Validity of the Nijmegen Questionnaire for Hyperventilation Syndrome

Yui Man (Vickie) Li Ogilvie

MHSc

2017

Validity of the Nijmegen Questionnaire for Hyperventilation Syndrome

Yui Man (Vickie) Li Ogilvie

A thesis submitted to Auckland University of Technology in partial fulfilment of the requirements for the degree of

Master of Health Science

Faculty of Health and Environmental Sciences

School of Clinical Sciences

2017

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature: 

Acknowledgements

Throughout the journey of completing my degree in Master of Health Science, there were many people who supported me and I would like to sincerely thank them here.

I would first like to thank my supervisors Associate Professor Nicola Kayes and Professor Paula Kersten for introducing me to outcome measurement science. I admire the passion and dedication they share in this specialised area of research. Thank you for always being honest, respectful, and encouraging.

Nicola, thank you for acknowledging my effort, big or small. I thank you for the positive and constructive feedback that is always available and for trusting me on my abilities to face challenges in many forms.

Paula, thank you for your unchanging dedication to this research project. Thank you for providing meaningful directions that assist me in seeking knowledge and learning that is beyond my comfort zone.

I would also like to thank the participants who gave their valuable time to share their experiences and knowledge on the topic with me.

I would like to acknowledge the Cardiothoracic Special Interest Group of Physiotherapy New Zealand for granting part of the annual scholarship fund in 2015 and 2016 to support my study for the degree of Master of Health Science.

I would also like to acknowledge the Research Office at the hospital for granting approval to carry out this study; Mātauranga Māori Committee for providing verification and recommendations for this study; Auckland University of Technology Ethics Committee for granting the ethical approval (15-197) to proceed with this study on 24 June 2015.

Lastly, I would like to thank my parents for being my safe haven.

Abstract

The Nijmegen Questionnaire is an outcome measure used by health professionals for the assessment of symptoms associated with hyperventilation syndrome in the clinical setting and other areas of health research. However, a review of existing literature suggests evidence on the psychometric properties of the questionnaire is limited. Further, in some instances, the methods used in the investigation of these properties are questionable. The aim of this study was to investigate two specific psychometric properties of the Nijmegen Questionnaire: content validity and internal construct validity.

Based on the principles of outcome measure development and testing, Qualitative Descriptive Methodology and Rasch analysis were employed. To assess content validity, data from six patients with hyperventilation syndrome and four health professionals with relevant clinical experience were collected using semi-structured interviews. Interview data were analysed using conventional content analysis to identify symptoms of hyperventilation syndrome, and organised into categories and sub-categories. Data were then mapped against current items of the Nijmegen Questionnaire. In addition, data from 239 questionnaires were collected and analysed using Rasch analysis to establish the internal construct validity of the Nijmegen Questionnaire.

Results indicated that perceived symptoms of hyperventilation syndrome divided into three categories: breathing symptoms, psychological symptoms, and physical symptoms. Each category had various sub-categories, which included symptoms that were conceptually congruent with each other. There was only one questionnaire item that did not map onto symptoms identified by participants. However, there were noted differences in symptom reporting and language used between data and the existing questionnaire items. Likewise there were differences between patients and health professional reports. Rasch analysis showed the current Nijmegen Questionnaire did not fit the Rasch model. Issues were identified concerning a misfitting item which was under discriminating and demonstrated differential item functioning for gender. This item was deleted. All items had disordered thresholds. Response categories were collapsed and items with local dependencies between them which were conceptually similar were combined into testlets. The revised version of the Nijmegen Questionnaire contained 15 items, was unidimensional and fit the criteria for the Rasch model. Subsequently, a conversion table was created for transforming raw total scores of the Nijmegen Questionnaire.

The study provides additional evidence on the psychometric properties of the Nijmegen Questionnaire that allows clinicians and researchers to ascertain the value of this outcome measure for the screening of hyperventilation syndrome. Given that all but one existing item mapped onto interview data, the study findings demonstrated that the Nijmegen Questionnaire technically meets criteria for content validity. However, there are some differences in the way patients and health professionals talk about symptoms which should be considered when interpreting findings. The Nijmegen Questionnaire did not meet strict criteria for internal construct validity until after deletion of one misfitting item and dealing with threshold disordering and local dependency. Therefore, those using this measure should rescore the items as done in this study, ignore responses provided to item NQ14 (Cold hands or feet) and then use the conversion table to convert ordinal raw scores to interval scores.

Future studies are recommended to explore whether the questionnaire wording could be improved to better reflect how patients perceive their symptoms. Further research should include a more diverse range of patient participants by including males, individuals of different ethnic, and socioeconomic background to ensure these findings are transferable to these populations. Furthermore, future work should explore test-retest reliability and responsiveness of the Nijmegen Questionnaire.

Chapter One: Introduction

1.1 Background of the Study

The Nijmegen Questionnaire was first introduced to me in 2005 during my physiotherapy undergraduate studies at Auckland University of Technology (AUT University). The questionnaire was presented to physiotherapy students as an adjunct to physiotherapy assessment for individuals with suspected hyperventilation syndrome. In the first few years of my physiotherapy practice immediately after graduating from AUT University, the Nijmegen Questionnaire was practically absent from my repertoire of clinical assessments due to the clinical areas I was involved in at the time. It was not until 2010 that I was reintroduced to the questionnaire while studying as a postgraduate student in respiratory physiotherapy rehabilitation. In the same year, I started working at a Hyperventilation Physiotherapy Clinic in Auckland, New Zealand.

At the clinic, my caseload included the assessment and treatment of individuals with suspected or confirmed hyperventilation syndrome. My physiotherapy role involved triaging and prioritising individuals who were referred to the clinic, assessment and treatment in the form of group and/or one-on-one sessions, and the provision of peer support for other allied health staff who were less familiar with the condition to support recognition and management.

During one of the clinic sessions, I chatted with one of my patients. After just a few clinic sessions, this patient reported overall reduction in their symptoms associated with hyperventilation syndrome. The patient also expressed confidence in managing their symptoms independently by applying various treatment strategies obtained from the clinic. Clinically, the patient was considered ready for discharge. As was routine practice, the patient was asked to complete the Nijmegen Questionnaire prior to discharge to screen for the presence of hyperventilation syndrome. Based on the subjective interview and my clinical observations, I expected the patient to report a score that was below the positive screening threshold, indicating a reduction in symptom experience. Despite the report of significant subjective improvement, the patient scored above the screening threshold with only a few points less than her score at initial assessment. The seemingly mismatched outcomes generated by the Nijmegen Questionnaire and other clinical observations prompted me to enrol in the Measuring Health and Wellbeing paper offered at AUT University as part of my study towards the

Postgraduate Diploma in Rehabilitation in Health Science. Through this paper, I started to explore the principles of development and evaluation of outcome measures and carried out a critical review on the evidence of validity and reliability of the questionnaire (Li Ogilvie & Kersten, 2015). With the encouragement and support from the paper leader, I decided to undertake this study to investigate the Nijmegen Questionnaire as part of the course requirement for Master of Health Science.

1.2 Purpose and Significance of the Study

The Physiotherapy Board of New Zealand (2016) stated that their vision is "Fostering Excellence in Physiotherapy: Physiotherapists actively making a difference" (Vision, para. 1). In order to achieve excellence in physiotherapy, I believe one of the first steps is to select and apply assessments that are relevant, conceptually sound, valid, and reliable for specific patients based on clinical reasoning. Without this as the foundation, it could be a challenge to identify and implement appropriate treatments to make a difference for our patients. I was interested in gaining further understanding on the Nijmegen Questionnaire by exploring its content and internal construct validity. The knowledge gained from this research could potentially:

- increase the utilisation of the Nijmegen Questionnaire among health professionals in the clinical and research settings;
- empower health professionals to make relevant inferences from questionnaire scores;
- increase efficiency in identifying individuals with hyperventilation syndrome, facilitating referrals for individuals to receive specialised physiotherapy and/or other relevant health services.

1.3 Aim of the Study

The aim of this outcome measurement study was to investigate the content and internal construct validity of the Nijmegen Questionnaire for hyperventilation syndrome by drawing on the guidelines for outcome measure development and testing (McDowell, 2006; Streiner & Norman, 2008). Qualitative Descriptive Methodology (Sandelowski, 2000) was employed to explore the content validity of the Nijmegen Questionnaire. Study participants included patients who were diagnosed with hyperventilation syndrome and health professionals who were familiar with hyperventilation syndrome.

They were recruited from the Auckland region and each participant took part in a semi-structured interview, with a focus on exploring symptoms they attributed to hyperventilation syndrome and the appropriateness of the response options used in the Nijmegen Questionnaire. Rasch analysis (Boone, Satver, & Yale, 2014) was used to examine the internal construct validity of the Nijmegen Questionnaire. This involved the collection and analysis of completed questionnaires from the Hyperventilation Physiotherapy Outpatient Clinic at one of the three tertiary hospitals in Auckland, New Zealand.

1.4 Research Question

Is the Nijmegen Questionnaire a valid outcome measure for individuals with hyperventilation syndrome?

1.5 Definition of Terms

Definitions of key terms used throughout this thesis are provided below.

Breathing pattern disorder/Dysfunctional breathing

Breathing pattern disorder/dysfunctional breathing describes an abnormal respiratory pattern where the associated changes in rate/volume/pattern of breathing are disproportionate to the metabolic demands of the body (Warburton & Jack, 2006).

Hyperventilation syndrome

Hyperventilation syndrome is a form of breathing pattern disorder which an individual presents with, including a range of apparently unrelated physiological symptoms associated with chemical changes (i.e. a reduction of carbon dioxide) in the cardiovascular/circulatory system. The reduced level of carbon dioxide within the bloodstream is the result of an acute or chronic increase in respiratory response (e.g. rate and/or volume) that exceeds the metabolic demands of the body (Lum, 1975).

Outcome measure

An outcome measure is an assessment tool that is designed to test and quantify a predefined construct (McDowell, 2006). The result obtained from an outcome measure provides a baseline of the construct being measured for the individual being assessed. In some instances it can be used as a diagnostic or screening tool as well as a means to evaluate the effectiveness of therapeutic interventions.

Construct

Outcome measure experts Streiner and Norman (2008) define a construct as constellations of relatively abstract variables that are difficult to observe directly; "a 'mini-theory' to explain the relationships among various behaviours or attitudes" (p. 257). The authors explain that the diagnosis of a syndrome is based on a group of symptoms, hence, syndromes are considered as hypothetical constructs within the field of outcome measure development and testing. Hyperventilation syndrome is the construct of interest for this study.

Item

An item refers to a question or statement that is part of an outcome measure. Within an outcome measure, there are usually a number of items which are designed to capture relevant information from individuals about the construct being measured (Streiner & Norman, 2008).

Content validity

Content validity consists of "a judgement whether the instrument samples all the relevant or important content or domains" (Streiner & Norman, 2008, p. 3). This is established through an evaluation process, involving experts to determine the appropriateness of the outcome measure.

Internal construct

Internal construct is one of the components under investigation in the process of construct validation when developing or testing an outcome measure (McDowell, 2006). It refers to the internal structure of the items within an outcome measure (e.g. if and/or how the different items relate to each other and their relationships to the scores generated; if there are differences between age/gender groups). The scores generated from the outcome measure are also examined for their fit with a well-developed statistical model within measurement science (Bowling, 1997).

1.6 Structure of the Thesis

Chapter One: Introduction

The current chapter has provided the background of the study, the purpose and significance of the study in relation to physiotherapy, and the study aim. The research question has been identified. Definition of terms and the overall structure of the thesis have been included.

Chapter Two: Literature Review

This chapter will provide materials synthesised from the literature that are relevant to this study. It begins with an overview of hyperventilation syndrome. The prevalence and impact of the syndrome will also be explored. The symptoms and assessments of hyperventilation syndrome will also be discussed with reference to the literature. The structure and application of the Nijmegen Questionnaire will be introduced here. An overview of psychometric properties and statistical techniques in measurement science that are relevant to this study (i.e. validity and reliability) will be provided. The literature review methods will be outlined, and results will be presented and discussed with a focus on gaps in knowledge in relation to the psychometric properties of the Nijmegen Questionnaire.

Chapter Three: Methodology and Methods

This chapter will present the rationale for the selected study design, drawing on guidelines for outcome measure development and testing. It will include an overview of the methodological approaches (i.e. qualitative and quantitative) underpinning the study and a detailed description of the methods employed for recruitment of participants, data collection, and data analysis. Ethical considerations that are relevant to the study will also be presented here.

Chapter Four: Findings

This chapter will present the results of the data analysis. It will include the symptoms of hyperventilation syndrome shared by the participants, presented in different symptom categories and sub-categories. It will also provide the results of the Rasch analysis.

Chapter Five: Discussion and Conclusion

This chapter will discuss the study findings in relation to the existing literature on the symptoms of hyperventilation syndrome and the psychometric properties of the Nijmegen Questionnaire. Informed by the study results, this chapter will include discussion of study strengths and limitations, implications for clinical and research practice, and suggestions for future research.

Chapter Two: Literature Review

2.1 Introduction

This chapter begins with a brief overview of hyperventilation syndrome, including data on prevalence and the impact of this diagnosis within New Zealand. The symptoms of hyperventilation syndrome and the assessments for individuals who suffer from this will be explored from a physiotherapy perspective. The structure of the Nijmegen Questionnaire and its application will be outlined. This will be followed by a description of terminologies used in measurement sciences that are relevant to this study, including psychometric properties such as validity and reliability. A comparison between classical test theory and item response theory will also be provided to illuminate their differences in the evaluation of an outcome measure. A narrative review of the psychometric properties of the Nijmegen Questionnaire will then follow.

2.1.1 Hyperventilation Syndrome

Hyperventilation syndrome is a breathing pattern disorder, which is often undiagnosed due to its multi-systemic and apparently unrelated symptoms (Mooney & Candy, 2008; van Doorn, Colla, & Folgering, 1983). Patients with hyperventilation syndrome are regarded as high healthcare users (Chaitow, Bradley, & Gilbert, 2002; Lum, 1975) due to the involvement of various medical or surgical services and array of investigations. Mooney and Candy (2008) demonstrated that the financial implications could be significant for both patients with hyperventilation syndrome and their healthcare providers. Early diagnosis and implementation of individualised physiotherapy education and treatment were proposed as cost effective management approaches for patients with hyperventilation syndrome (Mooney & Candy, 2008).

2.1.2 Prevalence and the Impact of Hyperventilation Syndrome

There is a lack of population based cohort studies in the literature and thus the prevalence of hyperventilation syndrome is unknown. However, two cross sectional studies based at a general practice of 7,033 clients in the United Kingdom, showed that

approximately one in 10 (9.5%) adults who visited a general practitioner in the community suffered from symptoms associated with breathing pattern disorder, and that dysfunctional breathing was more prevalent in women and in individuals diagnosed with asthma (Thomas, McKinley, Freeman, & Foy, 2001; Thomas, McKinley, Freeman, Foy, & Price, 2005). Thomas and colleagues (2005) also highlighted that none of the clients had received the diagnosis of dysfunctional breathing nor underwent treatment for it previously, which implied that there may be patients with undiagnosed breathing pattern disorders who suffer from symptoms that could be treated by physiotherapy management (Chaitow et al., 2002; Mooney & Candy, 2008).

In New Zealand, the number of patients who attended their first session at the Hyperventilation Physiotherapy Outpatient Clinic at one of the tertiary hospitals in Auckland was 139 (99 [71%] females and 40 [29%] males) in 2014. However, the capacity of this service was limited to six to 10 clinical hours per week. The average wait time for an individual to receive this publicly funded specialised physiotherapy service was between six weeks to six months, depending on their clinical priority. Occasionally, some patients decided to access similar services in the private sector at the cost of $100 New Zealand dollars per session, to reduce wait time to around one week. Regardless of whether it was within the public or private sector, there was a sense of uncertainty and urgency expressed by patients and, at times, their family members who seek this specialised physiotherapy service. The dominating concern was often identified as the absence of a definitive diagnosis to explain their various symptoms, which presented in a number of body systems: cardiovascular/circulatory, digestive, muscular, nervous, respiratory, and skeletal system. A delay in recognising the presence of hyperventilation syndrome in a patient could negatively affect the employment status for the person in the form of multiple days of sick leave due to persisting symptoms; the costs of public health funding in the form of multiple investigations/tests, and visits at the emergency department and specialist clinics (Mooney & Candy, 2008).

2.1.3 Symptoms of Hyperventilation Syndrome

The definition of hyperventilation is “ventilation in excess of metabolic requirements” (p. 687), which causes a drop in the partial pressure of carbon dioxide in the circulatory system (Hornsveld, Garssen, Dop, & van Spiegel, 1990). This depletion of arterial carbon dioxide levels that is brought on by changes in ventilation, can produce a battery of symptoms that are not secondary to an existing organic disease (Grossman & de Swart,

1984). The common symptoms associated with hyperventilation syndrome are often presented in two main categories in the literature (Grossman & de Swart, 1984; Hornsveld et al., 1990).

The first category concerns with physical symptoms of hyperventilation syndrome. This includes palpitations (pounding of the heart), precordial pain (pain in the chest), shortness of breath, dizziness, faintness, fatigue, visual problems, sweating, excessive sighing, muscle pain, paraesthesia, tremor, tetany (overly stimulated neuromuscular activity [Chaitow et al., 2002]), aerophagy (swallowing of air), belching, and flatulence. It is important to note that the symptoms are often grouped together under various sub-categories, based on their association with the respective body systems (e.g. cardiovascular, digestive, muscular, nervous, respiratory, or skeletal system). The second category groups together psychological symptoms associated with hyperventilation. This includes the feeling of intense fear, panic, and anxiety. In addition, phobic responses, depersonalisation, disturbances of memory and concentration are also evident in the literature.

The complexity associated with the presentation of these symptoms in affected individuals can increase the chance of misdiagnosis of hyperventilation syndrome (Grossman & de Swart, 1984; Lum, 1981). In order to provide timely assessment and intervention for individuals who suffer from hyperventilation syndrome, there needs to be an agreed process among health professionals to identify these patients.

2.2 Assessments for Hyperventilation Syndrome

In many areas of physiotherapy practice, a comprehensive subjective interview is an essential part of the overall assessment of an individual. Subjective questions relating to the history of the present complaint, past medical history, medications, social history, stress, exercise, nutrition, and sleep will provide valuable information towards the assessment of hyperventilation syndrome. In addition to subjective assessment, a respiratory physiotherapy assessment includes objective examination of posture and general appearance of the individual. Other common observations include respiratory rate, breath-holding time, breathing pattern, and palpitation. Moreover, specific outcome measurement is used by clinicians as part of the overall physiotherapy assessment for individuals with suspected hyperventilation syndrome.

Diagnostic and screening tools for hyperventilation syndrome include the hyperventilation provocation test and a number of patient self-reported questionnaires (Vansteenkiste, Rochette, & Demedts, 1991). The hyperventilation provocation test is the criterion for diagnosis and requires an individual to hyperventilate for a few minutes to reproduce presenting symptoms of hyperventilation syndrome (Hornsveld, Garssen, Dop, van Spiegel, & de Haes, 1996). Outcome measurements that assess symptoms associated with hyperventilation syndrome and dysfunctional breathing include the Nijmegen Questionnaire, the 33-item Hyperventilation Questionnaire, and the Self Evaluation of Breathing Questionnaire (Courtney & Greenwood, 2009; Rapee & Medoro, 1994; Vansteenkiste.et.al. 1991). However, only the Nijmegen Questionnaire is suggested in the literature for the screening of hyperventilation syndrome as part of the overall diagnostic process for patients (van Dixhoorn & Duivenvoorden, 1985). The 33-item Hyperventilation Questionnaire "measures the state levels of cognitive, affective and somatic responses" (Sabourin, Stewart, Watt, & MacDonald, 2013, p. 193). This questionnaire is used to evaluate a person's responses to 'arousal induction exercises', including, but not limited to hyperventilation (Rapee & Medoro, 1994; Sabourin et al., 2013). The Self Evaluation of Breathing Questionnaire evaluates the symptoms of dysfunctional breathing that are considered 'breathing sensations' but, its developers state it cannot "be relied upon to detect hyperventilation syndrome" (Courtney & Greenwood, 2009, p. 125). Another questionnaire, the Rowley Breathing Self-Efficacy scale (Rowley & Nicholls, 2006) is associated with the assessment of people with breathing pattern disorders. However, its focus is on investigating the person's perceived ability to control their symptoms in relation to breathing pattern disorders. This leaves the Nijmegen Questionnaire, which is widely used for the detection and diagnosis of hyperventilation syndrome (van Dixhoorn & Duivenvoorden 1985; van Dixhoorn & Folgering 2015).

2.2.1 Nijmegen Questionnaire

The Nijmegen Questionnaire (Appendix A) is a short, self-administered patient-reported outcome measure consisting of 16 symptoms related to hyperventilation syndrome. The frequency of occurrence for each symptom can be rated on a five-point ordinal scale (0 = *Never*, 1 = *Rare*, 2 = *Sometimes*, 3 = *Often*, and 4 = *Very often*) and the points are added up to produce a total score out of 64 (van Dixhoorn & Duivenvoorden, 1985; van

Doorn, Folgering, & Colla, 1982). A score of 24 is a positive screening of hyperventilation syndrome (Garssen et al., 1984; van Doorn et al., 1983; Vansteenkiste et al., 1991).

This questionnaire does not provoke symptoms, in contrast to the hyperventilation provocation test. It is considered to be a useful screening tool for hyperventilation syndrome within the multidisciplinary setting (Chaitow et al., 2002) given its high sensitivity (up to 91%) and specificity (up to 95%), as demonstrated by van Dixhoorn and Duivenvoorden (1985). This measure is also used routinely as an outcome measure to evaluate change (rather than simply as a screening tool) in New Zealand physiotherapy practice of patients with breathing pattern disorder, including hyperventilation syndrome. However, data on the validity and reliability of the Nijmegen Questionnaire have not been synthesised to date.

2.3 Validity and Reliability in Measurement Science

2.3.1 Validity

The examination of validity is paramount in the process of outcome measure development and it involves a number of sequential steps before the final goal of creating a valid outcome measure is achieved (Laver Fawcett, 2007; Pallant, 2016). The basic definition of validity in the context of measurement science is the degree to which a scale is measuring what it was designed to measure (Hambleton & Jones, 1993; McDowell, 2006; Streiner & Norman, 2008). Streiner and Norman (2008) further define the process of validating an outcome measure as a means to establish the level of confidence we can assume when inferences are made about individuals based on their scores from that outcome measure. Validity can be grouped into three types, namely content, construct, and criterion validity, with the latter looking at specificity and sensitivity specifically in some instances (Bowling, 1997; McDowell, 2006; Pallant, 2016; Streiner & Norman, 2008). The three aspects of content validity can be identified as discriminative, convergent, and divergent validity (Laver Fawcett, 2007; Streiner & Norman, 2008). Mokkink and colleagues (2010) provide similar definitions for the three types of validity in their publication of the COnsensus-based Standards for the selection of health status Measurement Instruments (COSMIN). However, they define the three aspects of construct validity as structural validity, hypotheses testing, and cross-cultural validity.

Content validity

In the literature, it is suggested that the content validity of an outcome measure relates to whether the items or questions included are representative of all the attributes to be evaluated within the specified conceptual basis while meeting the objectives identified for the given outcome measure (Bowling, 1997; McDowell, 2006). Additionally, the inclusion of a representative sample in the process of outcome measure development can lead to more accurate inferences of individuals being evaluated that are applicable to a greater variety of circumstances, hence increasing the content validity of the outcome measure developed (Streiner & Norman, 2008).

A sound conceptual basis is essential in the development of a health-related outcome measure (McDowell, 2006). The various aspects of a specified conceptual model articulate the concepts and populations that an outcome measure intends to evaluate and the relationships between the concepts (Scientific Advisory Committee of the Medical Outcomes Trust, 2002). A defined conceptual basis of an outcome measure supports its content and allows the results obtained to be interpreted alongside a broader body of theory that is associated with the conceptual definition (McDowell, 2006).

Construct validity

The presence of hyperventilation syndrome is recognised through the identification of a variety of physical and psychological symptoms (Grossman & de Swart, 1984). Such constellations of symptoms of hyperventilation syndrome are considered as hypothetical constructs (Streiner & Norman, 2008). The process of construct validation of an outcome measure is complex because there is no one single test or criterion standard to follow (McDowell, 2006). Construct validity of an instrument can only be established through a continuous process of learning, understanding, and testing of the constructs (McDowell, 2006; Streiner & Norman, 2008). Test developers need to look for a cumulative pattern of evidence to ascertain whether the emerging outcome measure relates to the theoretical constructs proposed when assessing the construct validity (Laver Fawcett, 2007). The different aspects of construct validity are defined below.

Discriminative validity

This is concerning whether an outcome measure is able to provide a valid measure, distinguishing individuals or population groups based on the construct of interest (Laver Fawcett, 2007). Streiner and Norman (2008) explain that it involves the comparison of

test scores obtained from two distinct population groups. For example, the researchers find two groups of patients, with one group complaining of symptoms of hyperventilation syndrome and another who do not. Both groups complete the Nijmegen Questionnaire and the researchers establish whether there are any statistically significant differences between these two groups.

Convergent validity

To evaluate convergent validity is to see "how closely the new scale is related to other variables and other measures of the same construct to which it should be related" (Streiner & Norman, 2008, p. 263). The correlation between the two scores obtained from two different measures attempting to measure the same construct, contributes to establishing construct validity of the measure being evaluated (Streiner & Norman, 2008). An example would be to compare the scores from the Nijmegen Questionnaire with scores from another scale for breathing dysfunction such as the Self Evaluation of Breathing Questionnaire (Courtney & Greenwood, 2009). Hypothetically, it can be assumed that the Nijmegen Questionnaire is invalid if the scores do not correlate.

Divergent validity

Divergent validity is commonly evaluated after convergent validity in the process of construct validation. While the construct of interest is expected to correlate with similar variables, it is also expected to not correlate with dissimilar variables. To assess divergent validity of a newly developed scale, is to find out the extent to which the scores of this new scale are not correlating with the scores of another scale that evaluates a dissimilar construct. For example, researchers compare the scores from the Nijmegen Questionnaire with the scores from the Asthma Control Questionnaire (Juniper, Guyatt, Ferrie, & Griffith, 1993). If a strong correlation is found between the scores, the new scale could be considered invalid.

Internal construct validity

Another aspect of construct validation is the evaluation of internal construct validity of a scale by examining the internal structure of the items (e.g. if and/or how the different items relate to each other and their relationships to the scores generated; if there are differences between age/gender groups). The scores generated from the scale are examined for fit with a well-developed statistical model within outcome measurement science, which contributes to establishing internal construct validity.

Cultural validity

The cultural background of the person being evaluated can affect test administration and data interpretation (Laver Fawcett, 2007). A cultural validation process is not simply having the outcome measure translated to a different language. It should also ensure the conceptual foundation of the outcome remains unchanged after the necessary adaptation of individual items or development of new items (Beaton, Bombardier, Guillemin, & Ferraz, 2000). Health professionals should select a valid and reliable assessment tool that is also culturally relevant to the people being assessed (Høegh & Høegh, 2009). There are existing cross-cultural adaptation guidelines and processes in the literature that can help enhance the level of cultural validity or adaptability of an outcome measure (Beaton et al., 2000; Høegh & Høegh, 2009).

