INVESTIGATING THE VALIDATION, RELIABILITY AND USABILITY OF AUTOMATED METHOD FOR TESTING AUDITORY SENSITIVITY (AMTAS) WITH MALAY INSTRUCTIONAL VIDEO IN CLINICAL AND NON-CLINICAL SETTINGS

NOR HIDAYAH BINTI MOHAMMED HATTA

UNIVERSITI SAINS MALAYSIA

2025

INVESTIGATING THE VALIDATION, RELIABILITY AND USABILITY OF AUTOMATED METHOD FOR TESTING AUDITORY SENSITIVITY (AMTAS) WITH MALAY INSTRUCTIONAL VIDEO IN CLINICAL AND NON-CLINICAL SETTINGS

by

NOR HIDAYAH BINTI MOHAMMED HATTA

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

August 2025

ACKNOWLEDGEMENT

Alhamdulillah, all praises to Allah S.W.T, the Most Gracious and Most Merciful. Without His guidance and mercy, this achievement would not be possible. First, my deepest gratitude to my main supervisor, Dr. Mohd Fadzil Nor bin Rashid, for his unwavering support, invaluable guidance, and encouragement. His dedication and insightful advice have shaped this dissertation and my academic growth. I am thankful to him for securing the research grant from Universiti Sains Malaysia through the JEPeM-USM funding, with study protocol code USM/JEPeM/KK/23080602. I also sincerely appreciate my co-supervisors, Professor Dr. Mohd Normani bin Zakaria and Dr. Wan Najibah binti Wan Mohamad, for their invaluable advice, constructive feedback, and support. My gratitude extends to the hospital directors of Hospital Tuanku Ja'afar Seremban and Hospital Rembau for approving my research. I would also like to thank the Otorhinolaryngology Department staff of both hospitals for their cooperation. Special thanks to my audiology colleagues for easing my journey in balancing work and studies. I am truly grateful to my beloved husband, my six children, parents, and in-laws for their patience and unwavering support, both emotionally and financially. Their encouragement and sacrifices have been the foundation of this success, allowing me to balance my responsibilities as a wife, mother, part-time student, and working professionally. This achievement is as much yours as it is mine. Lastly, I appreciate everyone who has contributed to this dissertation, especially the research participants for their time and willingness to take part in this study. Thank you from the bottom of my heart. Above all, Syukran Ya Rabb for granting me the time and knowledge to complete this thesis. Syukur, Amin.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

AMTAS	Automated Method for Testing Auditory Sensitivity
ANSI	American National Standards Institute
ASHA	American Speech-Language Hearing Association
BSA	British Society of Audiology
dB	Decibel
DOSM	Department of Statistics Malaysia
FVI	Face Validation Index
GSI	Grason-Stadler Incorporated
HR	Hospital Rembau
HTJS	Hospital Tuanku Ja'afar Seremban
Hz	Hertz
IKU	Institute for Public Health
MANSA	Malaysian National Society of Audiologists
MAUQ	mHealth app Usability Questionnaire
mHealth app	Mobile Health Application
M-MAUQ	Malay version of mHealth app Usability Questionnaire
MOH	Ministry of Health
MPANL	Maximum Permissible Ambient Noise Levels
NIH	National Institute of Health
PTA	Pure-Tone Audiometry
SoLLaT	School of Languages, Literacies and Translation
TAM	Technology Acceptance Model
WHO	World Health Organization

LIST OF APPENDICES

KAJIAN KESAHAN, KEBOLEHPERCAYAAN DAN KEBOLEHGUNAAN KAEDAH AUTOMATIK UNTUK UJIAN KEPEKAAN PENDENGARAN (AMTAS) DENGAN VIDEO ARAHAN BAHASA MELAYU DALAM SETING KLINIKAL DAN BUKAN KLINIKAL

ABSTRAK

Kaedah Automatik untuk Ujian Kepekaan Pendengaran (AMTAS) adalah alat penilaian pendengaran kendiri yang menggunakan tablet, fon kepala, dan perisian AMTAS untuk menentukan ambang pendengaran. Walaupun AMTAS telah melalui kesahan di peringkat antarabangsa selama dua dekad dan menunjukkan potensi yang baik, kajian seumpamanya belum pernah dijalankan di Malaysia. Objektif utama kajian ini adalah untuk menilai kesahan, kebolehpercayaan, dan kebolehgunaan AMTAS dengan video arahan dalam Bahasa Melayu di kalangan orang dewasa di Malaysia, melalui tiga fasa. Fasa I melibatkan penterjemahan secara verbatim video arahan AMTAS ke Bahasa Melayu serta pengesahan muka untuk soal selidik Kebolehgunaan Aplikasi Kesihatan Mudah Alih (M-MAUQ) versi Melayu. Berdasarkan maklumbalas daripada 30 peserta (purata umur = 45.62 ± 14.13 tahun), indeks kesahan muka (FVI) adalah 0.98 bagi Item-FVI dan Skala-FVI, menunjukkan kejelasan dan kefahaman yang tinggi. Fasa II, yang dijalankan di seting klinikal di Hospital Tuanku Ja’afar Seremban (HTJS) dan Hospital Rembau (HR) menilai kesahan melalui perbandingan ambang pendengaran antara AMTAS dan ujian audiometri nada tulen (PTA), kebolehpercayaan melalui ujian semula AMTAS, dan kebolehgunaan AMTAS menggunakan M-MAUQ. Seramai 100 peserta (purata umur = 44.72 ± 14.13) menjalani ujian kesahan, manakala 30 peserta (purata umur = 44.43 ± 14.63) menjalani ujian kebolehpercayaan. Semua 130 peserta terlibat dalam

