AJELS can be found on EBSCOhost databases
https://doi.org/10.59960/8.a8

"What's in my medicine?": Evaluating the readability and comprehensibility of patient information leaflets of selected Philippine nonprescription drugs

Anthony Lemuell S. Cabañas & Marilu Rañosa-Madrunio

Published: December 2020

Corresponding author
Anthony Lemuell S. Cabañas
University of Santo Tomas Graduate School, Manila, the Philippines
anthonylemuellc@gmail.com

Marilu Rañosa-Madrunio
University of Santo Tomas, Manila, the Philippines

Additional information is available at the end of the article

Abstract

In the absence of a medical professional, patient information leaflets (PILs) serve as consumers' source of legitimate information about what a medical drug contains, what risks are involved in taking it, and how it should be safely administered and handled. More than just being a reference material, PILs guide consumers to make informed healthcare decisions for themselves. Hence, PILs become problematic when they defeat their purpose of delivering comprehensive information and become a source of confusion to consumers. Guided by the user-based tripartite model of communicative effectiveness (Garner et al., 2012), this paper evaluated the readability and comprehensibility of existing PILs culled from three common nonprescription medications in Philippine households – fever reducers, cough and cold remedy, and anti-inflammatory drugs. Forty-two PILs were collected and subjected to a readability test using the measures of text readability and easability of Coh-Metrix. Of the 42 leaflets, three were randomly selected – one for each medication category – to undergo user-testing via cloze procedure taken by 41 adult participants. Results revealed that based on the readability test, the existing PILs were found to be more academic-oriented rather than spoken-like, thus requiring high-skilled reading. Moreover, a user-testing of the leaflets revealed that only less than half of the participants demonstrated an independent level of understanding of the materials. This was further supported by an analysis of the difficulty index of the leaflets' information content, which found that most of the existing PILs contain texts with high difficulty, raising concerns over the materials' suitability for nonspecialist readers.

Keywords: Health discourse, health literacy, medical linguistics, patient information leaflets, user-testing

1. Introduction

Patient information leaflets (PILs) are essential communication materials used by clinical practitioners to convey medical information to healthcare consumers. They normally come in the form of pocket-size documents placed inside the packaging of medications that are made available for public consumption. They generally serve two functions in relation to healthcare consumers: first, to provide information about the characteristics, content, benefits, and adverse effects of the medical product; and second, to give instructions about the safety procedures in handling or taking the product (Fage-Butler, 2011; Reguero, 2017). Moreover, outside a clinician's office, PILs serve as healthcare consumers' accessible source of medical information. Harwood and Harrison (2004) posit that PILs are practically beneficial to healthcare consumers because (a) they are readily available with the medical product, (b) they spare people the cost of frequently going to the clinic/hospital and having to pay for a doctor's consultation, and (c) they can be used by consumers as reference just in case there are questions they feel ashamed to ask a doctor, or they want to be away from the stressful environment of a consultation room. In addition, several studies found that a patient's retention of medical information is better when it is written rather than when it is spoken (Harwood & Harrison, 2004; Kessels, 2003; Pongpunna et al., 2018). This is mainly due to the fact that during consultation, a number of patients encounter difficulty following their doctor's talk, especially in cases where their attending doctor mentions unclarified medical jargons. When medical information comes in the form of a written document that they can read such as, in this case, PILs, patients have more time to encode, process, and store information, and hence, can make more informed healthcare decisions for themselves.

PILs are a mandatory requirement in the healthcare system of a number of countries across the globe, and they come in different terms. For instance, in a comparative study of information leaflets for medical risk or safety communication, Yamamoto (2014) found that different names are used for PILs in Japan, the European Union, and the US. Japan's healthcare agency refers to PILs as a drug guide for patients. In the US, PILs are sometimes called medication guides or patient package insert. For countries of the European Union, PILs are also known as package leaflets. While it is a general regulation internationally to require pharmaceutical industries to provide a pocket information leaflet for medical products, not all medical pocket documents may be considered a PIL. Technically, the type of medical document to be written and produced depends on the type of drug where the document is going to be used. Information leaflets differ for prescription and nonprescription drugs. Nonprescription drugs are packaged with PILs, which are medical documents that are drafted in plain language and are directed to nonspecialist readers. On the other hand, prescription drugs come with a more technical document called Package Inserts (PI) containing medical information mostly meant for a clinical professional's reference. The textual content for both is regulated by legal healthcare mandates, i.e., policies and guidelines are imposed to ensure that sufficient safety information about the medical product is provided to avert preventable medical injuries or accidents (Maglie, 2015).