Criterion validity

Criterion validity is defined traditionally as the correlation of an outcome measure with another measure that is considered the 'gold standard' in the same field (Bowling, 1997; McDowell, 2006; Streiner & Norman, 2008). The comparison could be used formatively when developing a new outcome measure to guide the item selection process by recognising the elements that correlate optimally with the criterion or 'gold standard' (McDowell, 2006). When assessing concurrent validity (a form of criterion validity), researchers correlate a new outcome measure with a measure that has been validated, i.e. both measures are administrated concurrently (Streiner & Norman, 2008). An example of this type of validity in hyperventilation syndrome would be the comparison between the hyperventilation provocation test and the Nijmegen Questionnaire in recognising hyperventilation syndrome. Some researchers from the 1970s and 1980s suggested that the diagnosis of hyperventilation syndrome is proven by the hyperventilation provocation test (Lum, 1975, 1981; van Doorn & Colla, 1986). The level of arterial carbon dioxide was also suggested in the literature as a criterion for comparison (Grossman & de Swart, 1984; van Doorn et al., 1982).

2.3.2 Reliability

In measurement science, reliability is defined as the degree to which an outcome measure is free from error (Mokkink et al., 2010; Bowling, 2009). It is further explained that by evaluating reliability, the level of variability in measurement scores that is due to measurement error can be established (Streiner & Norman, 2008). An outcome measure

is more likely to reflect the true outcomes of individuals being assessed if the measure has evidence of proven reliability in the field (Laver Fawcett, 2007). The two types of reliability in relation to patient-reported outcome measurement are internal consistency and test-retest reliability (Bowling, 2001). Internal consistency reflects the degree of interrelatedness among the items within an outcome measure, whereas test-retest reliability is the extent to which scores obtained on the same version of an outcome measure for people who have not changed are the same for repeated measurement over time (Mokkink et al., 2010).

There is a lengthy history associated with the definition of reliability. Reliability is believed to be formally derived from what is generally referred to as classical test theory (Hobart & Cano, 2009; Streiner & Norman, 2008).

2.3.3 Classical Test Theory and Item Response Theory

Test theories and test models (Table 2.1) can assist researchers in understanding the implications of measurement errors (Hambleton & Jones, 1993; Hobart & Cano, 2009) in:

- the estimation of individual ability and potential ways to reduce these errors;
- the correlations among variables;
- the reporting of true scores or ability scores and the associated confidence levels.

Table 2.1
Comparison between Test Theories and Test Models

Test Theories	Test Models
Provide a framework linking observable variables (e.g. test scores, item scores) to unobservable variables (e.g. true scores, ability scores)	Formulate within the framework of a test theory; specify the relationships among a set of test concepts along with a set of assumptions about the concepts and their relationships.
Can only be judged for their utility when they are specified in particular test models	Provide only partial representations of test data to which the test models are applied.

Note. A brief overview of the differences between Test Theories and Test Models (Hambleton & Jones, 1993; Hobart & Cano, 2009).

This general knowledge on test theories, test models, and statistical frameworks, allows researchers to be more equipped to carry out measurement related projects, for example, concerning test development, test score equating, and identifying biased test items. Two groups of authors (Hambleton & Jones, 1993; Hobart & Cano, 2009) outlined the differences between classical test theory and item response theory, including the advantages and disadvantages associated with the application of these two frameworks in outcome measure development. Table 2.2 provides a summary of the differences between the two theories.

Table 2.2
Differences between Classical Test Theory and Item Response Theory

Classical Test Theory	Item Response Theory
Item difficulty; item discrimination; test statistics (e.g. reliability) are dependent on the sample population in which they are obtained.	Statistical theory about test item; test performance; how performance relates to the abilities that are measured by the items in the test.
Theory about test scores that introduces three concepts: (1) test score, (2) true score, (3) error score.	Logistic test models; model parameter estimation.
Models linking test scores to true scores.	Models linking item scores to true scores.
Sample population dependency reduces utility.	Item dependency increases utility.
Emphasis is on obtaining sample population for: generating item statistics, test statistics, and producing statistically parallel tests.	Item responses can be ordinal/ratio; dichotomous/polychotomous; un/ordered item categories; one/many abilities underlying test performance. Relationship between item responses and the underlying ability/abilities can be specified.
Weak models: assumptions are fairly easily met by test data.	Strong models assumptions are stringent; less likely to be met by test data.
E.g. Assumptions (1) true scores and error scores are not correlated, (2) average error score in the examinee samples are zero, (3) error scores on parallel tests are not correlated.	E.g. Assumptions (1) relates to the dimensional structure of the test data, (2) relates to the mathematical form of the item characteristic function or curve.

Note. A comparison between classical test theory and item response theory (Hambleton & Jones, 1993; Hobart & Cano, 2009).

It is essential to acknowledge the longstanding position and popularity of classical test theory among researchers in scale development and testing. Comprising of a body of principles and associated techniques, the classical test theory serves as the basis for many existing outcome measures and as a reference point for modern measurement approaches (DeVellis, 2006). It helps to determine the usefulness of various scales or instruments in estimating the variables of interest, explicating the relationship between the scores on scale items and the variables represented by the items. The theory is concerned with different properties of individual items (e.g. true scores, error scores, difficulties of items), with a primary emphasis on items as a group (i.e. the reliability and validity of the scale as a whole). Advantages of the classical test theory have been argued to include familiarity of key concepts (e.g. concepts such as reliability and validity and their associated indices are both familiar and well understood), the methods and associated software are accessible and easy to use, the model fits specific measures well, and the lack of requirement for item optimisation (DeVellis, 2006).

However, this theory is weak as its assumptions can be easily met by test data. The focus is placed on scale-level instead of item-level data. Also, the item and scale statistics derived based on this theory can only be applied to a defined population group. It assumes that each item on the scale contributes equally to the final calculated scores, irrespective of the correlation between each item and the construct being measured. This is unsuitable for ordinal data, which are frequently derived from patient-reported scales.

In contrast, item response theory is based on strong assumptions. The scale is expected to be unidimensional, assessing a single construct. It is assumed that there are varying levels of item difficulties and person abilities, which are taken into account when scale statistics are derived. The focus is placed on item-level instead of scale-level data. The relationship between an item and the overall score is predictable based on a group of factors, and it can be plotted and observed using item characteristic curves. This theory is argued to be superior in scale development and testing when ordinal data are involved, allowing the summing of item scores to generate test scores (Hobart & Cano, 2009; Streiner & Norman, 2008).

2.4 Narrative Review on the Psychometric Properties of the Nijmegen Questionnaire

A narrative review of literature in relation to the development and validation of the questionnaire was carried out with the aim to explore the psychometric properties of the questionnaire. The synthesised findings from this narrative review provided the foundation for the current research and contributed to the formulation of the research question.

2.4.1 Narrative Review Question

What is the state of the evidence in relation to the validity and reliability of the Nijmegen Questionnaire for adults with hyperventilation syndrome?

2.4.2 Narrative Review Methods

Principles of systematic review (e.g. in terms of a pre-specified search strategy, explicit exclusion/inclusion criteria) informed the review methods. A literature search of the electronic databases (EBSCO Health databases, including CINAHL and MEDLINE) and health-related citation index (SCOPUS) was undertaken to identify all articles that examined the validity and reliability of the Nijmegen Questionnaire for hyperventilation syndrome in adults, in addition to articles that were relevant to the development of the tool. Specific key words and phrases combinations were used for the electronic searches. See Table 2.3 for an overview of key search parameters.

Table 2.3
Literature Search Parameters

Key words/phrases combinations	
Population	breathing pattern disorder dysfunctional breathing hyperventilation syndrome
Phenomenon of interest	outcome measure assessment Nijmegen questionnaire
Measurement concept	reliability validity

Note. The literature searches were completed between 26/08/2014 and 25/08/2016. The key words and phrases combinations are listed in alphabetical order. Quotation marks are displayed as how they were entered into the search box.

In the context of the scope of Master of Health Science and the time and resources available for the study, papers published up until 25/08/2016 were included. There was no limit set on publication date. The titles and abstracts of each paper from the initial searches were reviewed for relevance after removal of duplicates. The full text was read if information provided in the abstract was insufficient. The reference lists of the articles identified from the initial searches were hand-searched to identify potential relevant titles. Studies were included if: (1) the aim of the study was to examine the psychometric properties (e.g. validity, reliability) of the Nijmegen Questionnaire for hyperventilation syndrome in adults; (2) the study contained information relevant to the development of the Nijmegen Questionnaire for hyperventilation syndrome in adults (≥ 18 years old). Studies were excluded if: (a) the study was published in languages other than English or Dutch (with Dutch papers being translated by one of the thesis supervisors whose native language is Dutch); or (b) participants of the study were diagnosed with any organic cardiac, neurological, or respiratory disease.

Critical evaluation of the studies that met the review criteria was guided by the COSMIN (COnsensus-based Standards for the selection of health status Measurement INstruments) checklist, a standardised tool for evaluating the methodological quality of studies concerning measurement properties (Mokkink et al., 2010; Terwee et al., 2012). The checklist considers eight key attributes of an outcome measure:

1. Conceptual and measurement model;

2. Reliability;

3. Validity;

4. Responsiveness;

5. Interpretability;

6. Respondent and administrative burden;

7. Alternative forms;

8. Cultural and language adaptations.

The criteria for the evaluation of these attributes have previously been defined by the Scientific Advisory Committee of the Medical Outcomes Trust (2002). Other standards for evaluation of outcome measurement properties are also available in the literature. However, those standards pay no attention to studies that apply item response theory models, have not been established as methodological quality

assessments, and have not been presented in the form of checklists (Mokkink et al., 2010; Terwee et al., 2012).

The development of the COSMIN checklist began with a four-round Delphi study with 57 international experts from various backgrounds (psychology, epidemiology, statistic, and clinical practice) to compile a checklist of outcome measurement properties (Mokkink et al., 2010). Outcome measurement properties were included if consensus was reached (i.e. ≥ 67% of the panel experts indicated 'agree' or 'strongly agree' on a five-point scale) and they were: internal consistency, reliability, measurement error, content validity, construct validity, criterion validity, responsiveness, and interpretability. Terwee and colleagues (2012) developed a scoring system for the COSMIN checklist through discussions among experts, which was tested on 46 articles from a systematic review. The authors further explained how each measurement property on the COSMIN checklist was scored using one of the four response options (i.e. excellent, good, fair, and poor) with specific criteria for each option defined in detail. They considered the "worst score counts" as essential in summarising the criteria so that one poor rating from a psychometric attribute will result in an overall poor quality score for that measurement property. In conclusion, the COSMIN checklist with its scoring system was considered to be a suitable tool in providing an overview of the methodological quality of psychometric studies for the purpose of this thesis.

2.4.3 Narrative Review Results

An overview of the paper selection process is shown in Figure 2.1. A total of 365 articles were generated electronically after discarding duplicates. Sixteen were identified as potentially relevant based on their study titles and/or abstracts. Fourteen of these were rejected based on the exclusion criteria. The two remaining articles were read in their entirety and reference list checking led to three more titles. Upon further inspections, four of the five articles provided information about the development of the Nijmegen Questionnaire and its validity and reliability data (Table 2.4). However, only two provided original data. These two research studies were led by van Doorn (1983) and van Dixhoorn (1985) respectively. The critical evaluation of these two studies was guided by the relevant questions for each property from the COSMIN checklist (Appendix B). Table 2.5 demonstrates a summary of the evaluation.

Figure 2.1 Literature Search Process and Results

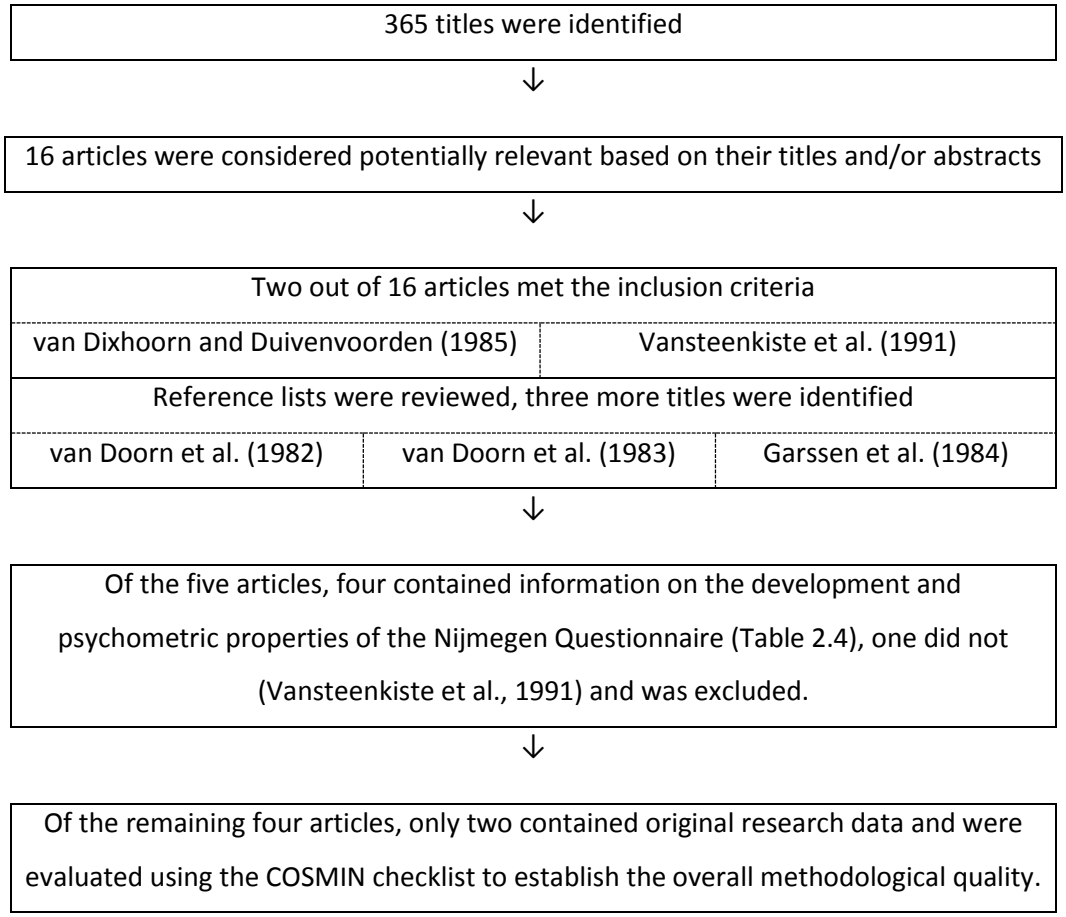


Figure 2.1. The process of literature search for the narrative review of psychometric properties of the Nijmegen Questionnaire.

Table 2.4

Summary of Studies in relation to the Narrative Review of the Psychometric Properties of the Nijmegen Questionnaire

Authors	Year	Study title	Purpose of study	Results
Van Doorn, Folgering, and Colla.	1982	Control of the end-tidal PCO_2 in the HVS: Effect of biofeedback and breathing instructions compared	To evaluate the efficacy of a behavioural management of HVS	Behavioural management supplemented with explanations about the mechanisms of HVS and coping strategies are useful
Van Doorn, Colla, and Folgering.	1983	Een vragenlijst voor hyperventilatieklachten [A questionnaire for HVS]	To investigate if a short questionnaire in which patients are asked to report the frequency of 16 common HVS is useful	The questionnaire was useful in patient screening and the provocation test could be used to rule out false positives
Garssen, Colla, van Dixhoorn, van Doorn, Folgering, Stoop, and de Swart.	1984	Het herkennen van het hyperventilatiesyndroom [Recognising the HVS]	To assess and review the NQ	*The NQ is able to discriminate (24 as the cut-off score) between individuals with and without HVS
Van Dixhoorn and Duivenvoorden.	1985	Efficacy of Nijmegen Questionnaire in recognition of the HVS	To establish the differentiating ability of the NQ by comparing individuals with and without HVS	The NQ is a suitable screening tool for early detection of HVS and in aid in diagnosis and management

Note. PCO_2 = partial pressure of carbon dioxide; HVS = hyperventilation syndrome; NQ = Nijmegen Questionnaire. *This study result was concluded from the study by van Doorn and colleague (1983).

Table 2.5

Summary of Study Evaluation using the COSMIN checklist in relation to the Nijmegen Questionnaire

Evaluated measurement properties	Studies with original research		Overall quality scores for each property	Questions for each property													
	Van Doorn, Colla, and Folgering (1983)	Van Dixhoorn and Duivenvoorden (1985)		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Reliability	✓		Poor	Good	Fair	Excellent	Poor	Excellent	Excellent	Good	Excellent	Good	Excellent	Poor	Poor	Poor	Excellent
Content validity	✓		Poor	Fair	Poor	Good	Fair	Poor	X	X	X	X	X	X	X	X	X
Structural validity		✓	Poor	Y	Good	Fair	Poor	Excellent	Excellent	Poor	X	X	X	X	X	X	X

Note. Only the measurement properties that are included in the two studies are presented here. Excluded properties are internal consistency, measurement error, cross-cultural validity, and responsiveness. ✓ = Study tested the specified measurement property. X = No further questions for methodological quality rating for the measurement property. Y = Yes, the scale consist of effect indicators, i.e. it is based on a reflective model. Each property has different number of questions within the COSMIN checklist as shown in the table. Adapted from *Rating the methodological quality in systematic reviews of studies on measurement properties*: *a scoring system for the COSMIN checklist* by Terwee and colleagues (2012). Please see Appendix B for the questions relevant to each property.

Table 2.5

Summary of study evaluation using the COSMIN checklist in relation to the Nijmegen Questionnaire (Continued)

Evaluated measurement properties	Studies with original research		Overall quality scores for each property	Questions for each property													
	Van Doorn, Colla, and Folgering (1983)	Van Dixhoorn and Duivenvoorden (1985)		1	2	3	4	5	6	7	8	9	10	11	12	13	14
Hypotheses testing (an aspect of construct validity)		✓	Fair	Good	Fair	Excellent	Fair	Good	Excellent	N/A	N/A	N/A	Excellent	X	X	X	X
Criterion validity	✓		Fair	Good	Fair	Excellent	Excellent	Excellent	N/A	Excellent	X	X	X	X	X	X	X

Note. Only the measurement properties that are included in the two studies are presented here. Excluded properties are internal consistency, measurement error, cross-cultural validity, and responsiveness. ✓ = Study that tested the specified measurement property. N/A indicates a lack of information from the study to answer the question listed. X = No further questions for methodological quality rating for the measurement property. Each property has different number of questions within the COSMIN checklist as shown in the table. Adapted from *Rating the methodological quality in systematic reviews of studies on measurement properties*: *a scoring system for the COSMIN checklist* by Terwee and colleagues (2012). Please see Appendix B for the questions relevant to each property.

Content validity

The conceptual and empirical basis for the inclusion of the 16 items on the Nijmegen Questionnaire was published over three decades ago (van Doorn et al., 1982). The researchers stated that the items were chosen out of a list of 45 complaints that were regarded as associated with hyperventilation syndrome for their clinical relevance by a group of specialists from various disciplines. These items were tested in two other studies with 40 and over 200 participants respectively, to assess the Nijmegen Questionnaire's effectiveness in differentiating between individuals with and without hyperventilation syndrome (van Doorn et al., 1982). This was an idiographic approach in item selection (McDowell, 2006), which employed empirical methods to select questions that best illustrate the eventual outcome after testing a larger number of items. The professional background of these specialists (physiology, psychology, and psychiatry) was published in a different paper in the following year (van Doorn et al., 1983). However, the authors offered no further details on the item selection process and there was no evidence to suggest the involvement of the target population in the process of content derivation. The Scientific Advisory Committee of the Medical Outcome Trust (2002) have suggested that to meet criteria of content validity both expert and layperson panels should judge the clarity, comprehensiveness, and redundancy of the items included in a measuring tool. This was only partially fulfilled by the developers of the Nijmegen Questionnaire. Considering the unavailability of this information, the level of adequacy regarding the selected items in relation to the conceptual basis of the Nijmegen Questionnaire warrants further investigation.

Furthermore, the title of the questionnaire appeared to only reflect its geographical origin (the city of Nijmegen in the Netherlands). The absence of association between the name and content of the questionnaire potentially negatively impacts on the face validity of the Nijmegen Questionnaire, which relates to its acceptability for individuals being assessed (Bowling, 1997; Laver Fawcett, 2007). Thus, based on the COSMIN criteria, content validity was rated as poor (Mokkink et al., 2010; Terwee et al., 2012).

Construct validity

In the 1985 publication by van Dixhoorn and Duivenvoorden (1985), the authors employed non-metric principal components analysis (NMPCA) to assess the construct validity of the Nijmegen Questionnaire. This was the first easily identifiable step in relation to construct validation process for the Nijmegen Questionnaire. The NMPCA was utilised to establish the dimensional structure of items included in the questionnaire and hence the structural validity (a form of construct validity) of the instrument (Tabachnick & Fidell, 2007; van Dixhoorn &

Duivenvoorden, 1985). Three components: respiratory, central tetany (overly stimulated neuromuscular activity [Chaitow et al., 2002]), and peripheral tetany were identified by the application of factor analysis and these followed the classic triad of hyperventilation syndrome related complaints (Lum, 1975). A key limitation of the study was an inadequate sample size to examine the structural validity of the Nijmegen Questionnaire; 75 patients were included, compared to sample size recommendations ranging between five to 10 people per questionnaire item (Thompson, 2004).

The construct validity of the Nijmegen Questionnaire was also examined using linear analysis of discriminance (van Dixhoorn & Duivenvoorden, 1985). It was performed to establish whether the questionnaire items were able to discriminate optimally between individuals with and without hyperventilation syndrome, hence assessment of discriminative validity (Streiner & Norman, 2008). The researchers found significant differences in the scores between the individuals with hyperventilation syndrome and those without across all components (van Dixhoorn & Duivenvoorden, 1985). In other words, participants with hyperventilation syndrome scored distinctly higher in all three groups of symptoms in the Nijmegen Questionnaire compared to those without the syndrome. Despite the appropriate application of statistical methods throughout the testing process, the methodological quality rating on the COSMIN checklist (Mokkink, 2010; Terwee et al., 2012) for construct validity was between fair to poor due to inadequate sample size, omission of clear hypotheses regarding the correlations, and how missing data were managed.

Criterion validity

Some evidence to support the criterion validity of the Nijmegen Questionnaire was presented by van Doorn and colleagues (1983). Participants with hyperventilation syndrome previously diagnosed by the hyperventilation provocation test (criterion/'gold standard') and those without the syndrome were asked to complete the Nijmegen Questionnaire and discriminant analysis was employed through the validating process. The authors found that the total scores of Nijmegen Questionnaire correlated strongly with the hyperventilation provocation test (van Doorn et al., 1983). While there was an adequate sample size of at least 100, the study did not provide sufficient information regarding the percentage of missing data and how they were managed. Thus, the evidence for the criterion validity of the questionnaire was deemed fair instead of excellent (Mokkink, 2010; Terwee et al., 2012). Moreover, van Dixhoorn and colleague (1985) demonstrated that the Nijmegen Questionnaire possessed a greater degree of specificity (94%) than sensitivity (89%). This suggests that the number of false alarms or false positives (i.e. people without hyperventilation syndrome who were identified as having

hyperventilation syndrome) was less than the number of false negatives (i.e. hyperventilation syndrome sufferers who were incorrectly identified as healthy). The authors concluded that the Nijmegen Questionnaire was a suitable screening tool for hyperventilation syndrome (van Dixhoorn & Duivenvoorden, 1985). They also suggested that a low score shall not be taken as a strong argument against the presence of hyperventilation syndrome (Dixhoorn & Duivenvoorden, 1985). Thus, the authors recommended additional subjective and objective information should be acquired from the individuals and a diagnostic test (e.g. hyperventilation provocation test) could be necessary (Dixhoorn & Duivenvoorden, 1985; van Doorn et al., 1983).

The COSMIN checklist manual states that the "COSMIN panel reached consensus that no gold standard exist for health-related patient-reported outcomes instruments. The only exception is when a shortened instrument is compared to the original long version" (Terwee et al., 2012, p. 38). Whether or not a criterion is considered to be adequate as a 'gold standard' in a measurement research, is based on the reader's judgement and the evidence provided within the research report (Terwee et al., 2012). A comprehensive literature review on the validity and reliability of the hyperventilation provocation test to determine if it can be considered a 'gold standard' seems like a logical step in expanding the investigation of this criterion. However, this is beyond the scope of the current study.

Decisions around the cut-off point for a screening tool had to be considered in relation to specificity and sensitivity (Laver Fawcett, 2007). McDowell (2006) proposed that "if the goal is to rule out a diagnosis, a cut-off point will be chosen that enhances sensitivity, whereas if the clinical goal is to rule in a disease, the cut-off point will be chosen to enhance specificity" (p. 32). The cut-off score of 23/64 for the Nijmegen Questionnaire was documented (Garssen et al., 1984; van Doorn et al., 1983; Vansteenkiste et al., 1991) and applied in the multidisciplinary health settings (Chaitow et al., 2002). However, the empirical evidence to support this was unclear in the literature. Van Doorn and colleagues (1983) supported their recommendation with original research by proposing 22/64 as the cut-off score. They also recommended that patients who were identified with hyperventilation syndrome to undergo the hyperventilation provocation test to rule out false positives. In the following year, Garssen and colleagues (1984) suggested the currently accepted cut-off score of 23/64 instead of 22/64 based on the summary of the research paper published by van Doorn and colleagues (1983) without carrying out their own evaluation of the patients. Although Garssen and colleagues (1984) recommended the steps in administrating the Nijmegen Questionnaire, there was a perceived reduction in credibility of this publication due to the lack of original research data.

Cultural validity

The Nijmegen Questionnaire was developed in the Netherlands. While this questionnaire had been widely used in the field of clinical practice and health research (Chaitow et al., 2002), there was no literature on its cultural validity in terms of clinical and research application in New Zealand (or other countries).

Reliability

Van Doorn and colleagues (1983) investigated the test-retest reliability of the Nijmegen Questionnaire. They concluded that the questionnaire was relatively stable given the coefficient of 0.87. However, by not determining the acceptable coefficient value prior to the study negatively affected their methodological quality rating on the COMSMIN checklist. The authors made the decision to retain all 16 items from the Nijmegen Questionnaire based on the range of bi-serial correlations obtained (.30 to .65) indicating that all items were associated with presentation of the hyperventilation syndrome. Based on the intercorrelations between all the items, which ranged from 0.03 to 0.52 (all items captured different aspects of hyperventilation syndrome), the authors suggested that the similarity between the retained symptoms of hyperventilation syndrome was minimal. Evidence for the reliability of the questionnaire was rated as fair because the authors did not report how missing data were managed and Kappa statistics were not presented (Mokkink, 2010; Terwee et al., 2012). Internal consistency of the Nijmegen Questionnaire has not been investigated to date.

2.4.4 Narrative Review Summary

The literature review identified a small number of studies concerning the development, validity, and reliability of the Nijmegen Questionnaire, of which only two studies contained original research. Despite the limited evidence presented over three decades, the questionnaire is still widely used in clinical and research practice. The methodological flaws identified in the two original research studies using the COSMIN checklist included the lack of target population involvement and missing items reporting, insufficient participants and statistical testing. Other measurement properties such as internal consistency, measurement error, responsiveness, and cultural validity have not been researched to date. Some of the methodological flaws could be addressed by designing and carrying out studies with more participants, with the application of more robust statistical tests to generate results that could be used to better evaluate the validity and reliability of the Nijmegen Questionnaire.