penilaian kebolehgunaan. Bagi kesahan, perbezaan purata ambang pendengaran antara AMTAS dan PTA adalah antara 0.30 dB HL hingga 3.40 dB HL, dengan lebih 95% berada dalam julat perbezaan ±10 dB. Persetujuan tahap pendengaran antara AMTAS dan PTA adalah tinggi (Kappa 0.95 - 0.96) dan kebolehpercayaan menunjukkan konsistensi yang baik bagi semua frekuensi (ICC = 0.64 - 0.99). Fasa III mengulangi prosedur Fasa II dalam seting bukan klinikal di sebuah pusat komuniti, melibatkan 22 peserta bagi kesahan (purata umur = 27.41 ± 7.92) dan 15 peserta untuk kebolehpercayaan (purata umur = 26.82 ± 8.29). Perbezaan antara ambang pendengaran AMTAS dan PTA adalah antara 3.18 dB HL hingga 7.73 dB HL, dengan kebolehpercayaan yang tinggi (ICC > 0.80). Kebolehgunaan menunjukkan penilaian tinggi dalam kedua-dua fasa II dan III, dengan skor purata melebihi 6.0 (daripada 7.0) dalam semua domain, mencerminkan tahap kepuasan dan kebolehgunaan yang tinggi. Ujian AMTAS juga terbukti secara signifikan lebih cepat berbanding PTA, dengan pengurangan masa purata sebanyak 3.03 minit dalam seting klinikal dan 2.68 minit dalam seting bukan klinikal. Kesimpulannya, AMTAS memenuhi piawaian PTA konvensional dari segi kesahan, kebolehpercayaan dan kebolehgunaan, dengan itu menyokong potensinya untuk digunakan dalam penilaian pendengaran di seting klinikal dan bukan klinikal bagi populasi di Malaysia.

INVESTIGATING THE VALIDATION, RELIABILITY AND USABILITY OF AUTOMATED METHOD FOR TESTING AUDITORY SENSITIVITY (AMTAS) WITH MALAY INSTRUCTIONAL VIDEO IN CLINICAL AND NON-CLINICAL SETTINGS

ABSTRACT

The Automated Method for Testing Auditory Sensitivity (AMTAS) is a self-administered hearing assessment tool that utilizes a tablet, headphones, and AMTAS software to determine hearing thresholds. While AMTAS has been validated internationally for over two decades and has shown strong potential, such research has never been conducted in Malaysia. This study aimed to assess the validity, reliability, and usability of AMTAS with a Malay instructional video among Malaysian adults through three phases. Phase I involved the verbatim translation of AMTAS instructional video and face validation of the Malay version of the mHealth App Usability Questionnaire (M-MAUQ). Based on responses from 30 participants (mean age = 45.62 ± 14.13 years), the face validity index (FVI) was 0.98 for both Item-FVI and Scale-FVI, indicating excellent clarity and comprehensibility. Phase II, conducted in clinical settings at Hospital Tuanku Ja'afar Seremban (HTJS) and Hospital Rembau (HR), assessed validity by comparing AMTAS and pure-tone audiometry (PTA) thresholds, reliability through AMTAS test-retest measurements, and usability using M-MAUQ. A total of 100 participants (mean age 44.72 ± 14.13) underwent validation testing, while 30 participants (mean age 44.43 ± 14.63) participated in reliability testing. All 130 participants were included in the usability testing. For validity, the mean difference between AMTAS and PTA thresholds ranged from 0.30 dB HL to 3.40 dB HL, with over 95% of results within an acceptable ±10 dB difference.

Agreement in hearing loss severity between AMTAS and PTA was high (Kappa = 0.95-0.96). Reliability demonstrated good consistency across all frequencies with intraclass correlation coefficient (ICC) ranged from 0.64 to 0.99. Phase III replicated Phase II in a non-clinical setting at a community centre, involving 22 participants for validation (mean age = 27.41 ± 7.92) and 15 for reliability (mean age = 26.82 ± 8.29). The mean difference between AMTAS and PTA thresholds ranged from 3.18 dB HL to 7.50 dB HL, with strong reliability (ICC > 0.80). Usability showed high ratings in both Phase II and III, with mean scores exceeding 6.0 (out of 7.0) across all domains, reflecting high user satisfaction and feasibility. AMTAS testing was also significantly faster than PTA, reducing test time by an average of 3.03 minutes per person in clinical settings and 2.68 minutes in non-clinical settings. In conclusion, AMTAS meets the standards of conventional PTA in terms of validity, reliability, and usability, supporting its potential for wider use in clinical and non-clinical settings among the Malaysian population.

CHAPTER 1

INTRODUCTION

1.1 Background of Study

Hearing is one of the five primary senses essential for daily activities. When any part of the ear, auditory organ, or nerves responsible for transmitting sounds to the brain malfunctions, it will result in hearing loss (Katz, 2015). An individual is diagnosed with hearing loss if they cannot hear as well as someone with normal hearing. The British Society of Audiology (BSA) classifies degree of hearing loss as mild (21 to 40 dB HL), moderate (41 to 70 dB HL), severe (71 to 95 dB HL), or profound (equal to or more than 95 dB HL) (BSA, 2018).

The World Health Organization's latest report on hearing indicates that over 1.5 billion individuals worldwide, constituting approximately 20% of the global population, are affected by hearing impairment. The majority of this demographic, approximately 1.16 billion people, experience mild hearing loss. In Southeast Asia, an estimated 401 million individuals are currently affected, with 109.4 million (5.5%) experiencing moderate or greater impairment. By 2050, the total number of individuals with hearing loss in the region is projected to rise to 666 million (WHO, 2021).

