et al., 2024, results). One study was able to train a machine learning model to classify Bipolar Disorder (BPD) and Major Depression Disorder (MDD) with an accuracy of 90.3% (Huang et al., 2024, p. 1) using data from fMRI brain imaging techniques shown in figure 2. Another study was able to train a decision tree

model to predict ADHD in youths with a predictive accuracy of 74% (Grazioli et al., 2024, p. 5). A third study was able to train a machine learning algorithm to predict young Schizophrenia disorder patients with an accuracy greater than or equal to 70% using data of the patient gut microbiomes (Stiernborg et al., 2024, p. 1). However, due to the current limitations of the data pool studies have drawn from, it remains to be seen if these models can be scaled to deal with greater amounts of data and patients (Portugal et al., 2023, conclusion).

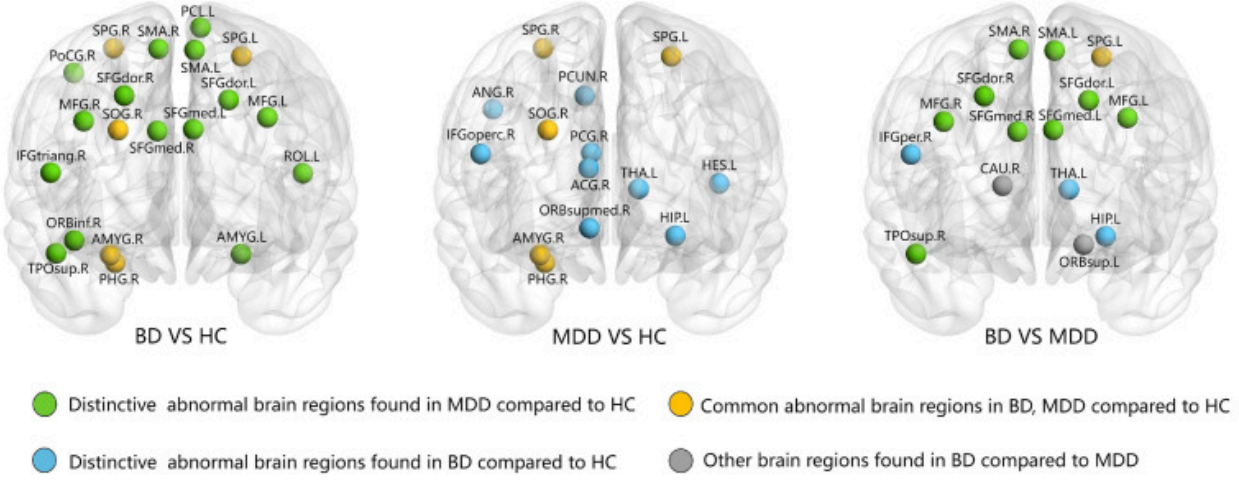

Figure 2. Graph of nodes in brain regions that signify BPD or MDD. Full description of acronyms can be located in the references. (Huang et al., 2024, 3.3. Alterations in Nodal Brain Network Properties).

It is clear that in many cases it may be possible for machine learning systems to have greater accuracy in diagnosing mental disorders than current doctors using the Diagnostic and Statistical Manual of Mental Disorders (DSM). While it is difficult to say the extent that computers can take over this diagnostic work due to limited sample sizes in data, the possibility that they could greatly assist doctors in diagnosing should not be underestimated. While it is unlikely that computers will completely replace doctors in diagnosing mental disorders, they will be able to assist in other ways. Doctors would be able to read computer given reports much like how radiologists assess reports given to them by their machines. The DSM could integrate certain machine learning models and provide guidance on diagnosing using certain metrics such

as how accurate the computer believes its diagnosis to be. Doctors could also provide questionnaires for patients to fill out and machine learning algorithms could flag patients most likely to have certain disorders such as ADHD. (Grazioli, et al., 2024 p. 2).