Chapter Three: Methodology and Methods

3.1 Introduction

The aim of this research was to investigate the content and internal construct validity of the Nijmegen Questionnaire for hyperventilation syndrome. The psychometric properties under investigation were limited (i.e. excluding sensitivity, specificity, and other psychometric properties) in the context of the scope of Master of Health Science and the resources available for the study (e.g. finance, time, and data source). This chapter presents the methodology that formed the foundation of this study and the methods used to investigate the content validity and the internal construct validity of the Nijmegen Questionnaire for hyperventilation syndrome. There are three parts in this chapter. The first part provides an overview of the selected research approach for outcome measure testing, including the Qualitative Descriptive Methodology and Rasch analysis and their relevance in measurement science. The second part describes the chosen methodology in more details and methods for the investigation of content and internal construct validity, respectively. The third part outlines ethical considerations and consideration for Māori that are relevant to the study.

3.2 Outcome Measure Development and Testing

This study employed guidelines for outcome measure development and testing, incorporating qualitative and quantitative research methods (Bowling, 2009; McDowell, 2006; Streiner & Norman, 2008). It was anticipated that the study could enhance the evidence base pertaining to various psychometric properties of the Nijmegen Questionnaire through the use of both qualitative and quantitative methods.

The decision on whether an outcome measure is robust for clinical and/or research application begins with the judgement on whether the items on the scale are appropriate (Bowling, 2009; Streiner & Norman, 2008). It also involves the critical review of other empirical evidence namely validity and reliability, supplementing the initial judgement on the overall appropriateness of the outcome measure (Hobart & Cano, 2009; McDowell, 2006; Streiner & Norman, 2008). These authors explained that a robust outcome measure can be expected to measure what it is designed to measure, consistently over time when other environmental and personal factors remain unchanged. In terms of establishing validity, different research methodologies are indicated depending on the specific types of validity under evaluation.

Neither qualitative nor quantitative research methods alone are likely to be adequate in establishing the overall validity and reliability of an outcome measure.

For example, content validity is concerned with whether the scale contains all the significant and relevant items that represent the trait being measured (Hobart & Cano, 2009). "No amount of statistical manipulation after the fact [devising the items] can compensate for poorly chosen questions [items]" (Streiner & Norman, 2008, p. 17). The qualitative research methods that are recommended to establish this include clinical observations, focus groups, and key informant interviews (Streiner & Norman, 2008). Using qualitative research methods, researchers can explore how people judge the relevance of items in relation to the trait being measured, contributing to the development of an outcome measure with the foundation of a robust conceptual basis.

Quantitative research methods are necessary in evaluating other psychometric properties (McDowell, 2006). For example, statistical analysis that explores the relationship between various items and item scores within a questionnaire provides useful information for researchers to make a judgement with regards to its construct validity. Using different quantitative research methods, researchers ensure the outcome measure is valid, reliable, and sensitive, allowing clinicians to confidently analyse and interpret data collected.

This study utilised the common principles of measurement science. The first part of the research question in the current study was concerned with the content validity of the Nijmegen Questionnaire, which was evaluated using a qualitative research approach, namely Qualitative Descriptive Methodology. The second part was concerned with the internal construct validity of the Nijmegen Questionnaire, which was examined quantitatively using Rasch analysis.

3.3 Investigation of Content Validity: Qualitative Descriptive Methodology

Common procedures in content validation include mapping the content of an outcome measure against an existing conceptual framework in the literature and/or asking patients and professionals in the area of interest to critically evaluate the content of the outcome measure (McDowell, 2006; Streiner & Norman, 2008). The Nijmegen Questionnaire purports to measure the frequency of symptoms in relation to hyperventilation syndrome (van Doorn et al., 1982) as experienced by patients, placing the patients at the centre of the assessment process. As such, I drew on a Qualitative Descriptive (Sandelowski, 2000) approach in the current study to

explore the symptoms of hyperventilation syndrome. These data were then compared and contrasted against existing Nijmegen Questionnaire items to critically assess its content validity.

Qualitative Descriptive Methodology is characterised as a commonly employed qualitative research approach (Sandelowski, 2000), which allows researchers to seek a rich description of an experience. In this study, the experience of symptoms in relation to hyperventilation syndrome was described from two different perspectives: (1) patients who experienced hyperventilation syndrome first hand and (2) health professionals who were identified as being familiar with hyperventilation syndrome. Research studies employing this particular approach present a descriptive summaries of data using everyday language (Sandelowski, 2000). It has been emphasised that the qualitative description in these studies involves interpretation with limited inference, allowing researchers to stay close to their data without the over rendering of their data (Sandelowski, 2000). Within the spectrum of qualitative methodologies, the Qualitative Descriptive Methodology is less likely to ascribe to a specific theoretical or philosophical lens (Sandelowski, 2000). Instead, research studies of this type draw from the general position of naturalistic inquiry (Guba & Lincoln, 1982; Lincoln & Guba, 1985). This form of inquiry emphasises the study of something in its natural state, to the extent that allowing the phenomenon of interest to emerge through the research process as it would if it were not under investigation (Lincoln & Guba, 1985; Sandelowski, 2000). Qualitative Descriptive Methodology was considered appropriate for the current study, since the intention was to explore and describe the participants' experiences with hyperventilation syndrome without a layer of interpretation in order to provide a comparison against the symptoms of hyperventilation syndrome that were already included in the Nijmegen Questionnaire.

3.3 1 Investigation of Content Validity: Methods

Participants

It is often overlooked that patients are an excellent source for item generation as they can report on the more subjective information pertaining to the phenomenon of interest (Streiner & Norman, 2008). Both patients and health professionals should be consulted with regards to the comprehensiveness and relevance of the items when investigating content validity of an outcome measure (Mokkink et al., 2010; Terwee et al., 2012).

For the investigation of content validity in this study, participants included:

a) patients who were diagnosed with hyperventilation syndrome; or
b) health professionals who worked/were working with patients with hyperventilation syndrome.

Patients or health professionals were included in the study if they were:

1. 18 years or older;
2. able to communicate in English in verbal and written form; and
3. able to provide informed consent in verbal and written form.

Patients were excluded from the study if they had:

a. a known organic cardiac, neurological, and/or respiratory disease.

Consistent with previously published articles associated with the development and validation of the Nijmegen Questionnaire between 1982 and 1985, only healthy individuals with "pure" hyperventilation syndrome were included (Garssen et al., 1984; van Dixhoorn & Duivenvoorden, 1985; van Doorn et al., 1983; van Doorn et al., 1982). This was to limit the potential of symptoms of other origin tainting the presentation of hyperventilation syndrome related symptoms. The crossover of symptoms could propose a risk in contaminating the research findings, impacting the chance of a direct comparison of the findings between this study and the existing literature. Therefore, patients with hyperventilation syndrome who were also diagnosed with an organic cardiac, neurological, and/or respiratory disease were excluded. Patients with psychiatric disorder were not excluded from the study by van Doorn and colleague (1983), the same stance was taken with the sampling approach in this study for consistency. The rationale for this was that an accurate history of psychiatric disorders and/or other psychological symptoms was likely to be challenging to ascertain due to the hospital patient confidentiality policy and the sensitive nature of this information.

Sampling

Purposeful sampling was used to select patients and health professionals to take part. This sampling technique is consistent with the Qualitative Descriptive Methodology (Patton, 1990), which aims to identify individuals who are most likely to provide information-rich data, informing and illuminating the research question (Sandelowski, 2000). For this study, individuals were purposely selected for their prior knowledge on hyperventilation related symptoms and Nijmegen Questionnaire. In addition to seeking information-rich cases, a

maximum variation sampling strategy was used to ensure a range of participants who could potentially provide varying perspectives on the topic (Patton, 1990).

Individuals from different age, gender, ethnic groups, and health professional (for health professionals only) disciplines were purposefully invited. Common themes which emerge from a diverse population would be of significant interest and value in capturing the researched experience given the level of heterogeneity these themes represent (Patton, 1990). This sampling technique was fitting for the aim of this study in seeking a rich description of the various hyperventilation related symptoms, allowing the researcher to compare and contrast these symptoms with the existing items on the Nijmegen Questionnaire.

For health professionals, the aim was to include individuals who held different clinical roles in the community- and hospital-based health services (e.g. medical specialists, nurse specialists, physiotherapists), in addition to the demographic characteristics listed above. This could provide contrasting perspectives on the symptoms shared by the physiologist, psychologist, and psychiatrist who were involved in the development of the Nijmegen Questionnaire (van Doorn et al., 1983), gaining views from the kinds of professionals who nowadays would be more likely to see these patients (i.e. more so than psychiatrist and/or psychologists). Patton (1990) stressed that this sampling method is not for the purpose of producing generalisable findings. Rather, it was to encourage the discovery of the prevailing themes that cut across great variations which can give insight to the phenomenon of interest. In this case, the end results were likely to be the symptoms recognised as associated with hyperventilation syndrome by a range of patients and health professionals with experience of the condition.

Sample size

It is argued that qualitative inquiry is to gather in-depth perspectives from a small sample (Patton, 1990). As such, sample size decisions are informed by a number of aspects relating to both the research question and methodological approach (Sandelowski, 1995). In the current study, the aim was to describe the symptoms of hyperventilation syndrome. To maximise the understanding of these symptoms, the number of participants was determined by what would be required to achieve maximum variation (in line with the sampling strategy described above). Van Doorn and colleagues (1982) mentioned the involvement of three health professionals (their age, gender, and ethnicity were not reported) who reached consensus on the items to be included in the Nijmegen Questionnaire. This study aimed to match this minimal sample size (minimum of three; maximum of five) and to include health professionals who were not specialists in physiology, psychology, or psychiatry (health professional

disciplines listed in the study by van Doorn and colleagues), aiming to capture perspectives not already included in the development of the questionnaire. There was no study in the literature that included patients when devising items for the Nijmegen Questionnaire. It was decided that a minimum of six and maximum of 10 patients were required to allow an equal number of representatives in both patients and health professionals (three from the study by van Doorn and colleague; three from the current study). The proposed sample size was achievable in the context of the resources available for the study (e.g. finance, time, and other participants-related factors.).

Recruitment

Patients were recruited from the Hyperventilation Physiotherapy Outpatient Clinic that operated within one of the district health boards in Auckland, New Zealand. Health professionals were recruited from Auckland-based healthcare services, including a community-based physiotherapy outpatient clinic, a hospital-based physiotherapy outpatient clinic, and two other specialised services within the hospital. The specialised services included respiratory medicine, and allied health services. Selecting local clinics had many advantages. First of all, the clinic locations were convenient for the researcher to manage recruitment logistics such as meeting with clinic staff and distributing research invitations. With potential participants who lived locally, it was likely to be less time consuming due to limited traveling time for the researcher and participants during the data collection process.

At the initial phase of recruitment, a hospital administrator and physiotherapy colleague distributed the Participant Information Sheet (Appendix C) to purposefully selected individuals. Interested individuals were invited to provide consent for their contact details to be passed on to the researcher or to get in touch with the researcher directly. However, after three months, only one individual had consented to participate. Given this, other recruitment strategies were explored and ethical approval was granted for those additional strategies. These included the distribution of study flyers via specialist services mailing list and by offering them to patients at clinic group sessions. The snowballing technique was also included.

After expressing interest to take part in the study, the researcher contacted potential participants and the Participant Information Sheet was provided. Individuals were encouraged to take time to review the study information and discuss this with their family/whānau. It was made clear that questions related to the study were welcomed at any point. A consent form (Appendix D) was signed by all participants prior to data collection.

Data collection

The data collection method for this current study was defined as key informant interviews, which entails the carrying out of "in-depth interviews with a small number of people who are chosen because of their unique knowledge. These can be patients who have, or have had, the disorder, for example, and who can articulate what they felt; or clinicians who have extensive experience with the patients and can explain it from their perspective" (Streiner & Norman, 2008, p. 19). Streiner and Norman (2008) discussed that the interview styles vary from being unstructured (informal, spontaneous conversations) to highly structured (interviewer with predesigned, carefully worded questions). They proposed, the less structured-type interview is generally more suited for investigating topics that are less known within the field of study. This study employed the semi-structured interview format.

Each participant took part in a semi-structured interview that was approximately one hour long. The broad focus of interview questions included identifying and exploring the symptoms associated with hyperventilation syndrome, including the evaluation of the appropriateness of the current items and response options of the Nijmegen Questionnaire. Two pilot interviews were carried out to test and review the interview guide, and for the researcher to gain interview experiences with individuals who were identified to have hyperventilation syndrome. Along with feedback from research supervisors on the questions, prompts, and interview style, the pilot interviews facilitated revisions to the interview guide (Appendix E) such as the type of language, phrases, and prompts used for the interviewing process, allowing the researcher to better capture the descriptions associated with the phenomenon of interest. Consent was obtained from both individuals prior to pilot interviews, allowing the inclusion of the data collected from these individuals in the analysis. The Voice Recorder application of the Samsung Galaxy S4 and S5 mobile phones was used for interview recording. Interviews were then transcribed verbatim by the researcher using Microsoft Word.

Data analysis

Data were analysed using conventional content analysis (Hsieh & Shannon, 2005). This is a research method used for analysing text data that allows the researcher to focus on the characteristics of language used to illuminate key concepts associated with the phenomenon of interest from the data (Hsieh & Shannon, 2005). It entails "the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (Hsieh & Shannon, 2005, p. 1278). It is the most appropriate approach for the current study because of its suitability within a research study that aims to describe a phenomenon, i.e. symptoms of hyperventilation syndrome. The literature review indicated

that this topic is yet to be investigated formally. Without any preconceived categories, this method of analysis enables the researcher to immerse in the data before recognising different codes and patterns within the data and constructing appropriate categories in relation to the researched phenomenon (Hsieh & Shannon, 2005).

The data analysis process commenced with the researcher, who listened to the audio recording of each interview multiple times as part of the transcription process. To familiarise herself with the data, each paper copy of the transcription was read and data related to symptoms of hyperventilation syndrome were highlighted by the researcher. Highlighted data were entered into a spreadsheet using Microsoft Excel during a second read. This spreadsheet was then checked against the highlighted paper copy during a third read. This demonstrated that the analytical process commenced prior to the formal coding process.

A number of techniques were used during the analysis of the pilot interview data. One of the techniques was using pen and paper to document highlighted symptoms (codes) in a list-form, including notes taken (see Appendix F for an example) by the researcher during the process. As part of the analytical process, the list was reviewed as a whole before grouping similar symptoms into different categories (see Appendix G for an example). Another technique was using different coloured pens for symptoms that belonged to different categories for clear visual definition on the page. Names were given to the different categories but, the symptoms and their categories were reviewed and revised as the researcher progressed through the analytical process and in discussion with the supervisory team. The computer based qualitative data management programme NVivo was trialled with two sets of transcripts at one stage following an introductory course attended by the researcher at AUT University. It was anticipated that the programme could improve efficiency of the analytical process. However, this strategy was abandoned due to the lack of time in gaining sufficient skills to navigate within the programme efficiently.

The prevailing techniques included using Microsoft Excel. The highlighted symptoms of similar nature were entered into a spreadsheet and grouped together to form categories. Sub-categories were then identified within each category. The categories, sub-categories, and symptoms were reviewed and revised throughout the analytical process. It was only after the preliminary categories and sub-categories were formed, the Nijmegen Questionnaire was reviewed by the researcher, to compare and contrast against the symptoms derived from interview data. The advantage was that it allowed the researcher to gain information on the symptoms of hyperventilation syndrome directly from the participants with minimal imposition of preconceived perspectives or theories.

Other techniques and strategies were implemented by the researcher to ensure the trustworthiness of the study findings. They included keeping a reflexive journal, having conversations with research supervisors, acting on feedback they provided, and documenting changes to strategies employed and coding decisions made. Specifically, feedback was provided at different stages of the coding and analytical process, ensuring that the researcher was staying close to the data as the categories, sub-categories, symptoms, and symptom clusters were developed. They were all important techniques and strategies in providing an audit trail, ensuring rigor was demonstrated throughout the study.

Once the preliminary categories, sub-categories, and symptoms/symptom clusters were formed, they were compared against the Nijmegen Questionnaire to identify any similarities and/or differences. Instead of being a one-step procedure, this comparison process was repeated as many times as required throughout the analytical process. This allowed the researcher to make changes to the previously identified symptoms/symptom clusters and revise sub-categories as indicated by interview data.

3.4 Investigation of Internal Construct Validity: Rasch Analysis

After establishing content validity, the common next step in the measure validation process is to evaluate the construct validity of the outcome measure (Laver Fawcett, 2007; McDowell, 2006; Streiner & Norman, 2008). As mentioned earlier, validation is often complex and involves a number of procedures depending on the types of construct validity under investigation.

The internal construct validity of the Nijmegen Questionnaire makes up the second part of the research question. Kersten and Kayes (2011) have suggested that internal construct validity is a psychometric property that is best evaluated by Rasch analysis. In the literature, statistical analysis of means, standard deviations, analysis of variances, and correlation coefficients are often calculated as part of the reporting of ordinal data using classical methods (Boone et al., 2014; Streiner & Norman, 2008). The interpretation of these results does not take into consideration the lack of equal interval in ordinal data. While there are other statistical techniques and methods of interpretation in the literature for assessing the internal construct validity, Streiner and Norman (2008) have highlighted that “many of these techniques draw on the statistical concepts of associations (correlation coefficients), repeated measures analysis of variance, and factor analysis”(p. 3). While these parametric statistical techniques and concepts are suitable for analysing ratio data, they are not suitable for ordinal

data (Hobart & Cano, 2009; Streiner & Norman, 2008). “Ignoring the parametric requirement of utilising linear measures [ratio data] can result in incorrect statistical conclusions” (Streiner & Norman, 2008, p. 301). Ordinal data need to be considered as frequencies in individual categories thus, non-parametric techniques are more suitable in this instance (Streiner & Norman, 2008). The individual item scores of the Nijmegen Questionnaire are summed to provide a total score despite the ordinal nature of the data. The summing and analysing of ordinal data using parametric techniques are problematic as “the score may decrease or increase but they are not evenly distributed as in the case with interval or ratio data” (Kersten & Kayes, 2011, p. 93).

Rasch analysis is based on the Rasch model, which is fundamentally different from the parametric statistical techniques that are based on classical test theory models (Boone et al., 2014; Kersten & Kayes, 2011). With the Rasch model, “the total score summarises completely how much of a construct the person has” (Kersten & Kayes, 2011, p. 93). It also takes into consideration the varying ability of individuals at the time of completing the outcome measure by analysing both test items and individuals’ abilities (Boone et al., 2014; Kersten & Kayes, 2011).

The Rasch model is often categorised under the umbrella of item response theory models. It is mathematically similar to the 1-parameter model within item response theory, yet there is a core philosophical difference between the two models (Boone et al., 2014; Hobart & Cano, 2009). While item response theory aims to search for the item response model that explains the data, Rasch measurement utilises only the Rasch model (Hobart & Cano, 2009). Rasch analysis is a quantitative method that examines the fit of data against the Rasch model and its theories (Kersten & Kayes, 2011). “If data do not fit this [model] researchers will seek to understand why and, if necessary, remove data, recollect data or reconceptualise the construct” (Hobart & Cano, 2009, p. 15).

Qualitative reflection is required when carrying out Rasch analysis, examining the fit between data and model assumptions in order to proceed to the next stage of analysis and the eventual outcomes (Boone et al., 2014). When a health-related outcome measure fits the Rasch model it is possible to transform its ordinal data into interval measures of individual responses. Additionally, in the likelihood of any missing items found within a questionnaire, Rasch analysis allows the individuals to be compared on the transformed single, equal interval scale. In contrast, missing data is often discarded when using statistical techniques within the classical test theory, producing confounding results (Boone et al., 2014).

In summary, the literature review demonstrated that there appeared to be an over reliance on statistical analyses based on the classical test theory in the evaluation of the validity of the Nijmegen Questionnaire, despite these drawing on parametric techniques which are inappropriate for ordinal data. In the current study, Rasch analysis was employed to evaluate the internal construct validity of the aforementioned questionnaire with the aim to provide a different perspective on the interpretation of the psychometric property of interest.

3.4.1 Investigation of Internal Construct Validity: Methods

Sampling

For the investigation of the internal construct validity, routinely collected Nijmegen Questionnaires completed by patients who attended the aforementioned Hyperventilation Physiotherapy Outpatient Clinic between 01/05/2013 and 30/04/2016 were extracted from patient clinical records. All Nijmegen Questionnaires within each patient clinical file were included regardless of the completion date. Files that were stored off-site were excluded due to the cost associated with accessing these files. The individual item scores and total scores of the questionnaires made up the data set for analysis, along with a selected range of person characteristics (e.g. age, gender, and ethnicity.). Permission for the researcher (who was also a staff member) to access the clinical files was granted by the Research Office at the hospital, given the letter of approval (Appendix H) issued by the Auckland University of Technology Ethics Committee (AUTEC).

A staff member at the Patient Information Department obtained the record of attendance for patients who attended the clinic. From this record, clinical files were retrieved by the Clinical Record Service for on-site access. The clinical files were moved from the Clinical Record Service to the Acute Allied Health office. The access to the offices was limited to staff and entrance was gained using a swipe card issued by the hospital. At the office, each file was hand-searched by the researcher for Nijmegen Questionnaires. The individual item scores from the questionnaires found in clinical files were entered into a Microsoft Access database. Access to this software was granted by the Information and Technology Department at the hospital. It was installed on a designated computer that the researcher had access to using a personalised login and password (unique to the researcher). Only the researcher had access to the questionnaire data set given the restricted access to the office, computer, and programme used. All clinical files were returned to Clinical Record Service immediately after use. These steps ensured that the research data were stored securely.

Sample size

For Rasch analyses, reasonably well targeted samples of 108 are reported to have 99% confidence that the estimated item difficulty is within +/- ½ logit of its stable value (Linacre, 1994). For poorly targeted samples, 243 are required for this level of confidence. We therefore erred on the side of caution and aimed to include 250 questionnaires (no upper limit was set for the number of questionnaires per patient).

Data collection

Parameters were set within the database to limit human errors during data entry. Examples included setting date format for assessment dates, a restricted number of options for gender and individual item scores. Total item scores were calculated by a pre-entered formula and the total item scores could not be calculated if there were any missing items. Each item score that was entered and the automatically generated total item score was checked against the questionnaire. The completed data set was converted to a Microsoft Excel file before being entered into IBM SPSS Statistics 22. Within the SPSS file, the data were formatted and saved as an ASCII file, suitable for import into the RUMM2030 (Andrich, Sherridan, & Luo, 2009) programme for Rasch analysis.

Data analysis

Using IBM SPSS Statistics 22, descriptive statistics were calculated for the Nijmegen Questionnaire data set. Summary statistics were calculated for demographic characterisitcs including age, gender, and ethnicity. The RUMM2030 (Andrich et al., 2009) programme was used for Rasch analysis, investigating whether the Nijmegen Questionnaire data fit the Rasch model. In the literature, the analytical steps (Kersten & Kayes, 2011; Siegert, Tennant, & Turner-Stokes, 2010) included (their order could vary):

1. Testing of the overall data fit to the Rasch model: The item-trait interaction chi-square probability should be non-significant
2. Checking of person fit to the Rasch model: Fit residuals should be within the range of +/-2.5, with a non-significant item fit chi-square probability, the mean fit residual should be closed to zero with a standard deviation value close to one.
3. Checking of individual item fit for the fit to the Rasch model: Fit residuals should be within the range of +/- 2.5 with a non-significant item fit chi-square probability, the mean fit residual should be close to zero with a standard deviation value close to one.

4. Identifying item(s) with poor fit to the Rasch model (using fit statistics outlined under 2.)
5. Identifying local dependency(ies) between items from the residual correlation matrix: The residual correlation should be < 0.2 above the mean residual correlation
6. Identifying items with disordered thresholds.
7. Analysing Differential Item Function (DIF) for demographic characteristics (e.g. age, gender, ethnicity, and assessment [time one; time two etc.]): Absence of DIF is shown if the analysis of variance (ANOVA) test is non-significant.
8. Testing of unidimensionality: Fewer than 5% of independent t-tests on estimates from testlets created from items with high positive and high negative loadings on the first principal component of the residuals should be significant.
9. (Potentially) Modifying the original scale by:
    a. deleting item(s) with poorest fit to the Rasch model;
    b. combining items with local dependencies;
    c. rescoring item(s) with disordered threshold(s).
10. Re-testing individual item fit and overall fit to the Rasch model
11. Distribution analysis of the participant-item thresholds

Rasch analysis of the Nijmegen Questionnaire for the current study incorporated the relevant steps listed above, each of which are discussed in more detail in the findings chapter to aid interpretation of data. The objective statistical outcomes generated by the RUMM2030 (Andrich et al., 2009) programme and the subjective interpretation of the outcomes in the context of the current items on the Nijmegen Questionnaire all contributed to the decision making throughout the analytical process.

3.5 Ethical Considerations and Consideration for Māori

This study was approved by the AUTEC (Appendix H) and the Research Office at the hospital (Appendix I). Approval was also sought from both the AUTEC and the Research Office at the hospital prior to the implementation of any amendments to the study design (e.g. recruitment strategies and/or localities, as described previously). The principles of partnership, participation, and protection was embodied within the study design and applied through the various research methods that were implemented. The Te Ara Tika guidelines (Hudson, Milne, Reynolds, Russell, & Smith, 2010) were also followed throughout the research process,

acknowledging bicultural social and cultural context in which the research was carried out i.e. in New Zealand, where Māori are the indigenous population. Consultation was carried out with the School of Rehabilitation and Occupation Studies Mātauranga Māori Committee at AUT University, where comments and recommendations were provided regarding Māori involvement and engagement in the study (Appendix J). The recommendations and relevant changes are provided in Table 3.1 below.

Table 3.1
Mātauranga Māori Committee Recommendations and Changes Made to the Study

Committee Recommendations	Changes Made
Consult further with Ko Awatea regarding appropriate phrases, greetings and possibility of inclusion of pepeha and mihi.	Māori greeting phrases, including pepeha and mihi were added to the Participant Information Sheet.
The Nijmegen Questionnaire may not be appropriate to form the basis of the open-ended interviews, seek alternatives which are available in the ICF browser for current information.	Semi-structured interviews were implemented using an interview guide.
Use *Te Whare Tapa Whā* (Ministry of Health, 2015) as a guideline to structure the interviews.	The four cornerstones of Māori (physical, spiritual, family, and mental) health were incorporated into the forefront of the interviewing process.

Note. ICF = International classification of functioning.
Recommendations from the committee were carefully considered and changes were made wherever appropriate in the context of the current study.

3.5.1 Privacy and Confidentiality

The privacy of study participants was paramount. Semi-structured interviews were carried out in private clinic rooms to ensure that the content of the interviews remained confidential. Various steps were taken to ensure the identity of the participants remained confidential. General descriptions were used to report recruitment localities for health professionals. A pseudonym was given to each participant. Digital files like the questionnaire data set were password protected. Identifiable information such as National Hospital Index and names were excluded during Nijmegen Questionnaires data entry. The researcher, who was also a staff member had clearance to access questionnaire data from patient clinical files with the approval by the AUTEC and the Research Office at the hospital. The researcher did not access any other medical information in patient clinical files.

3.5.2 Participant Distress

The recall of symptoms of hyperventilation syndrome was potentially distressing for some patients. It was possible for patients to report physical and/or psychological disturbances elicited by the interview process. If this occurred, patient was entitled to three free face to face counselling sessions at one of the Health, Counselling and Wellbeing Centres at AUT University.