In Malaysia, the National Health and Morbidity Survey (NHMS) 2023 reported that 3.2 % of adults experience difficulty in hearing, making it the second most common disability among Malaysian adults after difficulty in seeing (Institute for Public Health [IKU], 2024).

1.1.1 Impact of Hearing Loss and Demand for Audiology Services

Hearing loss can adversely affect various aspects of an individual's life, especially if left unaddressed or if their communication needs are unmet (Haile et al.,

2021). This sensory deprivation can diminish quality of life and hinder access to verbal communication, potentially increasing the risk of dementia (Livingston et al., 2020) and cognitive decline in older ages (Lin et al., 2011).

The emotional aspects affected by hearing loss can include loneliness, isolation, depression, and anxiety (Kramer et al., 2002; Monzani et al., 2008; Li et al., 2014). Older adults with untreated hearing loss are especially at significant risk of developing these issues (Jayakody et al., 2022).

Research suggests that initiating early auditory rehabilitation upon the individual's recognition of self's hearing difficulties, may improve outcomes especially for those with age-related hearing loss (Pronk et al., 2011). Despite this, research indicates that adults often postpone seeking hearing assessments and treatments, with delays averaging between 7 to 10 years after initially noticing hearing difficulties (Davis et al, 2007). This delay is most likely due to a lack of awareness regarding the severity and impact of their impairment (Fischer et al., 2011; Smith et al., 2011; Contrera et al., 2016).

The rising number of individuals with hearing loss outweighs the number of audiologists serving the population with hearing healthcare services (Windmill & Freeman, 2013). In Malaysia, 6.3% of the registered Persons with Disability (PWD) population - approximately 736,000 individuals - fall under the hearing disability category (Department of Statistics Malaysia [DOSM], 2024). In 2023, 3,856 civil servants with disabilities were serving the country, with 8.7% from the hearing disability category (DOSM, 2024). WHO estimated that 93% of low-income countries and 76% of middle-income countries have fewer than one audiologist per 1 million people (WHO, 2021). A global survey of audiology services across 64 countries found

that 86% of respondents reported an insufficient number of audiologists to meet community needs (Goulios and Patuzzi, 2008).

In Malaysia, as at 2023, nearly 800 certified local audiologists are employed in both public and private sectors across the country. Among them, 198 are serving in 48 government hospitals (Malay Mail, Oct 2023). With a population of 34.1 million Malaysians (DOSM, 2024), this equates to one audiologist for every 42,500 people - far below the recommended ratio of one audiologist per 500 individuals (Malay Mail, Oct 2023). This shortage is especially pronounced in rural and underserved areas, where geographic and logistical challenges further hinder access to hearing healthcare services.

To address these challenges, the Ministry of Health (MOH) Malaysia has introduced and implemented several initiatives such as teleaudiology, funding hearing health programs, community-based rehabilitation programs, and training more audiologists to enhance access to hearing healthcare services, especially in rural and underserved regions (Rashid et al., 2020; MOH, 2021; Quar et al., 2024; Rashid et al., 2024; Romli et al., 2024). A survey of 43 Malaysian audiologists found that around 50% believed tele-audiology could positively impact care quality and accessibility, with higher adoption willingness if it improved care quality (Rashid et al., 2020).

These advancements highlight the importance of innovative solutions to bridge the gap in hearing healthcare accessibility and ensure adequate support for individuals with hearing loss.

1.1.2 Pure-tone Audiometry and Automated Audiometry

Fundamentally, hearing health care includes hearing assessments categorised into subjective and objective tests. Subjective tests require patients to respond to the stimulus, whereas objective tests do not require a response. The most common

subjective test in hearing health care is pure-tone audiometry (PTA). PTA is the gold standard for measuring hearing sensitivity and identifying presence and severity of hearing loss (Hoff et al., 2024).

Hearing thresholds for air and bone conduction are assessed using calibrated audiometers and transducers. These measurements involve presenting pure-tone stimuli across frequencies from 250 Hz to 8000 Hz. According to Katz (2015), PTA is the gold standard for tests for hearing diagnosis as it provides comprehensive information on hearing severity and type of hearing loss across specific frequencies. PTA is conventionally carried out manually by an audiologist.

Provision of an alternative approach to PTA is long anticipated, as it is unrealistic to expect the number of audiologists to increase soon (D'Onofrio & Zeng, 2022). Implementing self-assessment technologies, such as automated hearing evaluation tools, is one way to resolve this deficit (Swanepoel & Hall, 2010). Automated healthcare services encompass procedures such as screening, diagnosis, and interventions that can be performed without direct involvement from healthcare professionals. In situations where specialist healthcare personnel are scarce or unavailable, this approach helps optimize services and healthcare resources (Margolis & Morgan, 2008; Swanepoel et al, 2010). Automated audiometry protocols have existed for several decades and have gained increased popularity, especially due to the Covid-19 pandemic and the expanding field of teleaudiology (Eikelboom et al., 2022). An example of an automated healthcare service is automated audiometry, which automatically records hearing thresholds (Mahomed et al., 2013). Automated audiometry refers to hearing assessments that are self-administered without direct involvement from trained professionals (Shojaeemend & Ayatollahi, 2018).

Margolis and Morgan (2008) highlighted the insufficiency of hearing tests

conducted by audiologists in the United States, amounting to less than half of the actual demand. They further emphasised the global shortage of audiologists, underscoring the imperative to enhance access to hearing testing, a primary objective in their pursuit of developing automated tests.