Effect on Actors

Patients are the most obvious actors in machine learning however other stakeholders such as hospitals, insurance companies, research institutions, and the communities of patients also feel the impacts of improved mental healthcare. These organizational actors must be recognized as well to fully consider the impacts this ML technology can have on society. The most important cultural actor is trust in technology, specifically artificial intelligence (Sperling et al., 2022). Other cultural actors such as trust in doctors and a desire to treat mental disorders whilst important are not unique to ML technology. The foremost technical actor is machine learning and AI. Other secondary technical actors include fMRI, ECG, gut microbiomes, and computers. All of these actors must be considered in order to properly assess the validity of machine learning to diagnose mental disorders.

As mentioned in the introduction, there are major barriers to healthcare mainly in the time and financial cost to receive care (Rivers et al., 2024). The introduction of ML Diagnosis aims to reduce these barriers for stakeholders by freeing up doctor's time and cutting down on costly misdiagnosis by increasing diagnostic accuracy. While this would benefit both patients and doctors allowing doctors to work more efficiently and patients to receive better treatment, these effects would also ripple out to other organizational actors. Insurance companies would be able to give better and more accurate rates due to the reduction of misdiagnosis and reduced treatment costs. Hospitals would be able to more efficiently use resources that previously needed

to be spent on diagnosing mental health patients and may be able to quickly diagnose patients going through a mental health crisis. The communities, friends, family, neighbors, coworkers, etc. of patients stand to have a reduced burden due to better treatment of the patient lessening negative symptoms.

These benefits will only be able to be seen if society can overcome the barriers to implementing ML diagnosis. There is a distinct lack of trust in AI from cries of robots and AI automation taking jobs, ChatGPT being the latest advancement, to Hollywood movies of AI taking over humanity like Skynet in Terminator. However, doctors and patients are right to be wary of AI as it is possible that a poorly implemented diagnostic model could harm patients with inaccurate diagnosis. It is thus important that these models be as transparent as possible in order to win the trust of doctors and patients. It is not such an easy task however due to the black box nature of machine learning algorithms. Researchers often do not know exactly how a model makes its decisions which can make it difficult to inspire trust. It is thus important that these models be tested rigorously and on large and varied datasets in order to distill trust from reliable results.

Future of Machine learning in psychology

Whilst it is mentioned above that machine learning in psychology has been limited to research and experiments It does not have to be the case forever. Typical timelines for drug discovery to clinical use can take up to 12 years. Different treatment therapies such as gene therapy can have an even greater timeline of up to 30 years (Biostock, 2023). However, due to the relatively straightforward nature of classification problems in AI, the model can only predict a mental disorder or predict no disorder. There are no multitude of possible side effects and

health concerns as with new drugs and gene therapies. It is therefore unlikely that the process of trials to clinical practice would take as long. Ultimately it is unclear what timeline machine learning assisted diagnosis could take as one cannot predict the future, however it should not be surprising if machine learning is in our doctor's offices in the near future.

Additional research will be needed in order to increase the accuracy and precision of ML assisted diagnosis to be able to ethically be used on patients. The ML diagnosis must be on par with current diagnosis in order for it to be considered for approval. Moreover, a concentrated effort must be made to win over trust from doctors and patients in this technology. Trust in AI remains a problematic actor in the implementation of ML diagnosis. It is important that patients, doctors, hospitals, and insurance companies trust the effectiveness of the ML AI models. As evidenced in Sperling (2022) without trust in the technology users, teachers and students in this case, do not use the technology for its intended purpose. Students and teachers treated the technology as a novelty but did not trust its answers and made sure to give lessons that did not rely on the accuracy of the model. If this trust can be established and ML models be trained to the required precision and accuracy

Conclusion

Machine learning in mental disorder diagnosis whilst not yet being used in clinical practice is quickly approaching more widespread adoption in medicine. This technology stands to be able to reduce common barriers to healthcare such as time and financial cost. However, there in turn exist both social and technical barriers to ML Diagnosis that must be first overcome. The successful introduction of ML Diagnosis to clinical practice will bring us one step closer to solving the current mental health crisis.

References:

CDC. (2023, November 29). *Suicide Data and Statistics | Suicide Prevention | CDC*. CDC.gov. Retrieved May 1, 2024, from https://www.cdc.gov/suicide/suicide-data-statistics.html#:~:text=Suicide%20rates%20increased%2037%25%20between,to%20their%20peak%20in%202021.