3.5.3 Voluntary Participation

In addition to being the researcher of the current study, my role as a clinician and an allied health colleague in some of the recruitment localities was acknowledged. There were pre-existing professional relationships between the researcher and study participants. Through careful planning, it was ensured that a third party (e.g. administrative staff or health professional colleague) was always involved when study information was first introduced to potential participants. Only upon the indication of interest, the researcher would then contact individuals to discuss study information in detail. This was to avoid any sense of obligation felt by potential participants to take part due to pre-existing professional relationships. Throughout the study process, only the role as a researcher remained to fulfil various research-related tasks. The researcher endeavoured to avoid imposing her role as a clinician or colleague. All consented participants were informed of their right to withdraw from the study at any time without their care being affected.

Chapter Four: Findings

This chapter presents the findings concerning the content validity and the internal construct validity of the Nijmegen Questionnaire. This chapter contains three parts. First, I will present the symptoms of hyperventilation syndrome described by the study participants addressing the content validity of the Nijmegen Questionnaire. This will be followed by the Rasch analysis results, addressing the internal construct validity. A summary of these findings will be presented at the end of this chapter.

4.1 Content Validity: Interview Findings

In this section, I will first present an overview of the study participants. This is followed by the symptoms attributed to hyperventilation syndrome, reported in the interview data. The symptoms are classified into three categories: breathing symptoms, psychological symptoms, and physical symptoms. Within each category, there are different sub-categories. While symptoms that are conceptually congruent are grouped together to form symptom clusters, there are symptoms that stand alone within its sub-category. The number of symptoms varies in each cluster. These symptoms, organised in their clusters and sub-categories are then compared with the items that currently form the Nijmegen Questionnaire, evaluating the content validity of the questionnaire. The appropriateness of the available response options are also presented here.

4.1.1 Participants

Approximately 130 patients and 120 health professionals would have received information about the study inviting them to take part. Only a small percentage (< 5%) of those approached, translated into study participation. The availability of time and commitment could have been a contributing factor. For patients, it could also be related to perceived lack of understanding with regards to the connection between the reported symptoms and hyperventilation syndrome.

Ten patients (one male and nine females) expressed interest in the study. Two females were excluded due to existing cardiac and/or respiratory condition. One male and one female did not attend the scheduled interview session and the researcher was unable to establish

further contact with these two individuals. Six female patients met the study criteria and consented to participate. Four health professionals (one male and three females) expressed interest in the study, met the study criteria, and consented to participate. Their professional disciplines included nursing, physiotherapy, and respiratory medicine. Table 4.1a and 4.1b provide a summary of the participants, including their pseudonyms and demographic details (e.g. gender, age, and ethnicity).

Table 4.1a

Patients Demographics Summary

Pseudonym	Gender	Age	Ethnicity
Abby	F	26	NZ European
Becky	F	25	NZ European
Cathy	F	56	NZ European
Dora	F	51	Chinese
Eva	F	34	South African
Flora	F	64	Māori

Note. F = Female; NZ = New Zealand.

Table 4.1b

Health Professionals Demographics Summary

Pseudonym	Gender	Age	Ethnicity
Jessica	F	57	NZ European
Kelvin	M	54	Chinese
Leena	F	Undisclosed	European
Margo	F	58	NZ European

Note. F = Female; M = Male; NZ = New Zealand.

All participants expressed that they enjoyed being part of the study. Health professionals appreciated the opportunity to share their views on the topic and showed interest in how the research findings could inform their practice. Individuals with hyperventilation syndrome welcomed the chance to reflect on their condition and some appreciated the opportunity to realise their ability in managing symptoms independently after physiotherapy education.

4.1.2 Breathing Symptoms

This category includes breathing-specific symptoms that are associated with hyperventilation syndrome as identified by study participants. These symptoms are divided into three sub-categories: altered capacity, altered pattern, and global changes and difficulties. Table 4.2 provides a summary of the symptoms, symptom clusters within each sub-category, and a selection of quotes from study participants to illuminate these symptoms.

Altered capacity

The first sub-category contained symptoms that relate to the changes in the individual's breathing capacity. Participants used phrases such as "hyperventilating" and "over breathing" to describe changes noted in their overall breathing capacity. Becky, simply stated "I…actually breathe more". Also, more specific changes in relation to the speed of breathing were described. For example, "breathing fast" and "quick breaths" were recalled by both patients and health professionals. Other symptoms included "can't take a deep breath" or "can't completely fill my lungs", with reference to the depth of breathing. Sometimes, changes to both the depth and speed of breathing co-existed as Margo recalled a patient's breathing was "faster and shallower".

Altered pattern

The second sub-category contained symptoms that relate to the changes in the pattern of breathing. Participants described "chest breathing" as a common symptom. Leena described a patient with hyperventilation syndrome as an "upper chest breather", with an "apical pattern of breathing". Using less technical phrases, patient Becky described, "I breathe in my chest". Becky went on to explain that she "can't take a deep breath in my diaphragm area". This statement clearly described her symptom which spanned two sub-categories, highlighting that changes in capacity and pattern overlap. Cathy recalled, "I've been taught how long to breathe in and how long to breathe out. You forget all that". This was reflected in Leena's more technical description of "altered inspiratory/expiratory ratio". Other reported changes in breathing pattern included "noisy breathing", "heavy breathing", breath-holding, and unable to breathe with the nose.

Global changes and difficulties

The third sub-category included symptoms that were associated with the global changes and difficulties as described by participants. They noted that the act of breathing had changed. Feeling "short of breath" was a common feature within this sub-category. Patients described "gasping" and "running out of puff". Many expressed that they were unable to breathe well

nor properly. Abby described that she experienced “air hunger” while other patients felt that they were “not getting enough air (or oxygen)”. Sighs and yawns were frequently reported symptoms. Leena recalled patients who had “pseudo or ugly sighs”. They were usually big and noisy, seemingly inappropriate in relation to the mood or energy level of the individuals. A general mismatch between overall exertion and perceived breathlessness was recognised by patients and health professionals. Leena described her patients were often “unduly breathless compared to how they should be”, in terms of physical exertion.

Table 4.2

Breathing Symptoms of Hyperventilation Syndrome

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Altered capacity	1 Hyperventilating/ Over breathing	"I am over breathing or would it be hyperventilating like that." (Flora)	"…the over breathing with the hyperventilation syndrome is…" (Jessica)
	2 Breathing more/ Deep breathing	"I need to keep taking these deep breaths" (Eva)	"…they could be breathing deeply." (Kelvin)
	3 Breathing fast/ Shallow breathing	"Probably just being able to breathe in a shallow way." (Dora)	"They're breathing fast." (Kelvin)
	4 Difficulty filling lungs/ taking deep breaths	"I can't take a deep breath in and I can't completely fill up my lungs." (Abby)	"…patient usually say that they can't get quite a full breath or they can't fill their lungs fully" (Kelvin)

Table 4.2

Breathing Symptoms of Hyperventilation Syndrome (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Altered pattern	1 Upper chest breathing	"I have to breathe in my chest." (Becky)	"They'll [patients will] often alluded to the fact that 'I've always been told I'm an upper chest breather'." (Leena)
	2 Noisy/Heavy breathing	"You know breathing, heavy breathing…" (Flora)	"Patient's partner also complained or noticed noisy breathing." (Leena)
	3 Altered rhythm of breathing	"You're not breathing in a good rhythm." (Cathy)	"So the mechanics can include apical pattern of breathing, altered inspiratory expiratory ratio…" (Leena)
	4 Breath-holding	"I tend to kind of hold my breath a bit more I guess." (Becky)	"…breathing but holding their breath in a way." (Margo)

Table 4.2

Breathing Symptoms of Hyperventilation Syndrome (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Global changes and difficulties	1 Gasp/Pant/Puff	"Not concentrating on slowing down your breathing so you [are] just gasping." (Cathy)	"…they run out of puff or much more breathless…" (Leena)
	2 Short of breath	"I do feel like short of breath like I'm not getting enough oxygen." (Eva)	"Becoming short of breaths without any exertion." (Jessica)
	3 Air hunger	"I feel like I get air hunger." (Abby)	Nil
	4 Sigh/Yawn	"I also sigh and yawn a lot." (Abby)	"The sighing, yawning…" (Jessica)
	5 Difficulty breathing	"…it's the struggling to breathe, you feel like you have to catch your breath…" (Dora)	"…feeling not able to get their breaths." (Jessica)

4.1.3 Psychological Symptoms

This category includes psychological symptoms that are associated with hyperventilation syndrome as identified by study participants. These symptoms are divided into two sub-categories: feelings and thoughts. Table 4.3 provides a summary of the symptoms, symptom clusters within each sub-category, and a selection of quotes from study participants to illuminate these symptoms.

Feelings

This sub-category was the larger one of the two, including psychological symptoms related to feelings that were experienced by individuals with hyperventilation syndrome. There were five symptom clusters and one symptom which stood alone. The first cluster of symptoms were anxiety, fear, and panic. The second cluster of symptoms included aggravating, agitated, stressed, and rushed. The third cluster of symptoms were chaotic, confused, overwhelmed, and frustration. The fourth cluster was poor tolerance and hypervigilance. The fifth cluster had symptoms that describe various uneasy feelings. The stand-alone symptom that did not match the other symptom clusters was labelled 'disconnected'. There was a mismatch in symptom reporting between patients and health professionals for the third and fourth symptom clusters. Only health professionals acknowledged these feelings and their attribution to hyperventilation syndrome.

Thoughts

This sub-category was smaller than the one identified above, including psychological symptoms that reflected the thoughts of individuals with hyperventilation syndrome. Participants described individuals who viewed their body, self, situation, and/or world as "out of control". Jessica explained that patients thought their life was "out of balance". While patients specifically described worrying thoughts, health professionals did not. Abby described in her words, "something (is) always at the back of your mind". Cathy recalled herself thinking that "something is aggravating". Lastly, both patients and health professionals recalled that individuals were thinking about what was wrong with them, while being unable to articulate or investigate further.

Table 4.3

Psychological Symptoms of Hyperventilation Syndrome

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Feelings	1 Anxiety/Fear/Panic	"Probably more just a little bit of anxiety, it feel[s] like you can't breathe properly." (Abby)	"A general sort of sense of anxiety." (Margo)
		"Stress, panic, fear." (Dora)	"They might be fearful." (Kelvin)
		"When you're not breathing right, you do panic." (Flora)	"Sort of a panic about them [patients]." (Margo)
	2 Aggravating/Agitated/Stressed/Rushed	"Something is aggravating or you're rushed but you are not concentrating on it [breathing]"; "Very, just very tense. Agitated." (Cathy)	"They do all kind of bring up the idea that they're stressed." (Jessica)
		"I'm feeling anxious kind of makes me feel little stressed out." (Eva)	
	3.Chaotic/Confused/Overwhelmed/Frustration	Nil	"...their world feels…chaotic or confused or they feel this strong sense of…frustration"; "They feel often overwhelmed." (Jessica)

Table 4.3

Psychological Symptoms of Hyperventilation Syndrome (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Feelings (Continued)	4 Poor tolerance/Hypervigilance	Nil	"Sense of…hypersensitivity or…hypersomatisation"; "They have a very poor tolerance of whatever"; "…this kind of hypervigilance." (Leena)
	5 Uneasy/Feeling different/ Not feeling so good/Something is always at the back of your mind	"Just this uneasy feeling." (Eva)	"A genuine feeling of not being quite right but having nothing tangible to investigate or to act on." (Leena)
		"I just generally not feeling well"; "…say that I'm not feeling very good." (Dora)	
	6 Disconnected	"You…feel a little disconnected from your body." (Abby)	"There is a 'disconnect' between those two things [they symptoms and hyperventilation syndrome]." (Jessica)
			"…particular the male guys. They are not connected with their symptoms or their bodies." (Leena)

Table 4.3

Psychological Symptoms of Hyperventilation Syndrome (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Thoughts	1 Out of control/balance	"Worry because I am not in control." (Flora)	"Feeling that their body is…out of control or their situation is out of control"; their life is…out of balance." (Jessica)
	2 Worry	"It's probably a combination of not being able to breathe properly...I get a little bit more worried." (Abby) "It's a bit worrying, mentally." (Becky)	Nil

4.1.4 Physical Symptoms

This category includes physical symptoms that are associated with hyperventilation syndrome as identified by study participants. These symptoms are divided into seven sub-categories: bodily regulations, bodily sensations, head/face/mouth/throat, heart/chest, fingers/hands, muscle/posture, and speech/voice. Table 4.4 provides a summary of the symptoms, symptom clusters within each sub-category, and a selection of quotes from study participants to illuminate these symptoms.

Bodily regulations

This sub-category included symptoms related to changes in overall body regulations. These symptoms reflected deviations from normally observed bodily regulations. The symptoms from the three symptom clusters were linked to changes in temperature, level of perspiration, gastrointestinal function, and sleep pattern. Both patients and health professionals reported changes in temperature regulation, sleep quality, and sleep quantity. Change in gastrointestinal function was only noted by one health professional.

Bodily sensations

This sub-category included symptoms related to overall body sensations reported by participants. The symptoms from one symptom cluster included dizziness, faintness, and light-headedness. Passing out, physical collapse and vision goes dark formed another symptom cluster. Tiredness was a stand-alone symptom in this sub-category. Two patients recalled dizziness as a common symptom. For example, Jessica explained that “a lot of patients have experience of…dizziness, faintness”. Three patients (Abby, Becky, and Dora.) and Jessica all reported light-headedness as a common physical symptom of hyperventilation syndrome. Abby also reported light-headedness and changes in vision together.

Head/face/mouth/throat

This sub-category included symptoms associated with the head, face, mouth, or throat region. There were eight groups of symptoms. Examples included headaches, pressure in the head, frowning/tensing of the face, and changes in facial appearance. Tight feeling in the throat, dry mouth, and gritting of teeth, and throat clearing were also noted. Both patients (Cathy and Dora) and health professionals (Jessica and Margo) reported headaches. While two patients (Cathy and Flora) described changes in their facial expression such as frowning, health professionals did not. Cathy described a tightening sensation around her throat that was

associated with swallowing. Leena reported “throat clearing” as a common physical symptom. Only Flora recalled during her interview, “I notice that I’m gritting my teeth together”.

Heart/chest

This sub-category included symptoms associated with the heart and chest region. There were three symptom clusters, including increased heart rate, pain and/or tightness around the chest. There was discrepancy in language use between patients and health professionals when describing symptoms. Heart palpitation was used to describe increased heart rate by two health professionals, however, only by one patient. Phrases such as “heart…beating fast” or “heart racing” were used by patients. Two patients and one health professional described chest tightness and two health professionals and one patient reported chest pain. One of the health professionals Leena recalled “chest pain kind of group of symptoms”.

Fingers/hands

This sub-category included symptoms associated with altered sensations in the fingers and hands. There were two symptom clusters. Sweaty fingers and palms were recalled by one patient. Eva explained, “I get sweaty palms and yea just feeling like kind of short of breath”; “I’m feeling anxious so I’ve always got that sweaty palms”. This quote included both psychological and physical symptoms of hyperventilation syndrome. Only health professionals described paraesthesia/tingling of the hands as another symptom within this sub-category.

Muscle/posture

This sub-category included symptoms related to changes in muscle tension and/or posture. Examples included increased muscular tension, aches, and pains (generally or of a specific muscle group). Changes in posture was only described by patients. Cathy described, “Your posture might change and yea, your muscles would tense up”. Another patient described herself as “fidgety”, others used phrases such as “hold myself a bit tightly”. There was also discrepancy between patients and health professionals when reporting aches and pains.

Speech/voice

This sub-category included symptoms related to changes to an individual’s speech and voice pattern. These symptoms were all reported from an observational standpoint (i.e. by health professionals or patients who observed others with hyperventilation syndrome). Symptoms included talking excessively or running out of breath when talking. Other changes included the tone of voice becoming anxious or panicky. Leena explained that patients often struggled to manage their breathing and talking at the same time.

Table 4.4

Physical Symptoms of Hyperventilation Syndrome

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Bodily regulations	1 Feeling hot/sweaty	"You felt sweaty or hot, really hot." (Cathy)	"They might be sweating." (Kelvin)
	2 Constipation/Irritable bowel	Nil	"The feeling of constipation or irritable bowel." (Jessica)
	3 Sleep disturbances	"…lying in bed trying to go to sleep and…I can't catch my breath properly." (Abby)	"Poor sleep quality generally." (Leena)
Bodily sensations	1 Dizziness/Faintness/ Light-headedness	"Sometimes the dizziness just last despite me trying different things to calm my breathing down." (Eva)	"The light-headedness, the dizziness are the ones that more likely to take them [patients] to a medical practitioner." (Jessica)
	2 Passing out/Physical collapse/Vision goes dark	"You feel like you're going to pass out." (Dora)	"A lot of patients have experience of…physical collapse." (Jessica)
	3. Tiredness	"Very tired feeling" (Dora)	Nil

Table 4.4

Physical Symptoms of Hyperventilation (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Head/face/mouth/throat	1 Headache	"Just a little bit of a headache, not intense." (Dora)	"Headaches, quite often associated with it." (Margo)
	2 Pressure/Exploding feeling	"…pressure in your head or…an exploding feeling." (Cathy)	Nil
	3 Frowning/Facial expression	"Even your facial muscle[s] will be tense." (Cathy) "They [observers] would see…a scary looking face." (Flora)	Nil
	4 Pale	"She [patient's daughter] goes, '…you've gone pale'." (Flora)	"Sort of a panic about them and…pale." (Margo)
	5 Tight feeling in the throat	"Tightening around…your throat." (Cathy)	Nil
	6 Gritting teeth	"I'm gritting/clenching my teeth." (Flora)	Nil
	7 Dry mouth	"I get a dry mouth." (Abby)	"…they have…dry mouth." (Jessica)
	8 Clearing throat	Nil	"Throat clearing is another common one [symptom]." (Leena)

Table 4.4

Physical Symptoms of Hyperventilation (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Heart/chest	1 Heart palpitations/beats fast/racing	"Slight difficulty and then also heart palpitations." (Eva) "Your heart might beat fast [or] racing." (Cathy)	"…[patients] come in...saying they have palpitation." (Kelvin)
	2 Chest restriction/tightness	"It's just kind of…tight, more at the bottom." (Becky)	"Feeling of tightness in their chest." (Margo)
	3 Chest pain	"Getting this pain across here [chest] like a stitch kind of pain…and thinking, 'Oh my Gosh, I am having a heart attack'." (Flora)	"The chest pain kind of group of symptoms." (Leena)
Fingers/hands	1 Paraesthesia/Tingling	Nil	"Some people have sort of tingling in their hands." (Margo)
	2 Sweaty fingers/palms	"I've always got…sweaty palms/fingers." (Eva)	Nil

Table 4.4

Physical Symptoms of Hyperventilation (Continued)

Sub-Category	Symptoms	Quotes from Patients	Quotes from Health Professionals
Muscle/posture	1 Tense muscles	"You muscles would tense up." (Cathy)	"Visibly tense or on palpation of trapezius." (Leena)
	2 Aches and pains	Nil	"Muscle aches and pains, they [patients] may not necessarily go to the doctors." (Jessica)
	3 Postural changes	"You wouldn't be standing in a relaxed posture." (Cathy)	Nil
Speech/voice	1 Voice changes	"You can hear the panic/anxiousness in their voice." (Flora)*	"Feeling that they can't project their voice." (Leena)
	2 Talking more/faster	"She...would talk talk talk talk talk talk." (Flora)*	"They only…say a few words and they…talk really fast." (Margo)
	3 Poor breathing control	Nil	"They get short of breath when they're talking." (Kelvin)
			"Not being able to keep up with their friends in terms of walking and talking." (Leena)

Note. *Flora was a patient with hyperventilation syndrome. This was her observation of others with hyperventilation syndrome.

4.1.5 Interview Findings and the Nijmegen Questionnaire Items

Table 4.5 provides an overview of the symptom categories, sub-categories, and symptoms of hyperventilation derived from interview findings.

Table 4.5
Overview of Symptom Categories, Sub-categories, and Symptoms.

	Breathing Symptoms	
Altered capacity	1	Hyperventilating/Over breathing
	2	Breathing more/ Deep breathing
	3	Breathing fast/ Shallow breathing
	4	Difficulty filling lungs/taking deep breaths
Altered pattern	1	Upper chest breathing
	2	Noisy/Heavy breathing
	3	Altered rhythm of breathing
	4	Breath-holding
Global changes and difficulties	1	Gasp/Pant/Puff
	2	Short of breath
	3	Air hunger
	4	Sigh/Yawn
	5	Difficulty breathing
	Psychological Symptoms	
Feelings	1	Anxiety/Fear/Panic
	2	Aggravating/Agitated/Stressed/Rushed
	3	Chaotic/Confused/Overwhelmed/Frustration
	4	Poor tolerance/Hypervigilance
	5	Uneasy/Feeling different/Not feeling so good/ Something is always at the back of your mind
	6	Disconnected
Thoughts	1	Out of control/balance
	2	Worry

Table 4.5

Overview of Symptom Categories, Sub-categories, and Symptoms. (Continued)

Physical Symptoms		
Bodily regulations	1	Feeling hot/sweaty
	2	Constipation/Irritable bowel
	3	Sleep disturbances
Bodily sensations	1	Dizziness/Faintness/Light-headedness
	2	Passing out/Physical collapse/Vision goes dark
	3	Tiredness
Head/face/mouth/throat	1	Headache
	2	Pressure/Exploding feeling
	3	Frowning/Facial expression
	4	Pale
	5	Tight feeling in the throat
	6	Gritting teeth
	7	Dry mouth
	8	Clearing throat
Heart/chest	1	Heart palpitations/beats fast/racing
	2	Chest restriction/tightness
	3	Chest pain
Fingers/hands	1	Paraesthesia/Tingling
	2	Sweaty fingers/palm
Muscle/Posture	1	Tense
	2	Aches and pains
	3	Postural changes
Speech/Voice	1	Voice changes
	2	Talking more/faster
	3	Poor breathing control

Table 4.6, 4.7, and 4.8 highlights the differences between the current items on the Nijmegen Questionnaire and the interview findings, including the symptom categories, sub-categories and symptoms associated with hyperventilation syndrome. Seven items matched fully with a symptom identified from interview data in terms of linguistic consistency and conceptual congruency. Eight items matched partly with one or more symptoms identified in the current study, with discrepancy in language or not entirely conceptually congruent. For example, while NQ10 (Tingling fingers) and NQ14 (Cold hands or feet) captured symptoms from the fingers/hands sub-category, there was incongruence between symptoms reported by the patient (i.e. sweaty fingers/palms) and the health professional (i.e. paraesthesia/tingling).One Nijmegen Questionnaire item (NQ12 Stiff fingers or arms) did not match with the interview data as the word 'stiff' was entirely absent from participant descriptions. Table 4.9 provides an overview of symptoms that are not captured by the Nijmegen Questionnaire later on in this section (p.68).

In total, 15 items from the Nijmegen Questionnaire (93.75%) fully or partially captured symptoms from nine of the 12 sub-categories (75%) identified from the interview data. Three sub-categories (altered pattern, thoughts, and speech/voice) did not match any current items. Symptoms relating to altered pattern of breathing were recognised by both patients and health professionals. These symptoms were concerned with noisy or upper chest breathing and changes in the rhythm of breathing. There were some differences between patients and health professionals regarding the remaining sub-categories, with changes in *thoughts* mostly identified by patients and changes in *speech/voice* predominantly observed by health professionals, with the exception of one patient, who observed these changes in fellow patients.

There were 28 symptoms from six different sub-categories that were not captured by any item on the Nijmegen Questionnaire. Examples included: four of five (80%) breathing symptoms from *global changes and difficulties*; four of six (66.7%) psychological symptoms from *feelings*. For physical symptoms: two of three (66.7%) from *bodily regulations* and one of three (33.3%) from *bodily sensations*; six of eight (75%) from *head/face/mouth/throat*; two of three from *muscle/posture*. In other words, between 33.3% and 80% of the symptoms from these sub-categories were not captured. The symptom reporting by patients and health professionals was comparable for four of six sub-categories: global changes and difficulties (Breathing symptoms); bodily regulations; bodily sensations; and muscle/posture (Physical symptoms). Health professionals identified marginally more symptomatic feelings (Psychological symptoms) than patients. In contrast, patients reported the majority of physical symptoms from the head/face/mouth/throat sub-category.

Table 4.6

Interview Findings and Items of the Nijmegen Questionnaire: Breathing Symptoms

	Altered capacity				[A]Altered pattern				Global changes and difficulties				
Questionnaire item	1. Hyperventilating/Over breathing	2. Breathing more/Deep breathing	3. Breathing fast/Shallow breathing	4. Difficulty filling lungs/taking deep breaths	1. Upper chest breathing*	2. Noisy/Heavy breathing*	3. Altered rhythm of breathing*	4. Breath-holding*	1. Gasp/Pant/Puff*	2. Short of breath	3. Air hunger*	4. Sigh/Yawn*	5. Difficulty breathing*
Faster or deeper breathing (NQ06)	P	P	P										
Short of breath (NQ07)										F			
Unable to breathe deeply (NQ11)				P									

Note. F = Full match (consistent language, conceptually congruent);

P = Part match (some discrepancy in language or not entirely conceptually congruent).

*No match.

[A]No match for this sub-category.

Table 4.7

Interview Findings and Items of the Nijmegen Questionnaire: Psychological Symptoms

	Feelings						[A]Thoughts	
Questionnaire item	1. Anxiety/Fear/Panic	2. Aggravating/Agitated/Stressed/Rushed*	3. Chaotic/Confused/Overwhelmed/Frustration	4. Poor tolerance/Hypervigilance*	5. Uneasy/Feeling different/Not feeling so good/Something is always at the back of your mind*	6. Disconnected*	1. Out of control/ balance*	2. Worry*
Feeling confused (NQ05)			P					
Feeling of anxiety (NQ16)	F							

Note. F = Full match (consistent language, conceptually congruent).

P = Part match (some discrepancy in language or not entirely conceptually congruent).

*No match.

[A]No match for this sub-category.

Table 4.8

Interview Findings and Items of the Nijmegen Questionnaire: Physical Symptoms

	Bodily regulations			Bodily sensations			Head/face/mouth/throat							
Questionnaire item	1. Feeling hot/sweaty*	2. Constipation/Irritable bowel	3. Sleep disturbances*	1. Dizziness/Faintness/Light-headedness*	2. Passing out/Physical collapse/Vision goes dark	3. Tiredness*	1. Headache*	2. Pressure/Exploding feeling*	3. Frowning/Facial expression*	4. Pale*	5. Tight feeling in the throat	6. Gritting teeth	7. Dry mouth*	8. Clearing throat*
Blurred vision (NQ03)				P										
Dizzy spells (NQ04)				F										
Bloated feeling in stomach (NQ09)		P												
Tight feelings around mouth (NQ13)											P	P		

Note. F = Full match (consistent language, conceptually congruent);

P = Part match (some discrepancy in language or not entirely conceptually congruent).

*No match.

Table 4.8

Interview Findings and Items of the Nijmegen Questionnaire: Physical Symptoms (Continued)

	Heart/chest			Fingers/ hands		Muscle/ posture			[A]Speech/ voice		
Questionnaire item	1. Heart palpitations/beats fast/racing	2. Chest restriction/tightness	3. Chest pain	1. Paraesthesia/Tingling	2. Sweaty fingers/palms	1. Tense muscles	2. Aches and pains*	3. Postural changes*	1. Voice changes*	2. Talking more/faster*	3. Poor breathing control*
Chest pain (NQ01)			F								
^Feeling tense (NQ02)						P					
Tight feelings in chest (NQ08)		F									
Tingling fingers (NQ10)				F							
[B]Stiff fingers or arms (NQ12)											
Cold hands or feet (NQ14)					P						
Palpitations (NQ15)	F										

Note. F = Full match (consistent language, conceptually congruent);

P = Part match (some discrepancy in language or not entirely conceptually congruent).