In 2002, Margolis and the team introduced AMTAS, an acronym for the "Automated Method of Testing Auditory Sensitivity". It is a self-administered hearing assessment tool to obtain a diagnostic or screening audiogram (Margolis et al., 2007). It is one of those automated hearing assessment tools and is the primary focus tool for this research. AMTAS is appropriate for patients who can follow instructions in a standard manual audiometry test (Eikelboom et al., 2013). Research indicates that over 80% of patients can complete the AMTAS evaluation. If a patient cannot perform the self-evaluation, the test may be aborted and manual testing administered (Grason-Stadler, Inc., 2016).

AMTAS methodology and validity have been documented with many years of research and publications in international peer-reviewed journals (Margolis et al., 2010; Margolis et al., 2011; Margolis & Moore, 2011; Margolis et al., 2013; Margolis et al., 2016). The differences observed between air and bone conduction thresholds for automated and manual audiometry were within the acceptable 10 dB test-retest variation (Eikelboom et al., 2013). Concerning this, an AMTAS validation study was most recently done in Singapore, concluded that despite differences in hearing thresholds obtained via AMTAS and manual PTA, these differences fall within this standard acceptable range (Yeo et al., 2023).

1.1.3 Technology Acceptance Model and mHealth Application Usability Questionnaire

It is significant to note that user acceptance is critical to the success of

implementing new technology (Taherdoost et al., 2013; Taherdoost et al., 2014). Generally, acceptance is interpreted as an antagonism to the term refusal and means the positive decision to use an innovation (Mathieson, 1991). Understanding the user's perception towards adopting new technology could help facilitate further growth of the development of that particular technology (Taherdoost et al., 2009; Taherdoost, 2017).

Over the years, several researchers have developed theories and models to assess user acceptance. The Technology Acceptance Model (TAM) developed by Davis (1989), is widely used to assess new technology acceptance. TAM consists of four constructs: (i) behaviour, (ii) perceived ease of use, (iii) perceived usefulness, and (iv) behavioural intention. The model posits that an application's perceived ease of use significantly influences the users' acceptance. When users anticipate that a system is easy to operate, they are more likely to adopt and utilize it (Davis, 1989; Davis, 1993).

Mobile health (mHealth) apps can be used to perform tasks in components such as wellness management, behaviour change, health data collection, disease management, self-diagnosis, rehabilitation, and also act as an electronic patient portal and medication reminder (Kao & Liebovitz, 2017; Roess, 2017). A number of researches on mHealth apps have been performed for the past years, and the results have proven that well-designed mHealth apps can empower patients, improve patient compliance, and reduce the overall cost of health care (Seto et al., 2012a, 2012b; Fairman et al., 2013; Parmanto et al., 2013).

All mHealth apps can be categorized according to the type of target users which are either the patients or the health care providers. The type of target user is not determined by the user's background, but rather the purpose of using the app. Individuals who utilize mHealth applications to maintain, improve, or manage their

personal health are referred to as patients. Contrarily, healthcare providers are professional who employ mHealth apps to deliver healthcare services such as medication prescription, laboratory ordering, consultation, and patient education. Among various methods for evaluating mobile app usability, questionnaires are the most frequently used due to their simplicity in terms of execution and data analysis (Zhou et al., 2019).

The mHealth App Usability Questionnaire (MAUQ) was developed and validated by Zhou and colleagues in 2019 to evaluate the usability of mHealth applications for both patients and healthcare providers. It comprises of four versions based on the type of app; interactive or standalone, and its target users (patient or healthcare provider). The questionnaire includes three subscales; ease of use, interface and satisfaction, and usefulness. Usability is determined by calculating the total and average scores of all items - the higher the overall average, the better the app's usability. In 2021, Mustafa and team, translated and validated the English version of MAUQ (standalone for patients) into a Malay version of MAUQ (M-MAUQ) for mHealth app research and usage in Malaysia. In this study, M-MAUQ was used to assess the usability and user acceptance of AMTAS.

1.2 Problem Statement

Hearing loss is a growing public health concern in Malaysia, yet access to timely and accurate hearing assessments remains limited, especially in rural and underserved areas. Automated audiometry systems, such as the AMTAS, offer a promising solution by enabling self-administered pure-tone testing without direct audiologist supervision. AMTAS has been validated in multiple international studies and shown to produce threshold results comparable to conventional manual

audiometry. However, its application in the Malaysian context, especially using linguistically and culturally adapted materials, remains unexplored.

Successful implementation of such automated tools requires rigorous evaluation of their validity (accuracy compared to the gold standard), reliability (consistency of results), and usability (user satisfaction and ease of use). One way to assess validity in this context is by calculating a validation score, defined as the mean difference in hearing thresholds (in dB HL) between AMTAS and PTA. A smaller validation score indicates greater agreement between the two methods and thus, stronger evidence of validity.

In a multilingual and multicultural setting like Malaysia, language accessibility is a critical factor influencing usability and test performance. Currently, AMTAS delivers instructions primarily in English, which may pose comprehension barriers for non-English speaking users. While some countries have adapted AMTAS into their native languages, there is currently no published research evaluating its use with linguistically adapted materials for the Malaysian population. Therefore, this study includes the development and integration of a Malay instructional video to improve user understanding and support its use among local populations.