Biostock. (2023 January 2). *Drug development - the four stages | Biostock. Biostock.se. Retrieved July 20, 2024, from https://www.biostock.se/en/2023/01/drug-development-the-four-phases/*

Grazioli, S., Crippa, A., Rosi, E., Candelieri, A., Ceccarelli, S. B., Mauri, M., ... Nobile, M. (2024, January 1). Exploring telediagnostic procedures in child neuropsychiatry: addressing ADHD diagnosis and autism symptoms through supervised machine learning. *European Child & Adolescent Psychiatry, 33*(1), 139-149.

Hawes, M. T., Schwartz, H. A., Son, Y., & Klein, D. N. (2023, October 1). Predicting adolescent depression and anxiety from multi-wave longitudinal data using machine learning. *Psychological Medicine, 53*(13), 6205-6211.

Huang, Y., Zhang, J., He, K., Mo, X., Yu, R., Min, J., ... Liu, M. (2024, February 15). Innovative neuroimaging biomarker distinction of major depressive disorder and bipolar disorder through structural connectome analysis and machine learning models. *Diagnostics (2075-4418), 14*(4), 389-404.

IBM. (2024), *What is ML? | machine-learning | topics | IBM.* IBM.com Retrieved June 30, 2024, from https://www.ibm.com/topics/machine-learning

Latour, B. (1992). Where are the missing masses? The sociology of a few mundane artifacts. In Bijker, W. E. and Law, J., eds. *Shaping technology/Building society: Studies in sociotechnical change.* Cambridge, MA: MIT Press, pp. 225-258

Othmani, A., Brahem, B., Haddou, Y., & Mustaqeem, . (2023, October 15). Machine-learning-based approaches for post-traumatic stress disorder diagnosis using video and EEG sensors: A review. *IEEE Sensors Journal, 23*(20), 24135-24151.

Peralta-Marzal, L. N., Rojas-Velazquez, D., Rigters, D., Prince, N., Garssen, J., Kraneveld, A. D., ... Lopez-Rincon, A. (2024, January 8). A robust microbiome signature for autism spectrum disorder across different studies using machine learning. *Scientific Reports, 14*(1), 1-11.

Portugal, L. C. L., Ramos, T. C., Fernandes, O., Bastos, A. F., Campos, B., Mendlowicz, M. V., ... De Oliveira, L. (2023, October 5). Machine learning applied to fMRI patterns of brain activation in response to mutilation pictures predicts PTSD symptoms. *BMC Psychiatry, 23*(1), 1-13.

Rivers, A. S., Sanford, K., Brewington, M., & Buchanan, A. (2024, January 1). Development of a new measure of encounters with health care barriers: The commonly experienced health care barriers index. Psychological Assessment, 36(1), 30 - 40.

Sperling, K., Stenliden, L., Nissen, J., & Heintz, F. (2022, December 1). Still W(AI)ting for the automation of teaching: an exploration of machine learning in Swedish primary education using actor-network theory. European Journal of Education, 57(4), 584 - 600.

Stiernborg, M., Prast-Nielsen, S., Melas, P., Skott, M., Millischer, V., Boulund, F., ... Lavebratt, C. (2024, January 1). Differences in the gut microbiome of young adults with schizophrenia spectrum disorder: using machine learning to distinguish cases from controls. *Brain, Behavior, and Immunity, 117*, 298-309.

Venturini, T. (2024, January 1). Bruno Latour and artificial intelligence. Tecnoscienza, 14(2), 101 - 114.

Vermani, M., Marcus, M., & Katzman, M. A. (2011). Rates of detection of mood and anxiety disorders in primary care: a descriptive, cross-sectional study. *The primary care companion for CNS disorders, 13*(2), PCC.10m01013. https://doi.org/10.4088/PCC.10m01013.

Wang, X., Yan, C., Yang, P., Xia, Z., Cai, X., Wang, Y., ... Chan, R. C.K. (2024, March 1). Unveiling the potential of machine learning in schizophrenia diagnosis: A meta-analytic study of task-based neuroimaging data. *Psychiatry & Clinical Neurosciences, 78*(3), 157-168.