*No match; ^this item could be variably interpreted in practice.

[A]No match for this sub-category; [B]no match for this item.

Table 4.9

Overview of Symptoms identified in interviews that are Not Captured by the Nijmegen Questionnaire

Symptom Sub-category	Symptom/Symptom cluster
[A]Altered pattern	1 Upper chest breathing
	2 Noisy/heavy breathing
	3 Altered rhythm of breathing
	4 Breath-holding
Global changes and difficulties	5 Gasp/Pant/Puff
	6 Air hunger
	7 Sigh/yawn
	8 Difficulty breathing
Feelings	9 Aggravating/Agitated/Stressed/Rushed
	10 Poor tolerance/Hypervigilance
	11 Uneasy/Feeling different/Not feeling so good/Something is always at the back of your mind
	12 Disconnected
[A]Thoughts	13 Out of control/balance
	14 Worry
Bodily regulations	15 Feeling hot/sweaty
	16 Sleep disturbances
Bodily sensations	17 Tiredness
Head/face/mouth/throat	18 Headache
	19 Pressure/Exploding feeling
	20 Frowning/Facial expression
	21 Pale
	22 Dry mouth
	23 Clearing throat
Muscle/posture	24 Aches and pains
	25 Postural changes
[A]Speech/voice	26 Voice changes
	27 Talking more/faster
	28 Poor breathing control

Note. [A]No match for this sub-category. Number one to 28 indicates the total number of symptoms that are not captured by the Nijmegen Questionnaire.

4.1.6 Response Options of the Nijmegen Questionnaire

Study participants also provided feedback on the appropriateness of the response options (0 = *Never*, 1 = *Rare*, 2 = *Sometimes*, 3 = *Often*, and 4 = *Very often*) currently used in the Nijmegen Questionnaire. Five of six patient participants were asked to feedback on the response options. Cathy commented that there was a "good range to choose from." Flora raised the question of whether all the options are necessary. She also stated, "Sometimes I think maybe the patient needs to express, 'Well you know look, this is happening to me quite a lot.' You [therapist] need to know that, the rating. To me that would probably benefit the therapist to understand what's going on in this person's makeup in their lives."

On the other hand, health professionals offered different opinions on the appropriateness of the response options. Jessica commented, "The frequency of the symptoms [is] the most important thing in terms of the effect that the patient [is] experiencing, matches perfectly what their experience is." For her, the response options and the items were equally important in terms of appropriateness. Kelvin related the fact that five response options is common. However, he questioned if seven response options could be more differentiating and stated, "Five is very common but I suspect that it won't allow you to tease it out quite as much." Kelvin explained that he was more familiar with other validated questionnaires (for assessing other respiratory conditions) with seven response options. In contrast, Margo explained, "I tend to go for…down to three questions '*Rare*, *Sometimes*, *Often*'." Although later in her interview she added, "*Never* is a good one to have". She also questioned, "*Often* versus *Very often* - what benefit are you getting between those two?" Jessica and Leena both disclosed that they would use the questionnaire without providing a hard copy for the patient. Jessica preferred to show patients the list of items while Leena preferred to read the items and response options out for patients, without them seeing the number corresponding to each option. They both wanted their patients to consider, on their own, what the options mean to them. Furthermore, Leena would sometimes relinquish the actual questionnaire altogether by using a blank piece of paper when she repeats the measure with a patient. She believed that could enhance objectivity for both patient and test administrator.

4.2 Internal Construct Validity: Rasch Analysis Results

4.2.1 Nijmegen Questionnaires

In total, 489 patient clinical files were searched. Data from 239 Nijmegen Questionnaires were located and extracted from these files and their data included in the study. The sample size was constrained by factors such as storage location, the cost involved in extraction, and missing questionnaires from patient files.

There were a total of 159 individuals who completed the 239 questionnaires. Each individual completed at least one questionnaire, with some completing more than one (with a maximum of five). Table 4.10 presents the gender, ethnicity, and age of the 159 individuals who completed the questionnaires. The mean age was 51 with a standard deviation of 16 (range 15-90). Table 4.11 presents the characteristics of the 239 questionnaires.

Table 4.10
Participant Characteristics

	Gender	
	Frequency	Percentage
Female	114	72%
Male	45	28%
	Ethnicity	
	Frequency	Percentage
NZ European	66	42%
Asian	41	26%
Pacific Islander	16	10%
Māori	13	8%
Other	23	14%
	Age	
	Frequency	Percentage
15-46	61	38%
47-57	44	28%
>57	54	34%

Note. NZ = New Zealand.
159 individuals completed the questionnaires.

Table 4.11

Questionnaire Characteristics

	Gender	
	Frequency	Percentage
Female	175	73%
Male	64	27%
	Ethnicity	
	Frequency	Percentage
NZ European	99	41%
Asian	66	28%
Pacific Islander	25	11%
Māori	20	8%
Other	29	12%
	Age	
	Frequency	Percentage
15-46	97	40%
47-57	66	28%
>57	76	32%

Note. NZ = New Zealand.

239 completed Nijmegen Questionnaires.

4.2.2 Rasch Analysis Results

The first step of analysis determined if the Nijmegen Questionnaire data collected satisfy the Rasch model expectations. The initial analysis included all the data (239 completed questionnaires) collected and all 16 items that made up the Nijmegen Questionnaire. Table 4.12 shows the distribution of response frequencies of this data set. Most people completed each item, with one or two questions being left blank for nine items. A floor effect was observed for 12 items (i.e. >25% of people scoring 0 = *Never*). There was no ceiling effect.

Table 4.12

Distribution of Response Frequencies of the Nijmegen Questionnaire

		Response categories					
		Never (0)	Rare (1)	Sometimes (2)	Often (3)	Very often (4)	Missing
Item	**Description**	**Freq (%)**	**Freq (%)**	**Freq (%)**	**Freq (%)**	**Freq (%)**	**Freq (%)**
NQ1	Chest pain	79 (33.1)	47 (19.7)	65 (27.2)	27 (11.3)	20 (8.4)	1 (0.4)
NQ2	Feeling tense	29 (12.1)	24 (10.0)	87 (36.4)	60 (25.1)	37 (15.5)	2 (0.8)
NQ3	Blurred vision	96 (40.2)	39 (16.3)	57 (23.8)	32 (13.4)	15 (6.3)	-
NQ4	Dizzy spells	65 (27.2)	40 (16.7)	76 (31.8)	40 (16.7)	17 (7.1)	1 (0.4)
NQ5	Feeling confused	94 (39.3)	51 (21.3)	52 (21.8)	24 (10.0)	18 (7.5)	-
NQ6	Faster or deeper breathing	40 (16.7)	41 (17.2)	77 (32.2)	48 (20.1)	32 (13.4)	1 (0.4)
NQ7	Short of breath	45 (18.8)	33 (13.8)	78 (32.6)	49 (20.5)	33 (13.8)	1 (0.4)
NQ8	Tight feelings in chest	62 (25.9)	40 (16.7)	67 (28.0)	38 (15.9)	31 (13.0)	1 (0.4)
NQ9	Bloated feeling in stomach	67 (28.0)	35 (14.6)	65 (27.2)	36 (15.1)	36 (15.1)	-
NQ10	Tingling fingers	94 (39.3)	42 (17.6)	55 (23.0)	24 (10.0)	22 (9.2)	2 (0.8)
NQ11	Unable to breathe deeply	80 (33.5)	42 (17.6)	55 (23.0)	34 (14.2)	26 (10.9)	2 (0.8)
NQ12	Stiff fingers or arms	99 (41.4)	40 (16.7)	47 (19.7)	27 (11.3)	26 (10.9)	-
NQ13	Tight feelings around mouth	153 (64.0)	38 (15.9)	25 (10.5)	11 (4.6)	11 (4.6)	1 (0.4)
NQ14	Cold hands or feet	81 (33.9)	32 (13.4)	46 (19.2)	33 (13.8)	47 (19.7)	-
NQ15	Palpitations	63 (26.4)	44 (18.4)	82 (34.3)	30 (12.6)	20 (8.4)	-
NQ16	Feeling of anxiety	35 (14.6)	38 (15.9)	72 (30.1)	51 (21.3)	43 (18.0)	-

Note. Freq = frequency

First stage of analysis: Testing of the overall fit

Initial fit of data to the Rasch model showed that one item (NQ14 Cold hands or feet) had an item fit residual of 4.577, which is outside the acceptable range of +/-2.5, indicating misfit. The mean item fit residual was 0.410 with a standard deviation of 1.499. In addition, the item-trait interaction chi-square was significant with probability of < 0.001, indicating lack of fit to the Rasch model (Table 4.13).

The next step of analysis examined the correlations between the residuals in order to identify any local dependency among the 16 items. In the literature, Marais and Andrich (2008) recommend that residual correlations should be smaller than 0.2 above the average residual correlation (in this case -0.063 + 0.2 = 0.137). High correlations between the residuals were identified, indicating local dependency between a number of items (Table 4.14). This suggested that the item responses of the Nijmegen Questionnaire depend not only on the severity of the symptoms of hyperventilation syndrome being measured but on responses to other items on the questionnaire.

The following step was to explore unidimensionality of the questionnaire. Items with high positive and high negative loadings on the first principal component of the residuals were combined into testlets and *t*-tests were carried out to examine if they were significant. Independent *t*-test on the two estimates derived from the subtests were carried out for each respondent. The scale is deemed unidimensional if the lower bound of the 95% confidence interval (CI) is smaller than 5% (i.e. fewer than 5% of tests are significant). It was found that 5.1% of t-tests were significant (95% CI 2.3% to 7.8%; lower bound of 95% CI 2.3%), suggesting the scale is unidimensional.

Table 4.13

Summary of Fit Statistics of the Nijmegen Questionnaire to the Rasch Model

	Item fit residual		Person fit residual		Chi-square interaction		PSI		α		Tests of unidimensionality 95% CI (%)		
Analysis	Mean	SD	Mean	SD	Value (df)	*p*	With extremes	Without extremes	With extremes	Without extremes	Significant *t*-test	Lower bound	Upper bound
First stage[a]	0.410	1.499	-0.265	1.609	109.4 (48)	0.000	0.889	0.880	0.898	0.890	5.1	2.3	7.8
Second stage[b]	0.385	1.145	-0.310	1.575	67.8 (45)	0.016	0.888	0.879	0.899	0.891	5.5	2.7	8.3
Third stage[c]	0.064	0.971	-0.215	1.202	41.9 (45)	0.604	0.844	0.826	0.893	0.869	5.8	2.9	8.6
Fourth stage[d]	0.063	0.858	-0.211	1.015	36.1 (45)	0.205	0.814	0.789	0.839	0.809	1.8	1.1	4.6

Note. SD = standard deviation; df = degrees of freedom; *p* = probability; PSI = Person Separation Index; α = alpha.

[a]Fit to the Rasch model of all 16 items. [b]Fit to the Rasch model after deleting item NQ14. [c]Fit to the Rasch model after rescoring response categories for items with disordered thresholds. [d]Fit to Rasch model after merging of items.

Table 4.14

Summary of Local Dependencies of the Nijmegen Questionnaire

Analysis	Individual item		Local dependency	
	Item	Description	Item	Description
First[a] stage and Second[b] stage	1	Chest pain	8	Tight feelings in chest
	2	Feeling tense	5, 16	Feeling confused, Feeling of anxiety
	3	Blurred vision	4	Dizzy spells
	6	Faster or deeper breathing	7	Short of breath
	7	Short of breath	11	Unable to breathe deeply
	10	Tingling fingers	12	Stiff fingers or arms
Third[c] stage	1	Chest pain	8	Tight feelings in chest
	2	Feeling tense	5, 16	Feeling confused, Feeling of anxiety
	6	Faster or deeper breathing	7	Short of breath
	10	Tingling fingers	11, 12	Unable to breathe deeply, Stiff fingers or arms
Fourth[d] stage	No local dependency			

Note. [a]Fit to the Rasch model of all 16 items. [b]Fit to the Rasch model after deleting item NQ14. [c]Fit to the Rasch model after rescoring response categories for items with disordered thresholds. [d]Fit to Rasch model after merging of items.

Item characteristic curves were then examined for all 16 items. The item characteristic curve has been described as a hypothetical illustration of the Rasch model (Kersten & Kayes, 2011). An expectation of the Rasch model is that the item response options reflect the level of symptoms of individuals being measured (i.e. the higher the level of symptoms, the higher the probability that the item would be scored). NQ14 (Cold hands or feet) was identified earlier as a problematic item given its high positive item fit residual 4.577 (outside the desired range of +/-2.5) and a statistically significant chi-square (p = 0.003125) (lower than the Bonferroni adjusted p value of 0.0032). This was reflected in the item response curve for item NQ14 of the Nijmegen Questionnaire (Figure 4.1).

Figure 4.1 Item Response Curve for item NQ14 Cold hands or feet

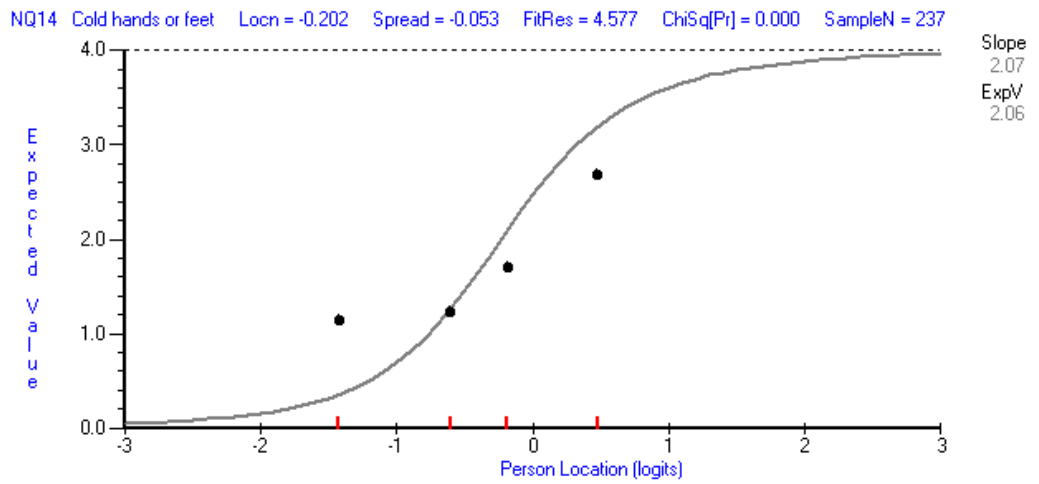
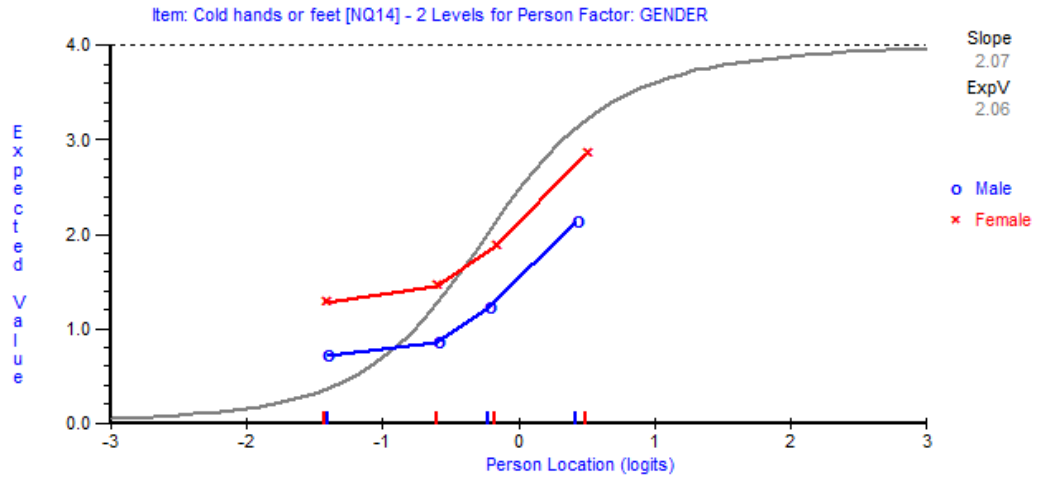


Figure 4.1 shows the expected value (i.e. expected raw scores on an item of a Nijmegen Questionnaire; ranged from 0 to 4) on the y-axis and the person location (i.e. person parameters estimates) in logits on the x-axis. The s-shape grey curve represents expected raw scores associated with the person location scores in logits (generated from Rasch analysis). The black dots show the probability associated with the four class intervals (four groups of respondents grouped by their symptoms of hyperventilation syndrome). If you were to draw a curve using the four dots, it appears to be similar to the grey curve but is flatter indicating a level of deviation from the Rasch model expectation. Along with the high positive item fit residual of 4.577, it was concluded that item NQ14 (Cold hands or feet) was under discriminating. This suggested that the item discriminates between people with mild or severe hyperventilation syndrome less than expected for an item of this difficulty. For the remaining 15 items of the Nijmegen Questionnaire, the black dots representing the probability associates with the class intervals were all close to their respective item characteristic curves hence, these items demonstrated good fit to the Rasch model (as also indicated by their item fit residuals which were within the range -/+ 2.5 and non-significant chi-square).

Next, Differential Item Functioning (DIF) for the 16 items was tested. All items were examined for DIF across all person variables. If an item displays DIF, it indicates the item is not invariant across various groups of individuals (Kersten, White, & Tennant, 2012). Consequently, the item is considered not to fit the Rasch model, which has the expectation that all items are unbiased across major population groups (e.g. age, gender, and ethnic groups). DIF analysis showed that item NQ14 displayed uniform DIF by gender (Figure 4.2). It is clear that the probability of reporting cold hands or feet is much lower for male than female individuals despite having the same severity of hyperventilation related symptoms. The remaining 15 items were invariant across different person factors, with stable item difficulty across time.

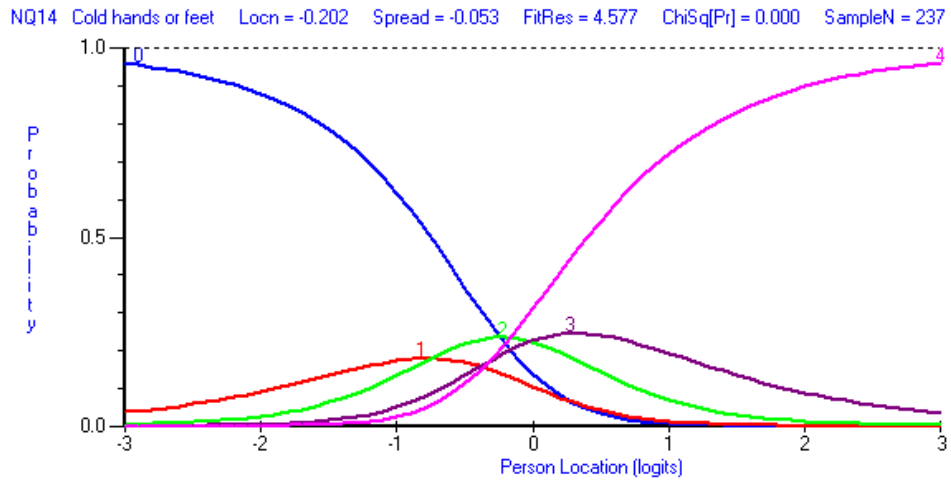
Figure 4.2 Differential Item Functioning for item NQ 14 Cold hands or feet



The final step in analysing all 16 items was to examine the category probability curves, determining if the item response categories were working as intended. The log-transformed item scores are generated from individuals' responses to the item response options. This applies to outcome measures with polytomous data (i.e. items with more than two response options). If the response options are reflecting the increasing or decreasing symptoms of hyperventilation syndrome, we would expect items to display ordered thresholds. "Thresholds are the points where the probabilities of a response of either 0 or 1, and 1 or 2 (and so forth) are equally likely" (Kersten & Kayes, 2011, p. 95). They are the transition points between item response options.

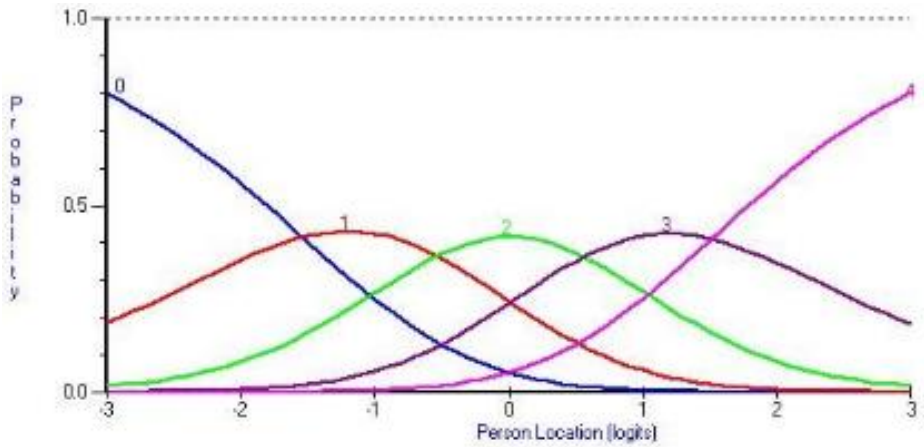
Figure 4.3 shows the category probability curves for item NQ14 with item response options range from 0 to 4 (0 = not experiencing symptom; 4 = experiencing the symptom very often). The *y*-axis represents the probability that someone selects an item response option (i.e. an individual marks 0, 1, 2, 3, or 4) for an item in relation to his/her overall level of symptoms (i.e. person location), which is displayed on the *x*-axis.

Figure 4.3 Category Probability Curves for NQ14 Cold hands or feet



In Figure 4.3, the locations of the thresholds points between 0 and 1, 1 and 2, and so forth are not in the expected order (e.g. the transition point between 0 and 1 is lower than that of 1 and 2). In theory, the transition point between 1 and 2 represents a higher level of symptoms and should be below 0 and 1. This indicates that as the frequency of cold hands or feet increases, there is not ever a time at which the item response options 1 and 2 are most likely. The remaining 15 items also have disordered thresholds. As an example, Figure 4.4 shows the category probability curves for a hypothetical item with ordered thresholds (Polytomous Rasch Model, n.d.).

Figure 4.4 Category Probability Curves for a Hypothetical Item with Ordered Thresholds

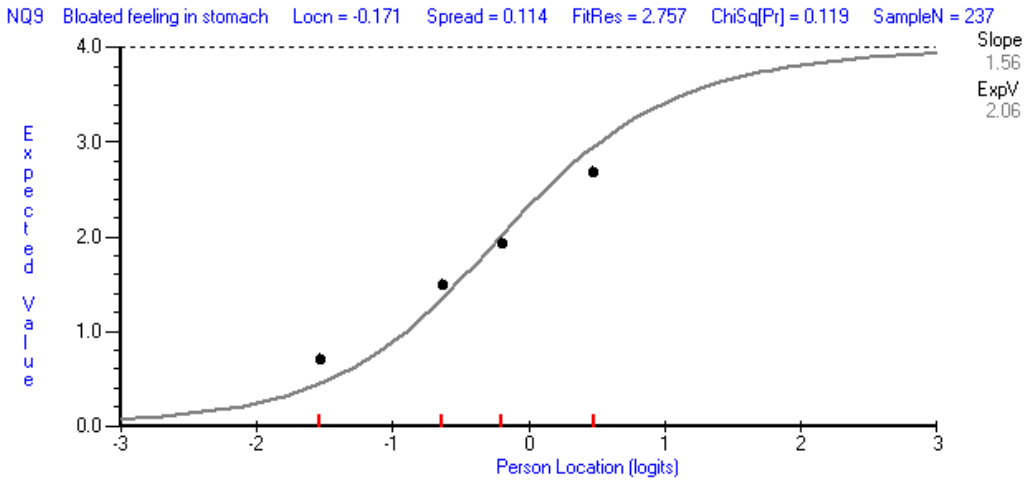


Second stage of analysis: Deleting of item NQ14 Cold hands or feet

Since item NQ14 did not fit the Rasch model (fit residual 4.577), it was deleted for the second stage of data analysis. Data collected from the remaining 15 items were analysed using the same steps and criteria described above. Item NQ9 (Bloated feeling in stomach) had an item fit residual of 2.757 which is just outside the acceptable range of +/-2.5. The mean item fit residual was 0.385 (expected value of 0) with a standard deviation of 1.145 (acceptable value should be < 1.4). The item-trait interaction chi-square was not significant with probability of 0.016 (which is greater than the Bonferroni adjusted *p* value of 0.0033), indicating fit to the Rasch model. Examination of the correlations between the residuals revealed local dependency between the same clusters of items identified from the first stage of analysis (i.e. NQ1 and NQ8; NQ2, NQ5, and NQ16; NQ3 and NQ4; NQ6 and NQ7; NQ7 and NQ11; NQ10 and NQ12). The 15-item scale remained unidimensional after deleting item NQ14 as demonstrated by the proportion of significant *t*-tests (95% CI 2.7% to 8.3%).

Figure 4.5 displays the item response curve for NQ9 (Bloated feeling in stomach), which had an item fit residual of 2.757 and a non-significant chi-square of 5.852 ($p = 0.119$). The observed scores for the class intervals or groups are represented by the black dots, deviated from the expected response pattern for this item, suggesting some deviation from the Rasch model expectation. This indicated that NQ9 was under discriminating, though not to the extent NQ14 was earlier. The remaining 14 items were then inspected and all demonstrated good fit to the Rasch model.

Figure 4.5 Item Response Curve for NQ9 Bloated feeling in stomach



All 15 items were invariant (i.e. unbiased, no DIF) across different age, gender, and ethnic groups, at initial and repeated assessment(s). However, as with the 16-item scale, all items had disordered thresholds.

Third stage of analysis: Rescoring of items

The next process was to address the disordered thresholds, guided by the output generated from the first two stages of analysis. There is no set guideline for collapsing response options in items with disordered thresholds. Nonetheless, the response options being collapsed needed to make sense and a uniform frequency distribution (i.e. collapsing *Sometimes* with *Rare* instead of with *Often*) is preferred (Boone et al., 2014). The Nijmegen Questionnaire has five possible response options and there are a few possible collapsing strategies. Table 4.15 shows the collapsing strategies that were chosen and applied through the testing process. They are labelled in order of the rescoring process.

Table 4.15

Rescore Strategy for Response Categories of the Nijmegen Questionnaire

Strategy	**Never (0)**	**Rare (1)**	**Sometimes (2)**	**Often (3)**	**Very Often (4)**
4 response options					
1st rescore	0	1	1	2	3
2nd rescore	0	0	1	2	3
3 response options					
3rd rescore	0	0	1	1	2

Note. Rescore strategies for the five response options of the Nijmegen Questionnaire. Response otions 1 and 2 were collapsed in the first rescoring process for 15 items.