Although many studies have explored the use of automated audiometry in both clinical and non-clinical settings, further evaluation is needed to understand its performance across different environments in the Malaysian context. Clinical settings, such as laboratories or sound-treated and quiet rooms in hospitals, offer controlled conditions ideal for benchmarking accuracy. Non-clinical settings, such as halls or community centres, better reflect real-world conditions where background noise and absence of professional supervision may affect test outcomes. Assessing AMTAS in

both settings is essential to determine its validity, reliability and feasibility for wider implementation across diverse Malaysian populations

Sociodemographic factors such as age, race, education level, and language proficiency have not been adequately explored in existing AMTAS literature but are particularly relevant in Malaysia's multilingual and multicultural population. Including these variables may help explore potential associations with test performance and usability perceptions, and support more inclusive implementation strategies.

This study aims to address these gaps by evaluating the validity, reliability, and usability of AMTAS with a Malay instructional video in both clinical (hospital-based) and non-clinical (community-based) settings. It also examines how sociodemographic factors relate to AMTAS outcomes and usability ratings. The findings will provide critical insights into the feasibility of implementing automated audiometry such as AMTAS and support efforts to expand access to hearing healthcare in Malaysia.

1.3 Research Aim and Objectives

This study aims to assess the validity, reliability, and usability of AMTAS with Malay instructional video among the adult population in Malaysia. To achieve this aim, the study sets forth the following specific objectives:

1.3.1 Phase I Objectives:

i) To translate and integrate the Malay instructional video of AMTAS, which is suitable for implementation across all populations in Malaysia.

ii) To determine the face validity of the Malay version of mHealth App Usability Questionnaire (M-MAUQ).

1.3.2 Phase II Objectives:

i) To determine the validity and test-retest reliability of AMTAS with Malay instructional video among the adult population in a sound-treated room at a general hospital.

ii) To determine the validity and test-retest reliability of AMTAS with Malay instructional video among the adult population in a quiet room at a district hospital.

iii) To evaluate the association between participants' sociodemographic factors and AMTAS validation scores.

iv) To determine the total score and domain (subscale) scores of M-MAUQ for AMTAS with Malay instructional video in clinical settings.

v) To assess the correlation between participants' sociodemographic factors and M-MAUQ domain scores for AMTAS with Malay instructional video

1.3.3 Phase III Objectives:

i) To determine the validity and test-retest reliability of AMTAS with Malay instructional video among the adult population in a non-clinical setting.

ii) To determine the total and domain scores of M-MAUQ for AMTAS with Malay instructional video among the adult population in a non-clinical setting.

1.4 Research Questions

1. How valid and reliable is AMTAS with Malay instructional video for assessing hearing among the adult populations across various settings in Malaysia?
2. What is the association between sociodemographic factors and the validation score of AMTAS with Malay instructional video?

3. How do usability scores, as measured by the M-MAUQ, reflect the effectiveness of AMTAS with Malay instructional video in clinical and non-clinical settings among the adult population in Malaysia?

1.5 Research Hypotheses

1.5.1 Null Hypotheses:

1. There is no significant difference in validation scores between AMTAS with Malay instructional video and conventional PTA.
2. There is no significant difference in AMTAS with Malay instructional video thresholds across repeated testing sessions.
3. There is no significant correlation between sociodemographic factors and AMTAS with Malay instructional video validation scores.
4. There is no significant correlation between sociodemographic factors and M-MAUQ domain scores for AMTAS with Malay instructional video.

1.5.2 Alternative Hypotheses:

1. There is a significant difference in validation scores between AMTAS with Malay instructional video and conventional PTA.
2. There is a significant difference in AMTAS with Malay instructional video thresholds across repeated testing sessions.
3. There is a significant correlation between sociodemographic factors and AMTAS with Malay instructional video validation scores.
4. There is a significant correlation between sociodemographic factors and M-MAUQ domain scores for AMTAS with Malay instructional video.

1.6 Significance of Study

The significance of this study lies in addressing the critical gap in research concerning the validity, reliability, and usability of the AMTAS with Malay instructional video among adult populations in Malaysia. Despite extensive research validating AMTAS in Western countries, its applicability in Asian countries, particularly Malaysia, remains unexplored, except for a study in Singapore. This study aims to fill this gap by translating AMTAS instructional videos into Malay and assessing its validity, reliability, and usability in various clinical and non-clinical settings across Malaysia.

Moreover, the study is novel in its approach to integrate the M-MAUQ, a validated tool in the Malaysian context, to measure the usability of AMTAS among local populations. The study not only contributes to the enhancement of AMTAS accessibility and appropriateness for the Malaysian population but also provides a pathway for future research in this area.

Furthermore, the study addresses challenges related to maintaining acceptable ambient noise levels during AMTAS assessments and ensuring consistency of hearing profiles obtained via AMTAS and PTA. By monitoring noise levels during testing and comparing hearing thresholds between AMTAS and conventional methods, the study aims to validate AMTAS's accuracy and feasibility for widespread use in clinical and non-clinical settings.

In summary, the study findings have the potential to significantly impact the field of audiology in Malaysia by providing validated and reliable tools for self-hearing assessments, ultimately improving the efficiency and accessibility of hearing healthcare services for the Malaysian population.

1.7 Thesis Overview

The thesis comprises six primary chapters, each meticulously structured in accordance with the study's framework and interlinked with one another. Below, the overview of each chapter has been thoroughly examined and delineated. Chapter 1 introduces the research aim and objectives, focusing on assessing the validity, reliability, and usability of the AMTAS with Malay instructional video among the adult population in Malaysia. Specific objectives are outlined for each phase of the study, detailing the translation process of AMTAS instructional materials and determination of the face validity of the M-MAUQ, as well as validation and reliability testing of AMTAS in various clinical and non-clinical settings.