Response options 1 and 2 were collapsed in the first rescoring process for 15 items. The thresholds map showed that disordered thresholds remained for 10 items (Figure 4.6). Response options 0 and 1 were then collapsed instead in the second rescoring process (all 15 items). With this change, there were only four remaining items (NQ9, NQ10, NQ13, and NQ15) showing disordered thresholds (Figure 4.7). NQ9, NQ10, and NQ13 appeared to have more problematic disordered thresholds (Figure 4.8, 4.9, and 4.10) when compared with NQ15. Category probability curves for NQ15 showed only borderline issue with disordered thresholds (Figure 4.11). The decision was to collapse response option 1 and 2 in the third rescoring process for item NQ9, NQ10, and NQ13. Ordered thresholds were achieved for all 15 items (Figure 4.12). Table 4.16 presents the available response options for all 15 items after the rescoring process.

Figure 4.6 Threshold Map after First Rescoring

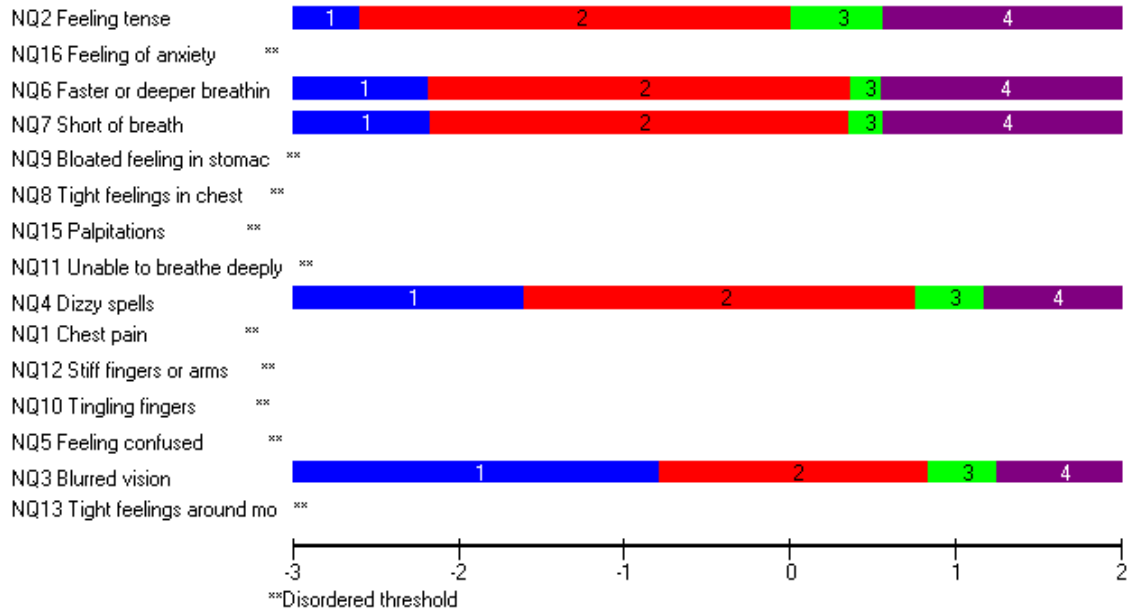


Figure 4.6. The easiest item to endorse is at the top of the graph and the hardest item to endorse is at the bottom.

Figure 4.7 Threshold Map after Second Rescoring (response options Never and Rare collapsed)

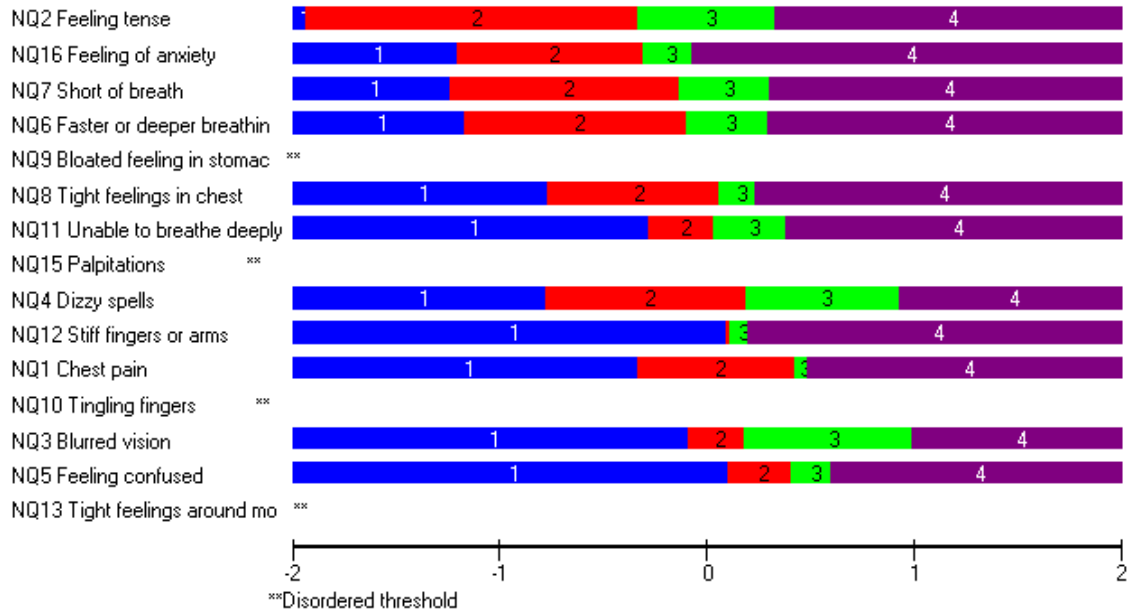


Figure 4.7. The easiest item to endorse is at the top of the graph and the hardest item to endorse is at the bottom.

Figure 4.8 Category Probability Curves for NQ9 Bloated feeling in stomach after Second Rescoring (response options Never and Rare collapsed)

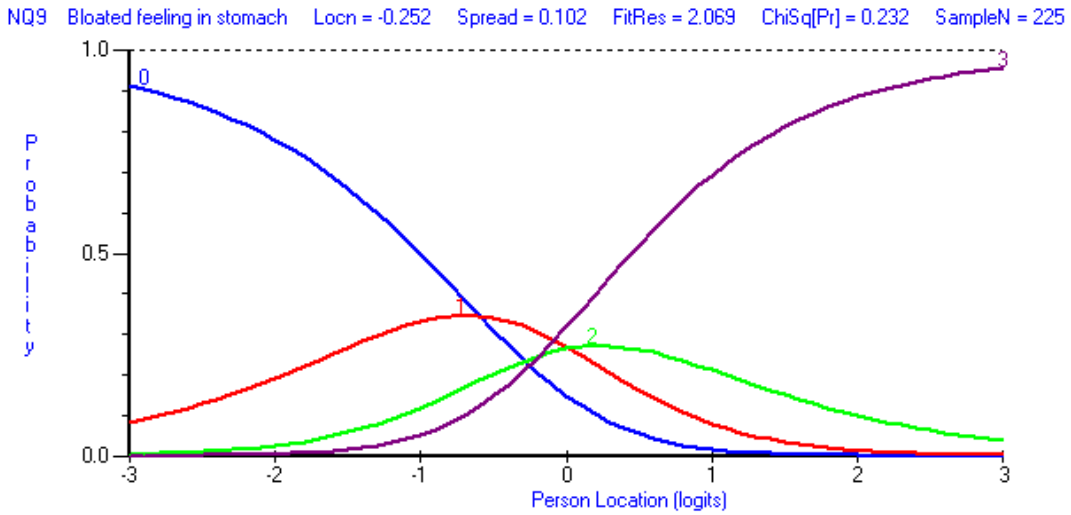


Figure 4.9 Category Probability Curves for NQ10 Tingling fingers after Second Rescoring (response options Never and Rare collapsed)

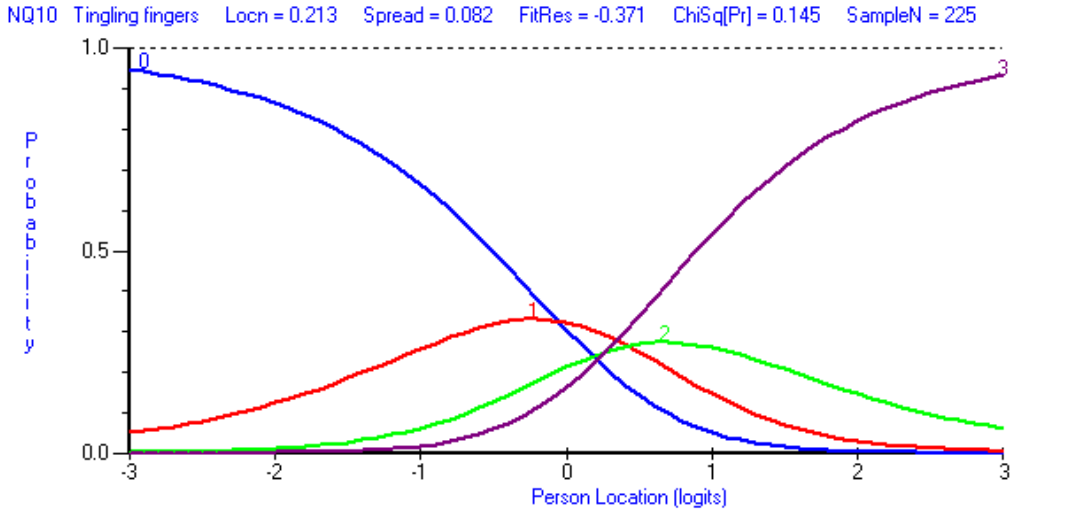


Figure 4.10 Category Probability Curves for NQ13 Tight feelings around mouth after Second Rescoring (response options Never and Rare collapsed)

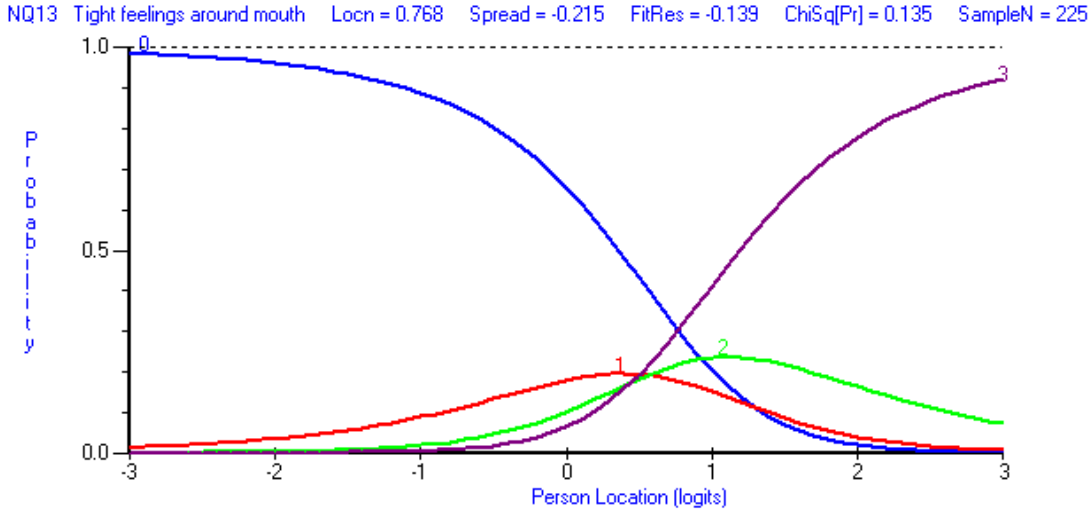


Figure 4.11 Category Probability Curves for NQ15 Palpitations after Second Rescoring (response options Never and Rare collapsed)

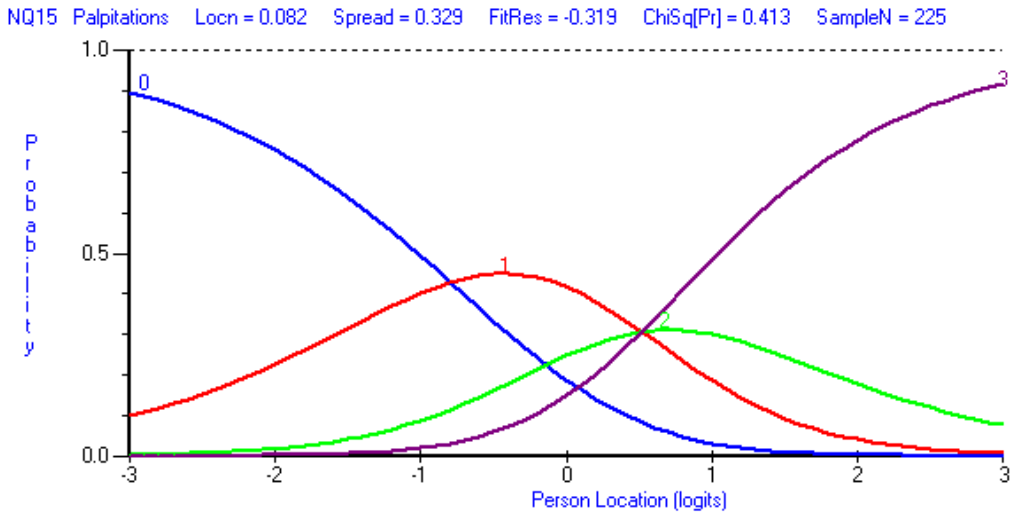


Figure 4.12 Threshold Map after Third Rescoring

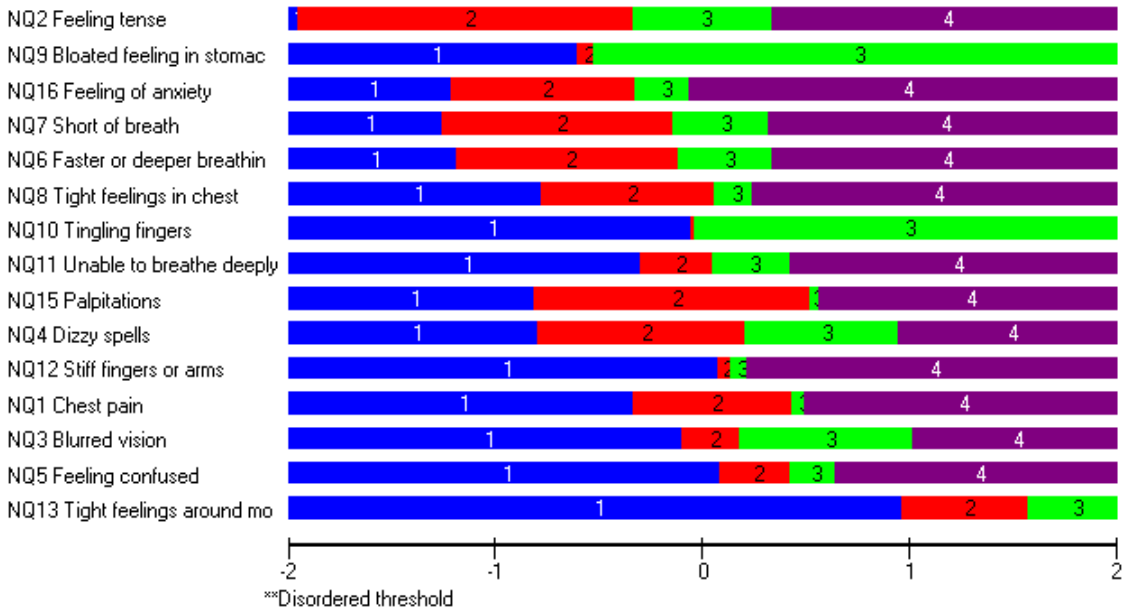


Figure 4.12. The easiest item to endorse is at the top of the graph and the hardest item to endorse is at the bottom.

Table 4.16

Response Options after Rescoring

Item	Maximum Score/Item	0	1	2	3	4
NQ1	3	0	0	1	2	3
NQ2	3	0	0	1	2	3
NQ3	3	0	0	1	2	3
NQ4	3	0	0	1	2	3
NQ5	3	0	0	1	2	3
NQ6	3	0	0	1	2	3
NQ7	3	0	0	1	2	3
NQ8	3	0	0	1	2	3
NQ9	2	0	0	1	2	2
NQ10	2	0	0	1	2	2
NQ11	3	0	0	1	2	3
NQ12	3	0	0	1	2	3
NQ13	2	0	0	1	2	2
NQ15	3	0	0	1	2	3
NQ16	3	0	0	1	2	3
New Total	42					

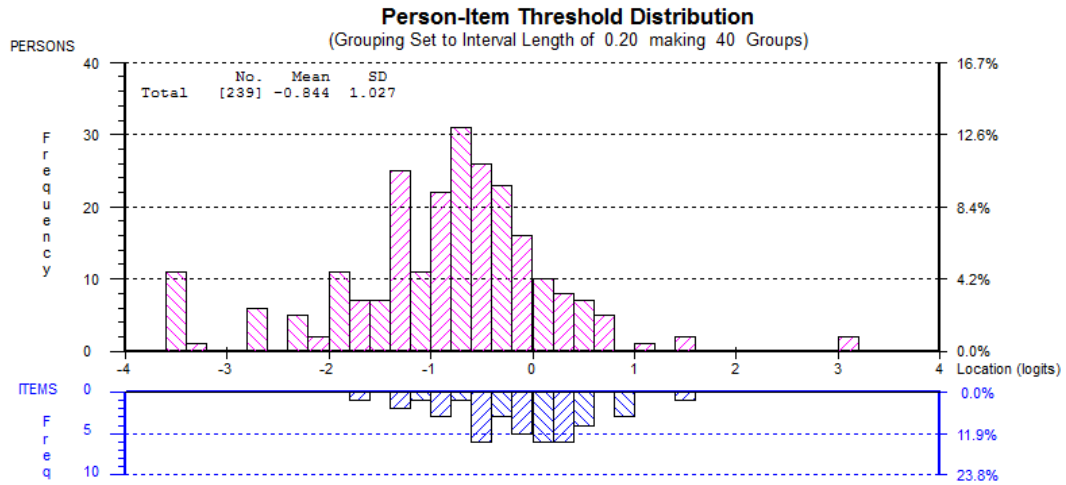
Following the rescoring process, the summary statistics were reviewed (Table 4.13). The individual item residuals were all within the range of +/-2.5 (including item NQ9 which previously had a fit residual > 2.5). The average fit residual statistics had a mean value of 0.064 and standard deviation of 0.971. The item-trait interaction chi-square probability was 0.604 which was non-significant. The scale remained unidimensional after the rescoring procedure with 5.8% of *t*-tests being significant with the lower bound of the 95% CI being 2.9% (95% CI 2.9% to 8.6%). The inspection of correlations between the residuals resulted in a reduced number of local dependencies and they were NQ1 and NQ8; NQ2, NQ5, and NQ16; NQ6 and NQ7; NQ10, NQ11 and NQ12. The next step in the process was to address the local dependencies as they can have a considerable influence upon fit (Lundgren Nilsson, & Tennant, 2011).

Fourth stage of analysis: Merging of items

The following items were merged to create subtests, essentially combining locally dependent items into new super items, thus removing the influence of local dependencies: NQ1 (Chest pain) and NQ8 (Tight feelings in chest); NQ2 (Feeling tense), NQ5 (Feeling confused), and NQ16 (Feeling of anxiety); NQ10 (Tingling fingers) and NQ12 (Stiff fingers or arms). While NQ10 (Tingling fingers) and NQ 11 (Unable to breathe deeply) also demonstrated local dependency, they were incompatible for merging considering the different nature of the two symptoms. Three new testlets were created for these super items in the subtest analysis, with the rest of the items remaining as individual items. Following this step, the average fit residual statistics had a mean of 0.063 and standard deviation of 0.858. The item-trait interaction chi-square probability was non-significant at 0.205. The subtest analysis showed that only 1.8% of *t*-tests were significant and that therefore the scale was unidimensional. This final version of the scale fits the Rasch model expectations (Table 4.13). With a Person Separation Index of 0.79, it is able to make distinctions between three groups of people with differing levels of severity of symptoms (Fisher, 1992).

Below is the person-item threshold distribution graph (Figure 4.7) of the 15-item Nijmegen Questionnaire (after rescoring and creation of subtests). The distribution of persons is displayed at the upper half of the graph and at the lower half is the distribution of items. All are mapped onto the logarithmic interval scale with a mean of 0 at the centre of the graph.

Figure 4.13 Person-Item Threshold Distribution for the Final Analysis



An outcome measure requires a range of items that spread along the construct being assessed according to their various levels of difficulty. Figure 4.13 demonstrates a good spread of item thresholds along the construct (spanning 3.5 logits). The majority of people are scoring at the middle and very few at the top of the scale. The mean person location is -0.844 logits, which suggests that the questionnaire is considerably well targeted to assess the frequency of hyperventilation related symptoms experienced by individuals. A conversion table (Table 4.17) was produced after this final stage of analysis, allowing the raw ordinal scores to be converted into interval equivalent scores (provided item NQ14 is deleted and the rescoring of remaining items is carried out as summarised in Table 4.16).

Table 4.17
Conversion Table for the Nijmegen Questionnaire

Raw total score	Logit	Interval score
0	-3.438	0.00
1	-2.71	4.62
2	-2.234	7.64
3	-1.923	9.62
4	-1.69	11.10
5	-1.502	12.29
6	-1.344	13.30
7	-1.207	14.17
8	-1.085	14.94
9	-0.975	15.64
10	-0.875	16.27
11	-0.782	16.86
12	-0.696	17.41
13	-0.616	17.92
14	-0.54	18.40
15	-0.469	18.85
16	-0.4	19.29
17	-0.334	19.71
18	-0.27	20.11
19	-0.208	20.51
20	-0.148	20.89
21	-0.088	21.27

Note. This table allows the conversion of the raw ordinal total scores to the equivalent interval total scores, when using the Nijmegen Questionnaire; for people without missing data.

Table 4.17

Conversion Table for the Nijmegen Questionnaire (Continued)

Raw total score	Logit	Interval score
22	-0.029	21.64
23	0.03	22.02
24	0.089	22.39
25	0.149	22.77
26	0.21	23.16
27	0.273	23.56
28	0.338	23.97
29	0.406	24.41
30	0.477	24.86
31	0.554	25.35
32	0.637	25.87
33	0.726	26.44
34	0.826	27.07
35	0.939	27.79
36	1.068	28.61
37	1.221	29.58
38	1.407	30.76
39	1.642	32.25
40	1.959	34.27
41	2.441	37.33
42	3.177	42.00

Note. This table allows the conversion of the raw ordinal total scores to the equivalent interval total scores, when using the Nijmegen Questionnaire; for people without missing data.

4.3 Summary of Interview and Rasch Analysis Results

Symptoms of hyperventilation syndrome were identified crossing three categories and 12 sub-categories, representing 46 symptoms/symptom clusters. The majority of the Nijmegen Questionnaire items mapped onto interview-identified symptoms either in full or partially. However, a number of symptoms/symptom clusters identified in the interview data were not well captured by the Nijmegen Questionnaire. Rasch analysis resulted in a 15-item unidimensional scale which was unbiased across the population and across different time point (after rescoring thresholds). A conversion table was made available for converting ordinal raw scores to interval scores, allowing the utilisation of parametric statistical techniques as indicated.

Chapter Five: Discussion and Conclusion

The aim of this measurement study was to investigate the content and internal construct validity of the Nijmegen Questionnaire for hyperventilation syndrome. The research question for the study was: Is the Nijmegen Questionnaire a valid outcome measure for individuals with hyperventilation syndrome?

In the literature, limited evidence was found with regards to the content and internal construct validity of the Nijmegen Questionnaire. The current study investigated these two psychometric properties, drawing on guidelines for outcome measure development and testing. This study has contributed new knowledge to the literature base since prior research had not explored content validity of the questionnaire by involving patients with hyperventilation syndrome and the internal construct validity had not been explored outside of the original work of the developers or using Rasch analysis.

In this chapter, I will discuss study findings in the context of existing evidence and discuss the implications for the use of this questionnaire. Strengths and limitations of the study will be reviewed along with recommendations for future research.

5.1 Content Validity of the Nijmegen Questionnaire

> "Content validation is the technical term for ensuring that the scale has enough items and adequately covers the domain under investigation." (Streiner & Norman, 2008, p. 24)

The current study demonstrated that 28 out of the 46 symptoms/symptom clusters identified by participants are not captured by the Nijmegen Questionnaire. The percentage of non-captured items for breathing symptoms, psychological symptoms, and physical symptoms was 67%, 75%, and 56% respectively. While these percentage values appeared high, we also need to consider 94% of items (15/16) in the Nijmegen Questionnaire were fully or partially congruent with symptoms/symptom clusters derived from the qualitative data. The only item that did not map onto interview data was NQ12 Stiff fingers or arms. Therefore, this study shows that the current items in the Nijmegen Questionnaire are representative of patients' and health professionals' view on symptoms of hyperventilation syndrome, though perhaps not fully. The original list of 45 symptoms mentioned by van Doorn and colleagues (1982) which informed the development of the Nijmegen Questionnaire is not available. However,

two additional lists of complaints/symptoms have been published by others (Grossman & de Swart, 1984; Hornsveld et al., 1990). These lists were published in the manuscripts' method sections as part of an outcome measure for symptoms of hyperventilation syndrome. These can be compared with the current qualitative findings for content validity evaluation. Hornsveld et al. (1990) named their list the Bodily Sensation Questionnaire (BSQ), including 32 symptoms often reported by patients with hyperventilation syndrome (Ruiter, Garssen, Rijken, & Kraaimaat, 1989). On the other hand, Grossman and de Swart did not provide references or a name for their list of 37 symptoms. While establishing content validity of the lists was not the aim for the two studies, these lists provided a wider context of complaints/symptoms than the original 16 Nijmegen Questionnaire items when considering its conceptual basis in relation to the study findings.

Twenty four symptoms/symptom clusters identified from this study are consistent with the majority of the symptoms from the above two lists. Seven of these are not currently captured by the Nijmegen Questionnaire: (1) upper chest breathing, (2) air hunger, (3) uneasy/feeling different/not feeling so good/something is always at the back of your mind, (4) feeling hot/sweaty, (5) tiredness, (6) headache, and (7) frowning/facial expression. There are seven complaint items from the lists that do not match with our qualitative study findings: (a) shivering, (b) trembling, (c) hand tremble, (d) stiffness in fingers or arms, (e) stiffness in legs, (f) nausea, and (g) fits of crying. However, stiffness in fingers or arms (which is in these lists) is captured by the Nijmegen Questionnaire. Two symptom categories (breathing symptoms and psychological symptoms) identified from this study are conceptually congruent with two of the four categories within the BSQ (Hornsveld et al., 1990; Ruiter et al., 1989). Difference in language use in the interviews and the lists was noted (i.e. *breathing* versus *respiratory*), demonstrating a contrast between layman and technical language used by patients/health professionals and researchers. The differences between the study findings and the existing lists of symptoms could be due to the distinct perspectives between patients/health professionals and researchers. Further, the different methodology and methods used in compiling the list/questionnaire could impact on symptom reporting and recording. Practitioners and researchers can vary in how the questionnaire is administered to meet certain needs or expectations that are relevant to their practice. This was certainly reflected by the feedback provided by the health professionals in the study, with regards to how they guide patients through the available response options on the Nijmegen Questionnaire.