Chapter 2 explores automated hearing assessment methods and provides background information on AMTAS. It discusses the development of the M-MAUQ, offering a comprehensive overview of relevant literature in these areas.

Chapter 3 details the methodology of the study, divided into three phases. It describes the translation process of AMTAS instructional materials into Malay, the application of AMTAS in different clinical and non-clinical settings, and outlines the process of data collection and analysis.

Chapter 4 presents the findings of the study based on the specific objectives outlined in Chapter 1. It analyses the data descriptively and statistically, providing insights into the validity, reliability, and usability of the AMTAS with Malay instructional video.

Chapter 5 discusses the overall findings of the study, comparing them with previous research in the field. It addresses any limitations of the study and suggests areas for future research.

Chapter 6 concludes the thesis by summarizing the key findings and their implications. It outlines future plans for further research and development in the field of automated hearing assessment and usability testing of AMTAS in Malaysia.

1.8 Summary

This chapter provides a comprehensive foundation for the research, covering key topics essential to understanding the context and significance of the study. It begins with an exploration of the background of the study, shedding light on the importance of hearing and the prevalence of hearing loss. The chapter then delves into the impact of hearing impairment and the rising demand for audiology services, highlighting the necessity for effective hearing assessment tools. It discusses traditional methods like PTA, alongside emerging self-hearing test technologies, offering insights into contemporary approaches. Additionally, it introduces theoretical frameworks such as TAM and M-MAUQ, which guide the investigation into the acceptance and usability of the AMTAS with Malay instructional video. The problem statement identifies the gap in research regarding the validation and usability of AMTAS in Malaysia, leading to the formulation of research objectives, questions, and hypotheses. Finally, the chapter underscores the significance of the study in advancing automated hearing assessment technology and improving access to accurate diagnostic tools for hearing impairment in the Malaysian population.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Hearing loss affects 1 in 5 people worldwide and by 2050, is expected to affect 1 in 4. Effective treatment relies on the precise diagnosis of hearing loss. However, this initial step is beyond reach for more than 80% of those affected (World Health Organization, WHO 2021). A large gap exists between audiologists need and capacity to provide hearing assessments. Automation of hearing tests could increase the number of hearing-impaired patients to serve (Margolis & Morgan, 2008). This chapter discusses matters relating to automated hearing tests, beginning with the transition from conventional hearing assessments to automated hearing assessments. It also covers automated hearing test validation studies in Malaysia, and highlights the Automated Method for Testing Auditory Sensitivity (AMTAS), as a prominent option among automated hearing tests currently available.

2.2 Hearing Assessment

Hearing assessment start by presenting sounds to individuals at controlled intensities and determining their detection through various methods. Detection is assessed via behavioural responses or through physiologic measurements (Germiller, 2007). Behavioural assessments are preferred because they verify the integrity of the entire auditory pathway from the ear to the highest centres of the cortex (Katz, 2015).

2.2.1 Pure Tone Audiometry

Most audiologists would agree that pure-tone audiometry (PTA) represents a key component of the assessment battery (Katz, 2015). PTA is a fundamental

diagnostic tool that guides healthcare providers towards providing appropriate treatment. Audiometry evaluation over the range of frequencies that are essential for daily communication, can determine the degree, configuration, and type of hearing loss. Such detailed information aids the healthcare team in determining the aetiology, assessing the prognosis, and selecting the most effective treatment strategy for the hearing impairment (Musiek et al., 2017).

2.2.1(a) Audiometer and Transducer

Audiometers are used to make quantitative measures of air conduction and bone conduction pure-tone thresholds. Air conduction thresholds assess the entire auditory pathway, while bone conduction testing aims to stimulate the cochlea directly, bypassing the outer and middle ear. Audiometers can select tonal frequency and intensity level and route tones to the left or right transducer (Katz, 2015). In this chapter, our emphasis is more on air conduction thresholds, as the key components of the validation study using AMTAS are diagnostic air conduction thresholds.

The three main types of transducers that can be used for air-conduction audiometry are supra-aural, circum-aural and insert earphones. Supra-aural earphones (e.g., Telephonics TDH39 and TDH49) rest on the ear, while circum-aural earphones Sennheiser HDA200 surround and cover the entire ear. Insert earphones (e.g., Etymotic Research ER3 and ER5) use a disposable foam tip to direct the sound straight into the ear canal (British Society of Audiology, BSA 2018). All transducers must be calibrated before use. Transducers are matched to the audiometer and should not be interchanged without calibration (Malaysian National Society of Audiologists, MANSA 2023).

2.2.1(b) Testing Environment

In terms of testing environment, ideally, hearing assessments are performed in a sound-treated booth or room with low background noise. It should be performed in an acceptable test environment that complies with Maximum Permissible Ambient Noise Levels (MPANL) for Audiometric Test Rooms (MPANL; Frank et al., 1993). In Malaysia, audiometric tests shall be conducted in an audiometric test booth or sound-treated room. In general, the ambient noise of the testing room should not exceed 35 dB (A) (Katz, 2015; BSA, 2018).

2.2.1(c) Measuring Pure-Tone Thresholds

Pure-tone thresholds represent the lowest level at which an individual responds to a pure-tone stimulus. Pure-tones are the simplest sounds characterized by frequency, amplitude, phase, and duration. Standard PTA typically assesses thresholds for frequencies between 250 and 8000 Hz, which is very similar to the range of frequencies (100 to 6000 Hz) important for speech understanding (Katz, 2015).