Overall, the current items of the Nijmegen Questionnaire capture 16 of the 24 hyperventilation related symptoms that were acknowledged by the two groups of authors (Hornsveld et al., 1990; Ruiter et al., 1989) and by the current study findings. Test developers tend to have far more symptoms/topics than what could be deemed appropriate for an outcome measure and it is up to the researcher to decide whether the retention or elimination of certain items could impact on the content validity of the measure (Bowling, 2009; Streiner & Norman, 2008). The risk of adding too many items entails duplication or making a measure unnecessarily long. If symptoms that are most often reported by patients are omitted while the items identified by professionals are retained by test developers, it could be perceived as bias towards professional opinion. Careful consideration is required when determining whether adding or removing of items will improve the overall content validity and reliability/internal consistency of the questionnaire. Test developers can be guided by statistical criteria pertaining to internal consistency of the scale when deciding on the length of the questionnaire (Streiner & Norman, 2008). The Cronbach's alpha coefficient is the most commonly used indicator of internal consistency. This statistic is sensitive to the number of items in the scale, with too few items resulting in an insensitive scale (e.g. alpha below acceptable values), and too many items no longer improving the Cronbach's alpha (Pallant, 2016; Streiner & Norman, 2008).

Current study findings also illuminated the complexity of symptoms of hyperventilation syndrome. Previously, one item (anxiety) was omitted during the validation study by van Dixhoorn and Duivenvoorden (1985). The authors stated that anxiety "should be seen rather as a cause or background feeling than as being inherent in the symptoms of HVS [hyperventilation syndrome]" (p. 200). Essentially, the question they raised is the difficulty for patients or health professionals to distinguish between psychological symptoms (which might characterise hyperventilation syndrome) versus the emotional response to symptoms or to an underlying psychological disorder that is undiagnosed. Study participants clearly articulated the importance of psychological symptoms in contributing to the overall presentation of hyperventilation syndrome, potentially contradicting the statement above by Dixhoorn and Duivenvoorden (1985). Practitioners who are making interpretations based on the Nijmegen Questionnaire should consider the following: Given that many symptoms identified by patients and health professionals are not currently captured by the Nijmegen Questionnaire, it should be used in conjunction with other subjective and objective assessments; While the cultural validity of the questionnaire was not tested in the current study, it is possible the variations in the language used to describe symptoms may become more accentuated in people with English as a second language.

Moreover, there was disagreement among health professionals on the appropriate number of response options (i.e. the preference of having more versus less response options). There is no other study that evaluated the appropriateness of the response options of the Nijmegen Questionnaire in the literature. However, there are conflicting general opinions on response options in the literature. For example:

> First, the person may select the first response option that seems reasonable. If the options are presented in written form, the respondent will opt for the one that occurs first, and give only cursory attention to those which appear later. Conversely, if the options are given verbally, it is easier to remember those which came most recently, that is, those towards the end of the list. In either case, the choices in the middle get short shrift. (Streiner & Norman, 2008, p. 108)

The above example suggests increasing the number of response options could be counterproductive in achieving overall perceptibility and differentiability of these options in the eyes of the respondents. In contrast, others have argued that "a greater number of categories [options] increases perceptions of variety, greater perceptions of variety increases self-determination, and greater self-determination increases … satisfaction" (Mogilner, Rudnick, & Iyengar, 2008). The differences between patient and health professional perspective on symptoms presentation, symptom reporting, and the appropriateness of the response options warrant further investigation as study findings clearly demonstrated gaps in existing knowledge.

On balance, the study findings show the Nijmegen Questionnaire is adequate at measuring level of severity of symptoms (in that it appears to contain enough items in relative terms to represent the symptoms of hyperventilation syndrome), but not necessarily the impact of the symptoms on the person. The latter requires a different measure to be used. Readers need to be aware of symptoms identified by study participants that are not captured by the questionnaire. There are variances in symptom descriptions with regards to conceptual and language congruency which practitioners must acknowledge when using the Nijmegen Questionnaire and interpreting the scores produced.

5.2 Internal Construct Validity of the Nijmegen Questionnaire

> Structural/internal construct validity is "the degree to which the scores of a health-related patient-reported outcomes instrument are an adequate reflection of the dimensionality of the construct to be measured" (Terwee et al., 2012, p. 9).

The Rasch analysis findings from this study indicated that the 16-item Nijmegen Questionnaire did not fit the strict Rasch model expectations for internal construct validity. Specific issues identified included a poorly fit item (NQ14 Cold hands or feet), non-unidimensionality, and disordered item thresholds for all 16 items. NQ14 was biased in its function when assessing hyperventilation syndrome between male and female individuals. Another issue with this item was that it was under discriminating. In other words, the item discriminates between people with mild or severe hyperventilation syndrome less than expected for an item of this difficulty. This suggested that NQ14 Cold hands or feet is not a good item for measuring overall symptoms of hyperventilation syndrome. Another item (NQ9 Bloated feeling in stomach) was found to misfit to the Rasch model in the second stage of analysis, after the deletion of NQ14. While NQ09 was also an item that under discriminates, there was no bias in terms of item function found in any person variables. Unidimensionality was achieved with the deletion of NQ14 but, retention of NQ9. This suggested that the item bloated feeling in stomach was a valid symptom in assessing hyperventilation syndrome. The systematic rescoring of response options and the merging of items with resonating meanings into testlets resulted in the final 15-item version of the Nijmegen Questionnaire, meeting strict criteria for internal construct validity. This study was also able to produce a conversion table that can be used to convert ordinal raw scores to interval scores, when parametric testing is indicated.

The internal construct validity was previously evaluated by van Dixhoorn and Duivenvoorden (1985) using non-metric principal components analysis (NMPCA), a parametric statistical technique. The difference in analytical techniques is significant in the interpretation and comparison of study results. The first issue is that parametric statistical techniques are generally suited to interval data and are based on the weak expectations of classical test theory (Bowling, 2009; Streiner & Norman, 2008). Therefore, the results generated by van Dixhoorn and Duivenvoorden (1985) could not be compared directly with the current study findings. However, I will highlight prior results concerning construct validity that can be extrapolated and interpreted with current study findings.

Van Dixhoorn and colleagues (1982) identified three questionnaire components: shortness of breath, peripheral tetany (overly stimulated neuromuscular activity [Chaitow et al., 2002]), and central tetany. The identification of this underlying relationship between

variables was consistent with the discovery of local dependencies among the current items of the Nijmegen Questionnaire in this study. Upon inspecting the items within each component, some of the local dependencies identified from the current were noted within the shortness of breath and central tetany components. This suggests that both studies revealed, the symptoms represented by these items were scored not just based on the severity of hyperventilation syndrome related symptoms but on the score for another item on the scale also. The locally dependent items were representing symptoms of similar nature. Also, NQ16 (Feeling of anxiety) was omitted from the 1982 validation study. This item was found to be locally dependent with NQ2 (Feeling tense) and NQ5 (Feeling confused). Van Dixhoorn and colleagues' decision to omit NQ16 was not supported by the current study findings.

Item NQ12 (Stiff fingers or arms) did not match with any participant-identified symptoms. However, it was found to be locally dependent with item NQ10 (Tingling fingers) which was fully conceptually and linguistically congruent with symptom identified by participants. It is important to bear in mind the small sample size for the interviews. Regardless of the lack of reporting by study participants, the fact that NQ12 was locally dependent suggests it measures something very similar to NQ10, which was reported. After creating the testlets the 15-item version of the Nijmegen Questionnaire fit the Rasch model and was uniform.

Item NQ14 (Cold hands or feet) was only partly congruent with interview findings. In addition, it was a misfitting item as highlighted by the Rasch analysis which resulted in the deletion of NQ14. Thus, both the interview and Rasch analysis findings from this study supported a 15-item version of the Nijmegen Questionnaire as a valid screening tool for hyperventilation syndrome.

5.3 Strengths and Limitations of the Study

By involving both patients and health professionals, this study met the criteria for the evaluation of content validity as described by the Scientific Advisory Committee of the Medical Outcome Trust (2002). The COSMIN checklist identifies several criteria to assess the methodological quality of measurement studies (Terwee et al., 2012). A self-assessment of the current study suggested that it meets all the criteria identified as critical to content validity, achieving an *Excellent* rating (Table 5.1). The qualitative findings included both patients' and health professionals' perspectives on symptoms attributing to hyperventilation syndrome. Studies employing the Qualitative Descriptive Methodology are able to produce findings that

are transferable to population with similar characteristics as the study participants (Sandelowski, 1995, 2000).

Despite the various adjustments made in the effort to recruit male participants, there was a lack of male patient-participant in the qualitative component of this study and there was only one male participant in the health professional group. Therefore, the study findings regarding content validity have limited transferability to a male population. Although more women than men suffer from hyperventilation syndrome and more women are treated at the recruitment locality, it created a gender bias in our qualitative study findings. One health professional from the study articulated that male patients were usually less willing to acknowledge or share their health-related symptoms in general. This is supported by evidence in the literature, "men often are unwilling and lack the motivation to engage with health-related information", suggesting that women may be generally more engaged, involved, attentive, and better informed in health decision making (Ek, 2013, p. 742).

It was evident from the study that people's experiences of hyperventilation syndrome can be confounded by psychological symptoms. For example, item NQ10 (Feeling tense) coded with the symptoms or symptom cluster from the muscle/posture sub-category identified by the researcher, but it could also be associated with a feeling of mental tension. The Nijmegen Questionnaire is a suggested screening tool for hyperventilation syndrome, based on reported symptoms. However, these symptoms are not exclusive to individuals with hyperventilation syndrome. With the nature of psychiatric and/or psychological disorders in terms of diagnosis/recognition, it was unrealistic to exclude patients with these problems. The mental health background of patients from the study was unexplored and could have affected their symptom reporting.

Table 5.1

COSMIN checklist for Content Validity

Questions to determine if a study meets the standards for methodological quality	Excellent	Good	Fair	Poor
1 Was there an assessment of whether all items refer to the relevant aspects of the construct to be measured?	✓			
2 Was there an assessment of whether all items are relevant for the study population?	✓			
3 Was there an assessment of whether all items are relevant for the purpose of the measurement instrument?	✓			
4 Was there an assessment of whether all items together comprehensively reflect the construct to the measured?	✓			
5 Were there any important flaws in the design or methods of the study?	✓			

Note. The definition of *Excellent* for different questions are: 1 = Assessed if all items refer to relevant aspects of the construct to be measured. 2 = Assessed if all items are relevant for the study population in adequate sample size (≥ 10). 3 = Assessed if all items are relevant for the purpose of the application. 4 = Assessed if all items together comprehensively reflect the construct to be measured. 5 = No other important methodological flaws in the design or execution of the study.

It must be noted that while there is a level of transferability of the interview findings, purposeful sampling methods carried some other disadvantages. With the restraint proposed by the inclusion and exclusion criteria, only the relatively healthy individuals with hyperventilation syndrome were included in this study. This limits the inferences that can be made outside of an otherwise healthy population. Patients with organic disease diagnosis other than hyperventilation syndrome are also treated at the Auckland-based hyperventilation physiotherapy clinics and it is not known if the results are transferable to them. The majority of the participants were recruited from the hospital, where access to potential participants was more readily available compared to other approved research localities. This led to an overrepresentation of data from the publically funded clinic and the underrepresentation of the privately funded practices in the community. The demographics of participants could also be affected by this as people who access public health services could be of different socioeconomic status compared to those who access health services privately (Chaitow et al., 2002).

On the investigation of internal construct validity, this study was rated *Excellent* (Table 5.2) on the COSMIN checklist (Terwee et al., 2012). Using a modern test theory approach (i.e. Rasch analysis) is the most appropriate for the evaluation of questionnaire-based measurement scales. Thus, this study provides new information on the internal construct validity of the Nijmegen Questionnaire that was lacking in the literature. This quantitative aspect of the study is generalisable to the wider population of individuals with hyperventilation syndrome. A conversion table was produced for clinicians and researchers for converting the raw ordinal scores to interval scores after which traditional parametric statistical techniques can be applied. The final solution from the Rasch analysis allowed issues of misfit, local dependency, and item threshold disordering to be addressed through the process of constructing the 15-item version of the Nijmegen Questionnaire. This was also in line with the Qualitative Descriptive findings. The study included 239 questionnaires. This is sufficient for samples with poor targeting (Linacre, 1994). Given our targeting was good, the sample size for this study was more than sufficient.

Table 5.2

COSMIN checklist for Construct Validity

Questions to determine if a study meets the standards for methodological quality	Excellent	Good	Fair	Poor
1 Was the percentage of missing items given?	✓			
2 Was there a description of how missing items were handled?	✓			
3 Was the sample size* included in the analysis?	✓			
4 Were there any important flaws in the design or methods of the study?	✓			
5 For IRT: Were IRT test for determining the uni/dimensionality of the items performed?	✓			

Note. IRT = Item response theory.
*Sample size for this study was 239. Number of questionnaires per item was 14.
The definition of *Excellent* for different questions are: 1 = Percentage of missing items described. 2 = Described how missing items were handled. 3 = At least seven questionnaires per each item on the questionnaire AND ≥ 100 questionnaires. 4 = No other important methodological flaws in the design or execution of the study. 5 = IRT test for determining uni/dimensionality performed.

5.4 Implications for Practice

The study findings illuminated a number of implications for practice. This section discusses these implications which I believe require attention from clinicians and researchers who are involved in the management of patients with hyperventilation syndrome.

Interview findings revealed two existing items that appeared to be a poor match to the symptoms of hyperventilation syndrome. Additionally, a number of symptoms identified by participants are not captured by existing items of the Nijmegen Questionnaire. The reason for the mismatch between items and symptoms could be multifaceted. On the one hand, the interpretation and description of these symptoms varied between patients and health professionals. This could cause symptoms to be missed or misinterpreted by both parties in the clinical encounter. The Nijmegen Questionnaire does contain a majority of items that reflect symptoms of hyperventilation syndrome. Given that the questionnaire is designed for self-reported symptoms for screening purposes, clinicians or researchers need to be aware of keeping clarification and definition to a minimum in the context of questionnaire completion. The use of standardised instruction will also be helpful in keeping assessment condition similar each time. However, health professionals could be asked to clarify what the questionnaire items mean. This reflects the value of the subjective assessment alongside standardised outcome measures to capture a more in-depth perspective of patients' symptomatic experience.

While the questionnaire is internally valid for repeat assessment (as there was no bias over time points in this study), no validation process to date has proved the ability of this questionnaire in measuring change (e.g. treatment effectiveness on hyperventilation syndrome). It is important to be aware of this when interpreting results from more than one assessment for individual patients. The same caution needs to be applied when using the Nijmegen Questionnaire as an outcome measure in research.

On symptom reporting and capturing, the Nijmegen Questionnaire contains symptoms that are from at least three different categories: breathing symptoms, psychological symptoms, and physical symptoms. Health professionals need to be mindful of this when assessing individuals with suspected hyperventilation syndrome and when appropriate, refer individuals to specialty medical or surgical services for further assessments to rule out other underlying diagnosis. Moreover, it can be particularly challenging with symptoms associated with psychological wellbeing. Patients can present with hyperventilation syndrome with or without psychological disorders. Hence, a comprehensive assessment relative to the health

professionals managing these patients is essential in ensuring the overall management of the patients' health.

On balance, the Nijmegen Questionnaire is an outcome measure that is suitable for its purpose in screening for hyperventilation syndrome in clinical and research setting with standards for application in place. The utilisation of the conversion table is recommended for converting ordinal raw scores to interval data when using the Nijmegen Questionnaire especially when parametric testing is indicated. It should be used in conjunction with other subjective and objective measures when assessing for hyperventilation syndrome.

5.5 Suggestion for Future Research

To further explore the psychometric properties of the Nijmegen Questionnaire, it might be possible to interview more people of different gender, ethnicity, and socioeconomic backgrounds at various clinic localities. The cultural validation of the questionnaire for its use in New Zealand is yet to be carried out.

A different research method that could be applied in future investigations would be clinical observations with an appropriate research protocol to allow frequency count of symptoms being reported at different clinic settings. Future Rasch analysis of the questionnaire could include a larger number of questionnaires with the goal of a more even distribution of males and females. Other psychometric properties such as test-retest reliability of the questionnaire can also be explored.

This study did not set out to examine the validity of the recommended cut-point of 24/64 to make a positive screening of hyperventilation syndrome. If we had been able to retain the original scoring of the questionnaire we would suggest to further examine the sensitivity and specificity using this cut-point. However, we rescored 15 items and deleted one, thus we cannot simply test if 24/64 is the best cut-point. A prospective study is required, which compares the revised Nijmegen Questionnaire with a 'gold standard' assessment to determine the optimal cut-point for positive screening. Such a study will also allow further assessment of the Nijmegen Questionnaire item locations and determine if items capturing more severe symptoms are warranted to enhance the scale's ability to measure change.

5.6 Conclusion

This outcome measurement study adds to the limited body of literature on the psychometric properties of the Nijmegen Questionnaire. The content validation findings from this study highlight that the existing questionnaire capture symptoms that are representative of hyperventilation syndrome, as well as observable discrepancies in the perception and description of these symptoms. There are symptoms that are currently not captured by the Nijmegen Questionnaire that are attributed to hyperventilation syndrome by both patients and health professionals. Rasch analysis of the existing questionnaire found that it did not fit the strict criteria for internal construct validity.

This study has produced *the Revised 15-item Nijmegen Questionnaire for Hyperventilation Syndrome* which is internally valid and has satisfactory content validity. With the amended scoring and the use of conversion table provided, this revised questionnaire is recommended for use in clinical and research practice. Future areas for research include a more diverse range of population with different gender, ethnic, and socioeconomic background when exploring content validity. Research methods such as theory testing and frequency counting of symptoms observed in the clinical setting can be applied. Other psychometric properties of the Nijmegen Questionnaire such as test-retest reliability need to be explored with modern statistical techniques to provide further information for health professionals using this questionnaire. Cultural validation and the investigation of the appropriate cut-point for positive screening are recommended for future research.

References

Andrich, D., Sherridan, B., & Luo, G., (2009). RUMM (2030) [For analysis assessment and attitude questionnaire data]. Perth: RUMM Laboratory Pty Ltd. Retrieved from http://www.rummlab.com.au/

Beaton, D., Bombardier, C., Guillemin, F., & Ferraz, M. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine, 25*(24), 3186-3191. Retrieved from http://www.spinejournal.com

Boone, W., Staver, J., & Yale, M. (2014). *Rasch Analysis in the Human Science*. Dordrecht, Netherlands: Springer.

Bowling, A. (Ed.). (1997). *Measuring health: A review of quality of life measurement scales* (2nd ed.). Buckingham, United Kingdom: Open University Press.

Bowling, A. (Ed.). (2001). *Measuring disease* (2nd ed.). Buckingham, United Kingdom: Open University Press.

Bowling, A. (Ed.). (2009). *Research methods in health* (3rd ed.). Berkshire, United Kingdom: Open University Press.

Chaitow, L., Bradley, D., & Gilbert, C. (2002). *Multidisciplinary approaches to breathing pattern disorders*. Edinburgh, Scotland: Churchill Livingstone.

Courtney, R., & Greenwood, K. (2009). Preliminary investigation of a measure of dysfunctional breathing symptoms: The Self Evaluation of Breathing Questionnaire (SEBQ). *International Journal of Osteopathic Medicine, 12*, 121-127. doi: 10.1016/j.ijosm.2009.02.001

DeVellis, R. (2006). Classical Test Theory. *Medical Care, 44*(Suppl. 3), S50-S59. doi: 10.1097/01.mlr.0000245426.10853.30

Ek, S. (2013). Gender differences in health information behaviour: A Finnish population-based survey. *Health Promotion International, 30*(3), 736-745. Retrieved from https://academic.oup.com/heapro

Fisher, W., Jr. (1992). Reliability, separation, strata statistics. *Rasch Measurement Transactions, 6*(3), 238. Retrieved from https://www.rasch.org/rmt/

Garssen, B., Colla, P., van Dixhoorn P., Folgering, H., Stoop, a., & de Swart, J. (1984). Het herkennen van het hyperventilatiesyndroom [Recognising the hyperventilation syndrome]. *Medisch Contact, 35*, 1122-1124.

Grossman, P., & de Swart, J. (1984). Diagnosis of hyperventilation syndrome on the basis of reported complaints. *Journal of Psychosomatic Research, 28*(2), 97-104. Retrieved from http://www.journals.elsevier.com/journal-of-psychosomatic-research

Guba, E., & Lincoln, Y. (1982). Epistemological and methodological bases of naturalistic inquiry. *Educational Communication and Technology, 30*(4), 233-252. Retrieved from https://www.jstor.org/journal/educcommtech

Hambleton, R., & Jones, R. (1993). An NCME instructional module on: Comparison of classical test theory and item response theory and their applications to best development. *Educational Measurement: Issues and Practice, 12*, 38-47. Retrieved from http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1745-3992

Hobart, J., & Cano, S. (2009). Improving the evaluation of therapeutic interventions in multiple sclerosis: The role of new psychometric methods. *Health Technology Assessment, 13*(12). doi: 10.3310/hta13120

Høegh, M., & Høegh, S.-M., (2009). Trans-adapting outcome measures in rehabilitation: Cross-cultural issues. *Neuropsychological Rehabilitation, 19*(6), 955-970. Retrieved from http://www.tandfonline.com/toc/pnrh20/current

Hornsveld, H., Garssen, B., Dop, F., & van Spiegel, P. (1990). Symptom reporting during voluntary hyperventilation and mental load: Implications for diagnosing hyperventilation syndrome. *Journal of Psychosomatic Research, 34*(6), 687-697. Retrieved from http://www.journals.elsevier.com/journal-of-psychosomatic-research

Hornsveld, H., Garssen, B., Dop, F., van Spiegel, P., & de Haes, J. (1996). Double-blind placebo-controlled study of the hyperventilation provocation test and the validity of the hyperventilation syndrome. *Lancet, 348*, 154-158. Retrieved from http://www.thelancet.com/

Hsieh, H., & Shannon, S. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-1288. Retrieved from http://journals.sagepub.com/home/qhr

Hudson, M., Milne, M., Reynolds, P., Russel, K., & Smith, B. (2010). *Te Ara Tika Guidelines for Māori research ethics: A framework for researchers and ethics committee members*:

Health Research Council of New Zealand. Retrieved from http://www.hrc.govt.nz/sites/default/files/Te%20Ara%20Tika%20Guidelines%20for%20Maori%20Research%20Ethics.pdf

Juniper, E., Guyatt, G., Ferrie, P., & Grifffith, L. (1993). Measuring quality of life in asthma. *American Review of Respiratory Disease, 147* (4), 832-838. Retrieved from http://www.atsjournals.org/loi/arrd

Kersten, P., & Kayes, N. (2011). Outcome measurement and the use of Rasch analysis, a statistics-free introduction. *New Zealand Journal of Physiotherapy, 39*(2), 92-99. Retrieved from https://pnz.org.nz/journal

Kersten, P., White, P., & Tennant, A. (2012). The consultation and relational empathy measure: An investigation of its scaling structure. *Disability and Rehabilitation, 34*(6), 503-509. Retrieved from http://www.tandfonline.com/loi/idre20

Laver Fawcett, A. (2007). *Principles of assessment and outcome measurement for occupational therapists and physiotherapists: Theory, skills, and application.* West Sussex, England: John Wiley & Sons.

Li Ogilvie, V., & Kersten, P., (2015). A critical review of the psychometric properties of the Nijmegen Questionnaire for hyperventilation syndrome. *New Zealand Journal of Physiotherapy, 43*(1), 3-10.

Linacre, J. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions, 7*(4), 328. Retrieved from https://www.rasch.org/rmt/

Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.

Lundgren Nilsson, Å., & Tennant, A. (2011). Past and present issues in Rasch analysis: The functional independence measure (FIM™) revisited.

Lum, L. (1975). Hyperventilation: The tip and the iceberg. *Journal of Psychosomatic Research, 19*(5), 375-383. Retrieved from http://www.journals.elsevier.com/journal-of-psychosomatic-research

Lum, L. (1981). Hyperventilation and anxiety state. *Journal of the Royal Society of Medicine, 74*(1), 1-4. Retrieved from http://jrs.sagepub.com/

Marais, I., & Andrich, D. (2008). Effects of varying magnitude and patterns of response dependence in the unidimensional Rasch model. *Journal of Applied Measurement, 9*(2), 105-124. Retrieved from http://jampress.org/

McDowell, I. (Ed.). (2006). *Measuring health: A guide to rating scales and questionnaires* (3rd ed.). New York, NY: Oxford University Press.

Ministry of Health. (2015). *Te Whare Tapa Whā: Māori Health Model. Retrieved from http://www.health.govt.nz/system/files/documents/pages/maori_health_model_tewhare.pdf*

Mogilner, C., Rudnick, T., & Iyengar, S. (2008). The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of consumer researcher, 35* (2), 202-215. doi:10.1086/588698

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., de Vet, H. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research, 19*(4), 539-549. Retrieved from http://www.springer.com/medicine/journal/11136

Mooney, S., & Candy, S. (2008). The real cost of effective treatment. A single case study of a patient with hyperventilation syndrome. *New Zealand Journal of Physiotherapy, 36*(2), 88. Retrieved from http://pnz.org.nz/journal

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Sydney, Australia: Allen & Unwin.

Patton, M. (1990). *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage.

Physiotherapy Board of New Zealand (2016). *Vision statement*. Retrieved from http://www.physioboard.org.nz/

Physiotherapy Board of New Zealand (2016). Home page. Retrieved from https://www.physioboard.org.nz/

Polytomous Rasch Model. (n.d.). In *Wikipedia*. Retrieved March 8, 2017, from https://en.wikipedia.org/wiki/Polytomous_Rasch_model

Rapee, R., & Medoro, L. (1994). Fear of physical sensations and trait anxiety as mediators of the response to hyperventilation in nonclinical subjects. *Journal of Abnormal Psychology, 103*(4), 693-699. Retrieved from http://www.apa.org/pubs/journals/abn/index.aspx

Rowley, J. & Nicholls, D. (2006). Development of the RoBE self-efficacy scale for people with breathing pattern disorders. *New Zealand Journal of Physiotherapy, 34*(3), 131-141. Retrieved from http://pnz.org.nz/journal

Ruiter, C., Garssen, B., Rijken, H., & Kraaimaat, F. (1989). The hyperventilation syndrome in panic disorder, agoraphobia and generalized anxiety disorder. *Behaviour Research and Therapy, 27*(4), 447-452. Retrieved from https://www.journals.elsevier.com/behaviour-research-and-therapy/

Sabourin, B., Stewart, S., Watt, M., & MacDonald, A. (2013). Development and psychometric evaluation of a brief version of the Hyperventilation Questionnaire: the HVQ-B. *Cognitive Behaviour Therapy, 42*(3), 193-202. Retrieved from http://www.tandfonline.com/toc/sbeh20/current

Sandelowski, M. (1995). Focus on research methods: Sample size in qualitative research. *Research in Nursing & Health, 18*(2), 179-183. Retrieved from http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1098-240X

Sandelowski, M. (2000). Focus on research methods: Whatever happened to qualitative description? *Research in Nursing & Health, 23*(4), 334-340. Retrieved from http://onlinelibrary.wiley.com/journal/10.1002/(ISSN)1098-240X

Scientific Advisory Committee of the Medical Outcomes Trust. (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. *Quality of Life Research, 11*(3), 193-205. Retrieved from http://link.springer.com/journal/11136

Siegert, R., Tennant, A., & Turner-Stokes, L. (2010). Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disability and Rehabilitation, 32*(1), 8-17. Retrieved from http://www.tandfonline.com/loi/idre20

Streiner, D., & Norman, G. (Eds.). (2008). *Health measurement scales: A practical guide to their development and use* (4th ed.). Oxford, England: Oxford University Press.

Tabachnick, B., & Fidell L. (Eds.) (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Allyn & Bacon.

Terwee, C., Mokkink, L., Knol, D., Ostelo, R., Bouter, L., & de Vet, H. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research, 21*, 651-657. doi:10.1007/s11136-011-9960-1

Thomas, M., McKinley, R., Freeman, E., & Foy, C. (2001). Prevalence of dysfunctional breathing in patients treated for asthma in primary care: Cross sectional survey. *British Medical Journal, 322*(7294), 1098-1100. Retrieved from http://www.bmj.com/

Thomas, M., Mckinley, R., Freeman, E., Foy, C., & Price, D. (2005). The prevalence of dysfunctional breathing in adults in the community with and without asthma. *Primary Care Respiratory Journal, 14*, 78-82. doi:10.1016

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.

van Dixhoorn, J., & Duivenvoorden, H. (1985). Efficacy of Nijmegen Questionnaire in recognition of the hyperventilation syndrome. *Journal of Psychosomatic Research, 29*(2), 199-206. Retrieved from http://www.journals.elsevier.com/journal-of-psychosomatic-research

van Dixhoorn, J., & Folgering, H. (2015). The Nijmegen Questionnaire and dysfunctional breathing. *1*, 1-4. doi: 10.1183/23120541.00013-2015

van Doorn, P., & Colla, P. (1986). Association of hyperventilation complaints with ventilatory variables. *Biological Psychology, 22*, 180-181. doi:10.1016/0301-0511(86)90050-5

van Doorn, P., Colla, P., & Folgering, H. (1983). Een vragenlijst voor hyperventilatieklachten [A questionnaire for hyperventilation syndrome]. *De Psycholoog, 18*(10), 573-577.

van Doorn, P., Folgering, H., & Colla, P. (1982). Control of the end-tidal PCO_2 in the hyperventilation syndrome: Effects of biofeedback and breathing instructions compared. *Bulletin Europeen De Physiotherpathologie Respiratoire, 18*(6), 829-836.

Vansteenkiste, J., Rochette, F., & Demedts, M. (1991). Diagnostic tests of hyperventilation syndrome. *European Respiratory Journal, 4*(4), 393-399. Retrieved from http://erj.ersjournals.com/

Warburton, C., & Jack, S. (2006). Can you diagnose hyperventilation? *Chronic Respiratory Disease, 3*(3), 113-115. Retrieved from http://journals.sagepub.com/home/crd

Appendix A: The Nijmegen Questionnaire

THE NIJMEGEN QUESTIONNAIRE

	Never 0	Rare 1	Sometimes 2	Often 3	Very Often 4
Chest pain					
Feeling tense					
Blurred vision					
Dizzy spells					
Feeling confused					
Faster or deeper breathing					
Short of breath					
Tight feelings in chest					
Bloated feeling in stomach					
Tingling fingers					
Unable to breath deeply					
Stiff fingers or arms					
Tight feelings round mouth					
Cold hands or feet					
Palpitations					
Feeling of anxiety					

TOTAL: **/ 64**

Appendix B: Relevant Questions from the COSMIN checklist

Reliability

1. Was the percentage of missing items given?
2. Was there a description of how missing items were handled?
3. Was the sample size included in the analysis adequate?
4. Were at least two measurements available?
5. Were the administrations independent?
6. Was the time interval stated?
7. Were patients stable in the interim period on the construct to be measured?
8. Was the time interval appropriate?
9. Were the test conditions similar for both measurements?
10. Were there any important flaws in the design or methods of the study?
11. For continuous scores: Was an intraclass correlation coefficient (ICC) calculated?
12. For dichotomous/nominal/ordinal scores: Was kappa calculated?
13. For ordinal scores: Was a weighted kappa calculated?
14. For ordinal scores: Was the weighting scheme described? E.g. linear, quadratic

Content validity

1. Was there an assessment of whether all items refer to relevant aspects of the construct to be measured?
2. Was there an assessment of whether all items are relevant for the study population?
3. Was there an assessment of whether all items are relevant for the purpose of the measurement? (discriminative, evaluative, and/or predictive)
4. Was there an assessment of whether all items together comprehensively reflect the construct to be measured?
5. Were there any important flaws in the design or methods of the study?

Structural validity

1. Does the scale consist of effect indicators, i.e. is it based on a reflective model?
2. Was the percentage of missing items given?
3. Was there a description of how missing items were handled?
4. Was the sample size included in the analysis adequate?
5. Were there any important flaws in the design or methods of the study?
6. For CTT: Was exploratory or confirmatory factor analysis performed?
7. For IRT: Were IRT tests for determining the (uni-) dimensionality of the items performed?

Appendix B: Relevant Questions from the COSMIN checklist (Continued)

Hypotheses testing (an aspect of construct validity)

1. Was the percentage of missing items given?
2. Was there a description of how missing items were handled?
3. Was the sample size included in the analysis adequate?
4. Were hypotheses regarding correlations or mean differences formulated a priori (i.e. before data collection)?
5. Was the expected direction of correlations or mean differences included in the hypotheses?
6. Was the expected absolute or relative magnitude of correlations or mean differences included in the hypotheses?
7. For convergent validity: Was an adequate description provided for the comparator instrument(s)?
8. For convergent validity: Were the measurement properties of the comparator instrument(s) adequately described?
9. Were there any important flaws in the design or methods of the study?
10. Were design and statistical methods adequate for the hypotheses to be tested?

Criterion validity

1. Was the percentage of missing items given?
2. Was there a description of how missing items were handled?
3. Was the sample size included in the analysis adequate?
4. Can the criterion used or employed be considered as a reasonable 'gold standard'?
5. Were there any important flaws in the design or methods of the study?
6. For continuous scores: Were correlations, or the area under the receiver operating curve calculated?
7. For dichotomous scores: Were sensitivity and specificity determined?

Appendix C: Participant Information Sheet

Participant Information Sheet

Date Information Sheet Produced:
23 June 2015

Project Title
An investigation of the internal construct and content validity of the Nijmegen Questionnaire for hyperventilation syndrome in adults.

An Invitation
E nga mana, e nga reo, e nga waka, e nga karangatanga maha, tena koutou, tena koutou, tena koutou katoa.

Tihei mauri ora (The breath of life). Kia ora koutou, and greetings.

You are invited to **share your experience/knowledge on hyperventilation syndrome** as part of a research based at the Hyperventilation Physiotherapy Clinic at Middlemore Hospital.

This research will contribute to my studies in the Master in Health Science at the AUT University. My name is Vickie Li Ogilvie and I am the researcher of this project. Your participation in this research is voluntary. Whether you choose to take part or not/withdraw from the research at any time, it will neither advantage nor disadvantage you.

What is the purpose of this research?
Research review has raised questions in regards to how accurate/valid the Nijmegen Questionnaire is in identifying people with hyperventilation syndrome. This research aims to gather further information about this questionnaire. In addition to statistically evaluate existing data from completed questionnaires, I hope to find out what constitute to hyperventilation syndrome from the point of view of affected individuals and health workers who work with them.

The research outcomes (will be presented in the format of: a Master thesis, conference papers, academic publications/presentations) will contribute to better understanding and use of the questionnaire by health workers which will benefit people with hyperventilation syndrome.

How was I identified and why am I being invited to participate in this research?
My aim is to gain perspectives on hyperventilation syndrome from two distinct groups of individuals: (1) People who experienced hyperventilation syndrome; (2) Health professionals who assessed individuals with hyperventilation syndrome. As such, I wish to speak with people from different backgrounds, who have been recently assessed/treated at a hyperventilation physiotherapy clinic and health professionals who have had experience working with people with hyperventilation syndrome.

You have received this invitation because you have attended the Hyperventilation Physiotherapy Clinic or Breathing Works **OR** you have referred someone to the aforementioned physiotherapy clinics.

You are able to take part in this research if you:

Appendix C: Participant Information Sheet (Continued)

- are 18 years old or above
- are able to communicate with the researcher without an interpreter present
- live in Auckland
- **have been assessed/treated at a hyperventilation physiotherapy clinic** (people with organic cardiac, neurological, and/or respiratory disease will be excluded) **OR**
- **have had experience working with people with hyperventilation syndrome**

What will happen in this research?

In this research, you will be invited to take part in a face to face interview with me, the researcher. The interview will last about one hour. I will arrange a time and place that suits you to conduct the interview. The location can be at your home or at another desired location that you prefer. A support person (e.g. family/whānau member or friend) is welcome to be present at the interview.

At the interview, the questions will **focus on what you believe to be the signs and symptoms of hyperventilation syndrome**. I will seek permission from you to audiotape the interview.

What are the discomforts and risks?

You will be required to give about one hour of your time to the project. Please allow additional time for scheduling the interview and setting up for the interview on the day. There is no known risk for taking part in this project and it is highly unlikely that you will experience any discomfort at the interview.

However, I am aware that some people might feel vulnerable when disclosing personal information at the interview. If you experience any physical or psychological discomfort, you will be encouraged to seek advice from your general practitioner.

What are the benefits?

This research will contribute to my studies in the Master in Health Science at the AUT University.

There will be no direct benefits to you, although some people may find reflecting on their experiences with hyperventilation syndrome helpful.

For the wider community, the additional knowledge that results from this research on the Nijmegen Questionnaire will better equip health professionals in recognising people with hyperventilation syndrome. This can lead to timely delivery of tailored management for individuals with the condition or the referral to other health services as needed without delays.

How will my privacy be protected?

The information you provided, including all interview recording will remain strictly confidential and will be stored in a secure location through the course of the project. Data from your interview will be transcribed. All personally identifiable information will be removed from transcripts and you will be allocated with a study ID. No one other than myself (including your treating clinician if you are a patient OR your patients if you are a treating clinician) will be aware that you are participating in the study or be privy to anything you say in your interview.

No material that could personally identify you will be included in any interview excerpts, reporting or dissemination activities associated with the study.

Appendix C: Participant Information Sheet (Continued)

At the completion of the research, all your information will be stored at the locked offices at AUT University and on password protected computer for ten years. Hard copies of research data will be disposed of by confidential documentation destruction service and electronic data will be deleted from password protected devices.

What are the costs of participating in this research?

There is a non-monetary cost of your time – about one hour. If it is necessary for you to travel to an interview location, a $20 petrol voucher is provided to compensate your travel costs.

What opportunity do I have to consider this invitation?

You can consider this invitation over the next one to two week(s). If you are interested in discussing this information sheet further in person, please contact me to arrange a meeting as soon as possible. If you are interested in taking part, please contact me via contact details listed below OR complete and return the reply slip using the prepaid self-addressed envelope.

How do I agree to participate in this research?

You will be asked to complete a consent form to indicate your wish to take part in this research before we begin the interview process.

Will I receive feedback on the results of this research?

A brief summary of the research will be provided for you via email/post after the completion of the project.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Paula Kersten, *pkersten@aut.ac.nz*, 9219999 ext 9180.

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, Kate O'Connor, *ethics@aut.ac.nz* , 921 9999 ext 6038.

Whom do I contact for further information about this research?

Researcher Contact Details:

Vickie Li Ogilvie, Physiotherapist, *vickie.liogilvie@middlemore.co.nz*, 09 2708896.

Project Supervisor Contact Details:

Paula Kersten, Primary Project Supervisor, *pkersten@aut.ac.nz*, 9219999 ext 9180.

Nicola Kayes, Secondary Project Supervisor, *nkayes@aut.ac.nz*, 9219999 ext 7309.

Approved by the Auckland University of Technology Ethics Committee on 24 June 2015 AUTEC Reference number *15/197*.

Appendix D: Consent Form

Consent Form	

Project Title: ***An investigation of the internal construct and content validity of the Nijmegen Questionnaire for hyperventilation syndrome in adults***

Project Supervisors: ***Paula Kersten and Nicola Kayes***

Researcher: ***Vickie Li Ogilvie***

- ○ I have read and understood the information provided about this research project in the Information Sheet dated 23 June 2015.
- ○ I have had an opportunity to ask questions and to have them answered.
- ○ I understand that notes will be taken during the interviews and that they will also be audio-taped and transcribed.
- ○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.
- ○ I am over 18.
- ○ I am not suffering from any organic cardiac/neurological/respiratory disease.
- ○ I agree to take part in this research.
- ○ I wish to receive a copy of the report from the research (please tick one): Yes○ No○

Participant's Signature: .....

Participant's Name: .....

Participant's Contact Details:

……..

……..

……..

……..

Date:

Approved by the Auckland University of Technology Ethics Committee on 24 June 2015 AUTEC Reference number 15/197

Note: The Participant should retain a copy of this form.

Appendix E: Interview Guide

Introduction

- Thank you for participating
- Remind the participant about confidentiality
- Remind the participant about study aim

Starting Questions

For patients

How would you describe what it feels like to have hyperventilation syndrome?

Can you tell me about the symptoms that you associate with this condition?

How would someone know that you were experiencing hyperventilation syndrome if they were watching you?

What would they miss?

Could you think of a specific incident where you were experiencing hyperventilation syndrome and tell me about those symptoms?

For health professionals

How would you describe the signs and symptoms of hyperventilation syndrome?

How do you determine if someone is suffering from hyperventilation syndrome?

What other symptoms would a family member/friend/support person identify from an individual with hyperventilation syndrome?

Any cases that stood out to you that are different from what you told me already?

Prompting Questions

Can you tell me more about/expand on that?

When you said X, what did you mean?

You talked about X, Y, and Z, is there anything else you would like to add?

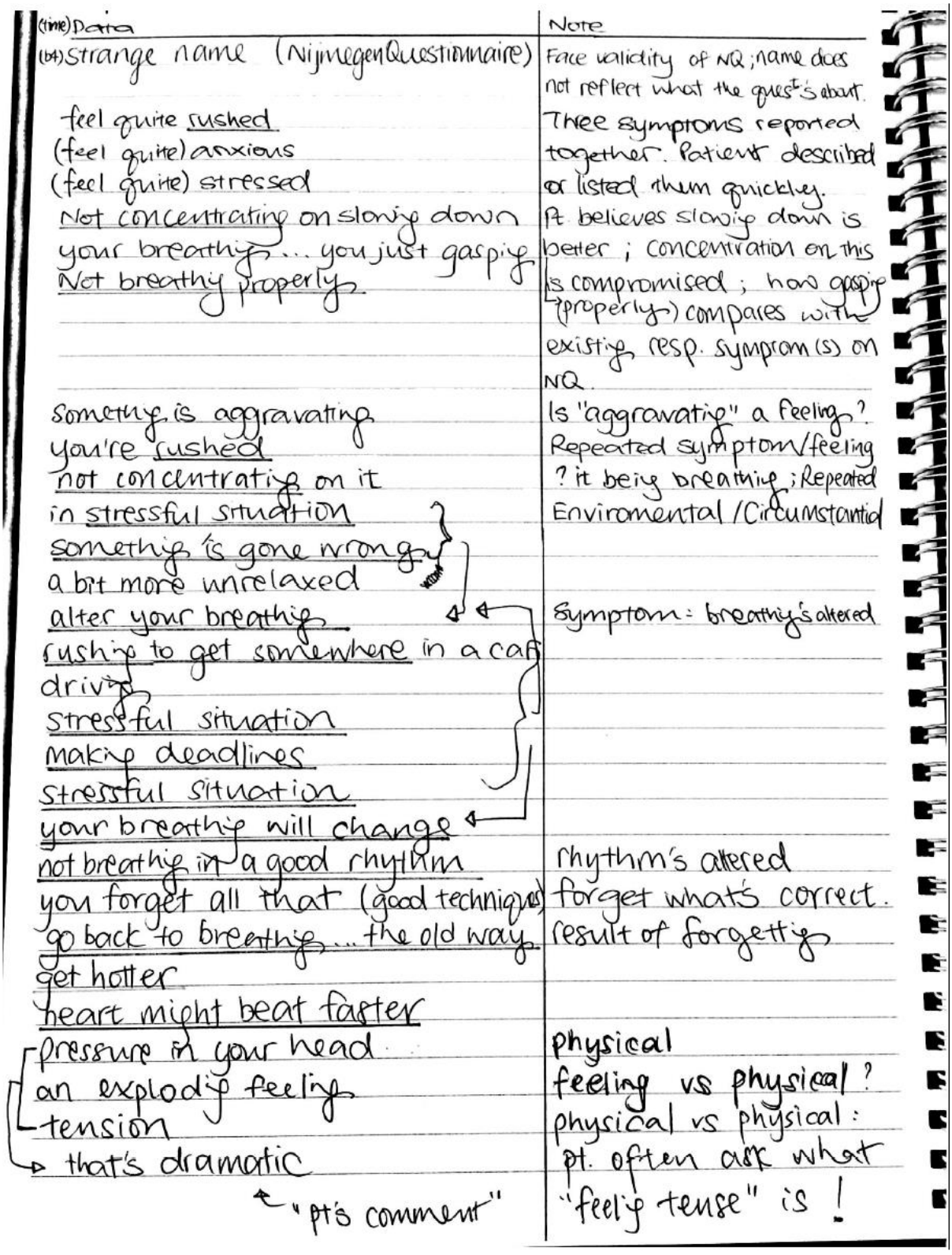
Nijmegen Questionnaire Questions

From your perspective, what are your views on the appropriateness of the questionnaire?

- Appropriateness of individual complaints
- Appropriateness of the response options
- Appropriateness of the language used
- Any important areas that are not currently included
- If you were to use this questionnaire, do you think it would give an accurate account of the symptoms associated with hyperventilation syndrome? Why?

Appendix F: Example of Interview Data

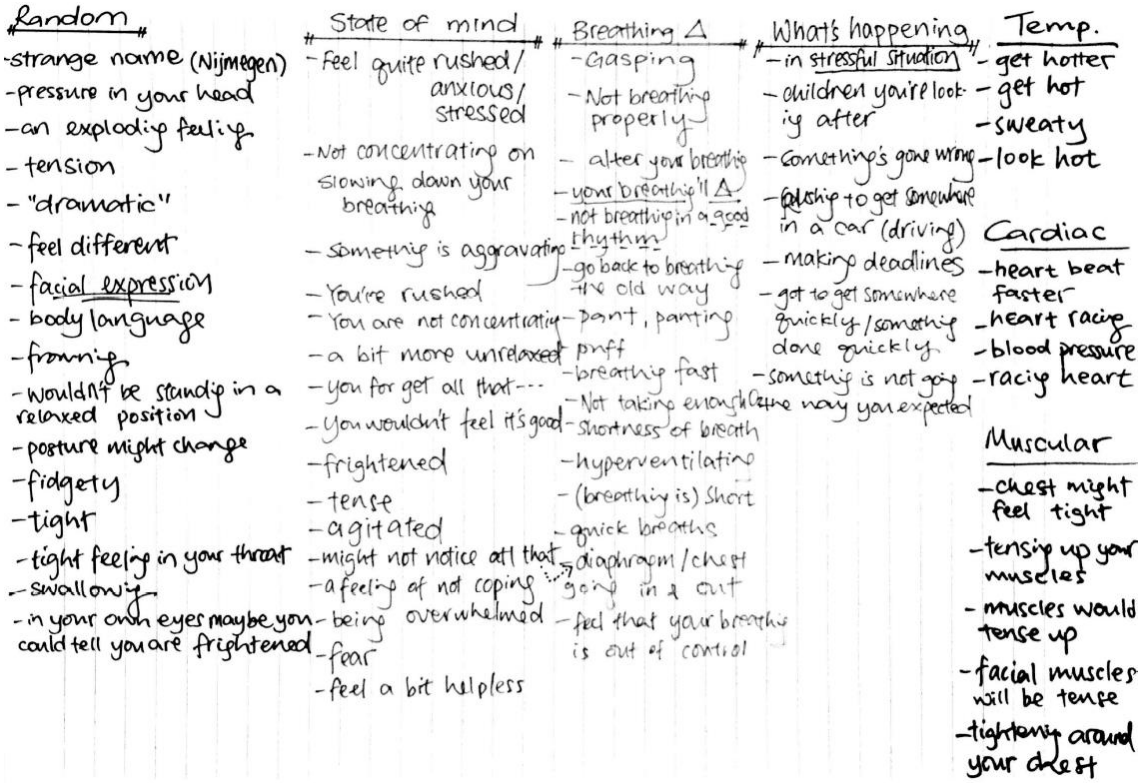
(List of Symptoms and Researcher's Notes)



(time) Data	Note
(04) Strange name (NijmegenQuestionnaire)	Face validity of NQ; name does not reflect what the ques^{t}'s about.
feel quite rushed	Three symptoms reported together. Patient described or listed them quickly.
(feel quite) anxious	
(feel quite) stressed	
Not concentrating on slowing down your breathing... you just gasping Not breathing properly.	Pt believes slowing down is better; concentration on this is compromised; how gasping (properly) compares with existing resp. symptom(s) on NQ.
something is aggravating	Is "aggravating" a feeling?
you're rushed	Repeated symptom/feeling
not concentrating on it	? it being breathing; Repeated
in stressful situation	Enviromental / Circumstantial
something is gone wrong	
a bit more unrelaxed	
alter your breathing	Symptom: breathing's altered
rushing to get somewhere in a car driving	
stressful situation	
making deadlines	
stressful situation	
your breathing will change	
not breathing in a good rhythm	rhythm's altered
you forget all that (good technique)	forget what's correct.
go back to breathing... the old way.	result of forgetting
get hotter	
heart might beat faster	
pressure in your head	physical
an exploding feeling	feeling vs physical?
tension	physical vs physical:
that's dramatic	pt. often ask what
← "pt's comment"	"feeling tense" is!

Appendix G: Example of Interview Data

(Symptoms and Preliminary Symptom Categories)



Random
- strange name (Nijmegen)
- pressure in your head
- an exploding feeling
- tension
- "dramatic"
- feel different
- facial expression
- body language
- frowning
- wouldn't be standing in a relaxed position
- posture might change
- fidgety
- tight
- tight feeling in your throat
- swallowing
- in your own eyes maybe you could tell you are frightened

State of mind
- feel quite rushed/ anxious/ stressed
- Not concentrating on slowing down your breathing
- something is aggravating
- You're rushed
- You are not concentrating
- a bit more unrelaxed
- you for get all that---
- You wouldn't feel it's good
- frightened
- tense
- agitated
- might not notice all that
- a feeling of not coping
- being overwhelmed
- fear
- feel a bit helpless

Breathing Δ
- Gasping
- Not breathing properly
- alter your breathing
- your breathing'll Δ
- not breathing in a good rhythm
- go back to breathing the old way
- pant, panting
- pnff
- breathing fast
- Not taking enough
- shortness of breath
- hyperventilating
- (breathing is) short
- quick breaths
- diaphragm/chest going in & out
- feel that your breathing is out of control

What's happening
- in stressful situation
- children you're looking after
- something's gone wrong
- rushing to get somewhere in a car (driving)
- making deadlines
- get to get somewhere quickly/something done quickly.
- something is not going the way you expected

Temp.
- get hotter
- get hot
- sweaty
- look hot

Cardiac
- heart beat faster
- heart racing
- blood pressure
- racing heart

Muscular
- chest might feel tight
- tensing up your muscles
- muscles would tense up
- facial muscles will be tense
- tightening around your chest

Appendix H: Auckland University of Technology Ethics Committee Approval

AUTEC
SECRETARIAT

24 June 2015

Paula Kersten
Faculty of Health and Environmental Sciences

Dear Paula

Re Ethics Application: **15/197 An investigation of the internal construct and content validity of Nijmegen Questionnaire for hyperventilation syndrome in adults.**

Thank you for providing evidence as requested, which satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC).

Your ethics application has been approved for three years until 24 June 2018.

As part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through http://www.aut.ac.nz/researchethics. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 24 June 2018;
- A brief report on the status of the project using form EA3, which is available online through http://www.aut.ac.nz/researchethics. This report is to be submitted either when the approval expires on 24 June 2018 or on completion of the project.

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to obtain this.

To enable us to provide you with efficient service, please use the application number and study title in all correspondence with us. If you have any enquiries about this application, or anything else, please do contact us at ethics@aut.ac.nz.

All the very best with your research,

Kate O'Connor
Executive Secretary
Auckland University of Technology Ethics Committee

Cc: Vickie Li Ogilvie vickie.liogilvie@gmail.com

Appendix I: Research Office Approval

25 June 2015

Dear Vickie

Thank you for the information you supplied to the Research Office regarding your research proposal:

Research Registration Number: **2043**

Ethics Reference Number: **15/197**

Research Project Title: **An investigation of the internal construct and content validity of Nijmegen Questionnaire for hyperventilation syndrome in adults**

I am pleased to inform you that the Research Committee and Director of Hospital Services have approved this research with you as the Co-ordinating Investigator.

Your study is approved until **24 June 2018** as specified on your HDEC ethics application.

Amendments:

- All amendments to your study must be submitted to the Research Office for review.
- Any substantial amendment (as defined in the *Standard Operating Procedures for HDECs*, May 2012) must also be submitted to the Ethics Committee for approval.

All external reporting requirements must be adhered to.

Please note that failure to submit amendments and external reports may result in the withdrawal of Ethical and Organisational approval.

We wish you well in your project. Please inform the Research Office when you have completed your study (including when a study is terminated early) and provide us with a brief final report (12 pages) which we will disseminate locally.

Yours sincerely

Research Advisor
Under delegated authority from the Research Committee and Director of Hospital Services

Appendix J: Verification of Māori Consultation Process

School of Rehabilitation and Occupation Studies

Verification of Māori Consultation Processes

This document provides verification that the research project named below was discussed with the School of Rehabilitation and Occupation Studies Mātauranga Māori Committee, AUT University. Specific comments and recommendations are indicated below.

Research Title: An investigation of the internal construct and content validity of the Nijmegen Questionnaire for hyperventilation syndrome in adults		
Researcher(s): Vickie Li Ogilvie		**Date: 03/06/2015**
Discussion Areas	**Addressed**	**Comments/ Recommendations**
Whakapapa: Relationships		
Researcher experience in field		C1
Consultation with local stakeholders		C2, R1
Consenting process	X	
Clarity of data usage	X	
Dissemination of findings		C3
Benefits to participants		C3
Tika: Validity of the research		
Clear purpose of project	X	
Relevance to Māori		C1
Likely outcome for participants, communities, other stakeholders		C3
Participant recruitment methods	X	
Māori involvement in project (participants, researchers, etc)	X	
Manaakitanga: Responsibility and respect		
Participants' access to appropriate advice	X	
Participants treated with dignity and respect	X	
Privacy and confidentiality	X	
Whānau support	X	
Transparency of research process		R2, R3
Mana tangata: Power & Authority		
Reciprocity (acknowledgements, compensation, gifts)	X	
Risks of participation identified	X	
Ownership of outcomes	X	
Informed consent process	X	

Comments

1. New emerging researcher with experience in this field, aware that >20% of current clients identify as Maori
2. Has had approval from Ko Awatea and is aware of literacy differences in terms of description of pain, anxiety, breathing etc and requires further exploration of the layman's terms to describe these characteristics for development of the open-ended interviews.
3. The information gathered from this study will inform clinical practice, reduce waiting times, enable appropriate management and improve knowledge regarding physical and psychological domains of this condition

Recommendations

1. Consult further with Ko Awatea regarding appropriate phrases, greetings and possibility of inclusion of pepeha and mihi
2. The Nijemegen Questionnaire may not be appropriate to form the basis of the open-ended interviews, seek alternatives which are available in the ICF browser for current information
3. Use Te Whare Tapa Wha as a guideline to structure the interviews

Appendix J: Verification of Māori Consultation Process (Continued)

General recommendations for recruitment of Māori in this and future studies:

Feedback on these comments and recommendations is to be provided by:

Signature: Date: 03 June 2015

Representative, Mātauranga Māori Committee