The recommended method for threshold determination is the Modified Hughson-Westlake technique (Hughson & Westlake, 1944; Carhart & Jerger, 1959). The procedure begins by presenting a tone to the better-hearing ear at a level presumed to be audible, typically around 40 dB HL. If the individual responds, the intensity is decreased in 10-dB steps until there is no response. Subsequently, when there is no response, the examiner raises the intensity in 5-dB steps until a response is obtained again. Following this down 10 up 5 rule, the tester continues until a threshold estimate is obtained. MANSA (2023) recommends that the threshold should correspond to the level at which responses were obtained for at least 50% of presentations, often requiring two out of three responses at the same intensity.

Hearing threshold levels are often categorized by severity terms rather than numerical values at different frequencies. Common classifications include mild (21-40 dB HL), moderate (41-70 dB HL), severe (71-95 dB HL) and profound (>90 dB HL) (BSA, 2018). The original intent of the classification system is to provide a general understanding of the degree of hearing impairment and its associated impact. Many audiologists use these categories routinely when describing results to other professionals or patients during counselling (Katz, 2015).

2.3 Automated Audiometry

The PTA is the gold standard for measuring hearing sensitivity and identifying the presence and degree of hearing loss. It is conventionally carried out manually by an audiologist or other skilled operator–using standardised methodology. However, automated protocols have existed for several decades and have gained increasing popularity, especially in the light of the recent COVID-19 pandemic and the growing field of teleaudiology (Eikelboom et al., 2022). When considering the topic of automated audiometry, it is important to note that the methods used to acquire the hearing thresholds are robust and evidence based. Although these automated pure-tone testing methods have existed for quite some time, they have not been used extensively for diagnostic audiometry (Margolis & Morgan, 2008).

2.3.1 Definition of Automated Audiometry

Automated audiometry are merely devices that automate or program some of the procedures used in pure-tone audiometry (Siemiński, 1978). Automated audiometry encompasses all hearing tests that are self-administered using pre-programmed protocols without the continuous involvement of an audiologist. More specifically, automated audiometry is calibrated pure-tone threshold audiometry in any

setting (i.e., hearing health care, occupational health, and community settings) that is self-administered from the point the test starts (Wasmann et al., 2022). Some platforms are designed as stand-alone mobile tools for basic screening (e.g., smartphone apps), while others are integrated into clinical diagnostic equipment and offer both automated and manual modes (e.g., KUDUwave, AMTAS) (Shojaeemend & Ayatollahi, 2018). Automated audiometry employs validated threshold-seeking procedures such as the Hughson-Westlake method or the method of adjustment (Jerger, 2018). These systems are adaptable to various environments and have been applied in clinical, community, occupational, and home-based settings (Shojaeemend & Ayatollahi, 2018).

2.3.2 History of Automated Audiometry

The evolution of audiometry started as early as 150 years ago when tuning fork tests were the first tools used to obtain information on hearing acuity. Weber test was developed in 1845, and Rinne test was developed in 1855 (Feldmann, 1997). The evolution of audiometry progressed to the development of PTA in the 19th century (Jerger, 2018) and George von Bekesy's invention of self-recording audiometry (Békésy, 1947). The subsequent decades witnessed sophisticated attempts at automation, addressing the need for large-scale hearing loss screening (Jerger, 2018).

In 1947, Nobel Prize winner Georg von Békésy introduces an automated instrument for measuring pure-tone thresholds, making it the first used for automated audiometry (Békésy, 1947). This method, known as Békésy audiometry, requires the patients to maintain a high level of concentration. During the test, the patient controls the stimulus intensity by pressing a response button whenever they hear the tone, and releasing it when they no longer hear the tone. This approach is commonly known as the "method of adjustment", where the listener controls the stimulus intensity (Jerger, 2018). In the new Bekesy audiometer, the intensity level is automatically reduced

when the response button is pressed, and when the response button is released, the intensity level is automatically increased. Bekesy audiometry remained a prevalent automated pure-tone procedure in audiology until the 1970s, when it was replaced by the availability of more sensitive evoked potential tests (Margolis & Morgan, 2008).

Another method used in automated audiometry is in accordance with performing manual audiometry, also employing adaptations of the Hughson and Westlake threshold-seeking method. This approach is known as the "method of limits", where the strength of the stimulus is reduced until there is no response (a descending step), then increased from below this point until a response first appears (an ascending step). This sequence of descending and ascending steps is repeated three or more times. A threshold is defined as the midpoint between the average of descending and ascending steps (Jerger, 2018). This method has also been modified in some cases to include forced-choice responses from the patient. Here, the listener is required to listen and respond by either indicating that a sound was heard or not. This can be done, for example, by pressing the appropriate button on a touchscreen monitor after a signal is presented (Margolis & Morgan 2008).

2.3.3 Emergence of Automated Audiometry

During the 1950s and 1960s, saw many sophisticated attempts at automation, driven primarily to facilitate the need for large groups hearing loss screenings (Jerger, 2018). The increasing demand has led to the emergence of self-hearing tests as a promising tool in hearing health services, addressing challenges in both crowded urban clinics and underserved remote regions. Rudmose (1963) added that the number of audiometric examinations made today has grown to such a magnitude that it is only natural that some of the measurement techniques become automated.

In a renowned study by Margolis and Morgan (2008), they stated the rationale for automating pure-tone audiometry by comparing the need for hearing tests with the capacity of audiologists to administer these tests. They analysed the need for testing based on the prevalence of hearing impairment, the number of patients with normal hearing seen, and an assumption of the testing frequency. Capacity was based on the number of audiologists and the number of audiograms performed daily. Time savings were estimated from the average duration of an audiogram, and an assumption was made that 80% of the time could be automated. Their analysis showed a large gap between the need and the capacity of audiologists to provide testing. Even with 80% automation, this gap would only be partially closed.

Modern audiometers are predominantly PC-based and incorporate internal microprocessors that facilitate software-driven testing system (Margolis & Morgan, 2008). It is an essential tool in the diagnosis and management of hearing loss and is used in a variety of settings, including hospitals, clinics, and schools (Margolis & Morgan, 2008; Shojaeemend & Ayatollahi, 2018). One example on that account is the KUDUwave, developed in 2008 by GeoAxon Holdings, a South African company. This portable hearing assessment device features sound-attenuating headphones with a fully integrated functional audiometer within the headset. It offers both an automatic and manual testing mode and can also be implemented remotely via the internet (Storey et al., 2014).

Margolis & Morgan (2008) noted that the pure-tone audiometry protocol is clear-cut and straightforward, and meet the requirements to being automated. Pure-tone threshold testing follows a well-defined sequence of steps, that can be effectively executed by a computer. Moreover, the advent advances in technology and the widespread availability of powerful, cost-effective computers, integrated into

diagnostic test equipment have made automation of diagnostic pure-tone audiometry both feasible and practical.

2.3.4 Benefits and Challenges of Automated Audiometry

Automated audiometry has the advantage of being more cost-effective and is, therefore, frequently used in hearing loss screening programmes as well as in population-based studies (Hoff et al., 2024). Manually conducted PTA, on the other hand, is generally viewed as being more reliable, owing to the possibility of the operator adapting the method to respond to the individual needs of the patient. This is especially important when it comes to more difficult-to-test populations, such as older individuals. However, the operator can also be a source of bias in manual audiometry (Margolis et al., 2016). Clinical expectations or time constraints may cause audiologists to deviate from protocols, potentially influencing test results – a risk that computerised automated procedures can minimize (Margolis et al., 2007).

Another important advantage of automated audiometry is that the test can be either self-administered or administered by a general nurse, hence not dependent on the availability of audiological expertise–a scarcity in many parts of the world (Hoff et al., 2024).

Automated audiometry has the potential to decrease the time professionals spend evaluating patients, allowing them to focus more on difficult and complexed patients, such as children, certain elderly patients, and people with developmental disorders. By reducing the need for one-on-one clinician involvement, automated audiometry can lower testing costs and allow for hearing assessments to be conducted with minimal supervision, including in remote and decentralized locations. This approach lends itself well to telemedicine (Margolis & Morgan, 2008). Maximizing

professional productivity by using automation improves the overall efficacy in hearing health care (Swanepoel & Hall, 2010).

While often associated with screening purposes, certain automated audiometry systems, such as AMTAS, are capable of generating results that are clinically acceptable for diagnostic use. This expands the utility of automated audiometry in more comprehensive clinical evaluations (Margolis & Moore, 2011; Eikelboom et al., 2013). Automated audiometry platforms also vary in design – from mobile-based applications for mass screening to advanced boothless diagnostic systems integrated with clinical audiometers (Shojaeemend & Ayatollahi, 2018). This flexibility allows clinic to select solutions that best match their testing needs and infrastructure.

The field of audiology in the United States is progressing into a doctoral-level profession. This progression necessitates engaging in professional activities that are required to that particular advanced level. The true value of the audiology profession is not in the ability to performing routine hearing assessments but in interpreting test outcomes and implementing rehabilitative strategies, including fitting hearing aids, managing cochlear implants, and providing aural rehabilitation. Automated audiometry enables audiologists to use their time more effectively to deliver these said clinical services and procedures, reflecting a doctoral level of training (Margolis & Morgan, 2008).

The use of automated audiometry may also have present certain challenges, such as measuring the impact of environmental noise on the test results, recording bone-conduction hearing thresholds and ensuring the overall quality of automated test results. Further research is necessary to compare different computerized solutions and address these related challenges to optimize automated audiometry practices (Shojaeemend & Ayatollahi, 2018).

Although an automated audiometer cannot replace an audiologist, a system that can achieve accuracy similar to manual audiometry can be most beneficial to meet the growing demand for hearing health services. Margolis & Morgan (2008) conclude that the audiology profession can serve more patients with greater efficacy by exploiting the technology that already exists and advocate that automation will and should be utilized for all the benefits it contributes.

2.4 Validity and Reliability of Automated Hearing Tests in Clinical Setting

The historical progression from early electronic audiometers to the current state of automated audiometry underscores the need for continued exploration and validation (Jerger, 2018). Automated audiometry must be validated for accuracy and reliability before being implemented in widespread clinical use. Several successful clinical validations, accuracy and execution of automated audiometry in clinical settings had been reported over the years (Yu et al., 2011; Storey et al.,2014; Nyein et al., 2020; Bean et al., 2022; Serpanos et al., 2022; Wiseman et al., 2023; Lee et al., 2024; Hoff et al., 2024), a systematic review of their accuracy (Mahomed et al., 2013), and a review of their administration and assessment techniques (Shojaeemend & Ayatollahi, 2018) were published over the years.

Validity is measured by the accuracy and reliability of automated threshold audiometry to that of manual threshold audiometry (Mahomed et al., 2013). Accuracy involves comparing between two different techniques assessing the same variable (Bland & Altman, 1999), with manual audiometry serving as the gold standard, and automated audiometry as the comparison method for determining auditory thresholds. Test-retest reliability refers to the ability of a test to produce consistent results when administered multiple times under identical conditions (Dobie, 1983).