As a highly regulated document, PILs are a standardized text because their structure is determined by a set of institutionalized conventions informed and established by members of the medical industry. While regulations on the organization of information in PILs differ from one country to another, it has been found that there are identifiable and obligatory segments, which seem to be generic in the structure of most PILs. Frade (2015) observed that the information content in most PILs may be generally divided into four main sections. These include (a) **identification of medicine**, which includes the name and description of the product, the formulation, and strength; (b) **patient information**, which includes information about therapeutic indications, safety instructions on the administration of the medicine, description of overdosage and undesirable effects, storage conditions, and expiration date; (c) **healthcare professional information**, which covers information about the pharmaceutical form and efficacy of the medical product, details about contraindications, i.e., medical factors or circumstances that may render the product harmful to use for a patient, medicine interaction, precautions and warnings, and a detailed description of the drug's adverse effects; and finally, (d) **legal basis**, which specifies a medical product's registration description, approval stamp from the healthcare agency, and the name and address of the pharmaceutical company or manufacturer.

Moreover, healthcare policies demanding a more consumer-centered approach to drafting PILs appear to be a recent legal trend observed in a number of countries (mostly western states); more consumer-based regulations were legislated in light of the plain language movement, and the growing literature in medical communication studies and applied linguistics highlighting the number of reported cases of healthcare consumers who found most of the content of PILs to be incomprehensible and who claim to be excluded from the drafting and evaluating process of PILs (Dickenson et al., 2001). In the Philippines, the regulatory body that filters and approves medical products being distributed in the market is the Food and Drug Administration (FDA), which operates under the auspices of the Department of Health (DOH). Among the regulatory directives for marketing and selling medical drugs in Philippine stores is for pharmaceutical companies to provide healthcare consumers accurate product information based on FDA's prescribed product labelling requirements. This is specified under the DOH Administrative Order (AO) 2016-0008 or the Revised Rules and Regulations Governing the Generic Labelling Requirements of Drug Products for Human Use, which sets the guidelines on how medical products should be properly identified and described. As per the aforementioned AO, nonprescription drugs are to be supplied with a PIL "prepared and written in layman's terms" upon its distribution in the market. The AO also indicates that information which appear in a medical drug's label or leaflet have to be written in English and/or Filipino. It is worth noting that despite the inclusion of the Filipino language in the labelling mandate, most PILs of medical drugs sold in Philippine stores are drafted in English. Hence, in this study, the PILs that the researchers managed to collect and examine are written in English as they are the most widely available.

Despite the legislation of healthcare regulatory policies standardizing medical product labels and ensuring accessible medical information, PILs continue to be a source of confusion to healthcare consumers. This could be attributed to two reasons. First,

manufacturers have been solely relying on readability estimates to evaluate PILs. Clerehan et al. (2005) posit that this reliance on readability estimates disregards the fact that readability tools can only measure the explicit features of a text, i.e., lexical or syntactic complexity; they could not accurately account for the implicit elements such as actual context or reader-reception, which is used to gauge the text's comprehensibility (or the degree in which the text can be understood by a reader). Second, the drafting and evaluating process of PILs has normally been participated by experts only, excluding healthcare consumers whose feedback is equally crucial and valid as, after all, they are the end users of the material (Dickenson, 2001; Fage-Butler, 2015).

Accordingly, the type of medical information that is made available for public consumption can have consequences as it can put healthcare consumers' safety at stake. Hence, any text that concerns the well-being of healthcare consumers has to be put under critical lens because as Hughman (2009) posits, communication of healthcare texts can significantly affect an individual's health welfare and outcomes. The inadequate and incomprehensible content of most available healthcare materials has recently gained scholarly interest among applied linguists and text analysts examining better ways of evaluating and drafting consumer-centered healthcare information materials (Asekhave & Zethsen, 2003; Clerehan et al., 2005; Fage-Butler, 2011). Respectively, the existing linguistic studies on PILs have emphasized the need to further examine the readability and comprehensibility of these healthcare materials to get a better picture as to how exactly healthcare consumers receive and respond to the medical information made available for them to read.

Among the first few recent studies on PILs was that of Azekehave and Zethsen's (2003) investigation of the patient package inserts in Denmark. They initially conducted a questionnaire study where they measured participant's degree of satisfaction of Danish PILs. The participants who joined were found to be generally better educated than the Danish national average, and while there were village participants, a majority of the participants were young professionals living in urban areas. It was revealed that surprisingly, only a few respondents expressed dissatisfaction with PILs. However, the researchers contend that this may have something to do with the fact that most participants were educated, and they felt like they already know how to read and understand the content of the PILs. It is worth noting that in the study, when the participants were asked if PILs need to be improved for other people, a majority of them expressed agreement, and a great percentage of those who agreed were professionals coming from the health sector. The researchers observed that the PILs were deemed generally reader-friendly to a participant who is well-educated, but to participants coming from the weakest third of the Danish population, the reception shows stark difference. A further examination on the information content of the leaflet found that in terms of language features, Danish PILs contain texts that show asymmetrical sender-receiver relationship where most of the content were expressed from the point of view of the expert as evidenced by the persistent use of medical jargons, nominalizations, depersonalized/passive clauses, inconsistent terminologies, and information-packed sentences making it more difficult for patients to understand and relate to medical information. The study also scrutinized the sequencing of information in the existing document structure that the Danish

government prescribes for PILs; the analysis showed that the prescribed structure does not seem to place the most relevant and important information first, resulting in patients experiencing inconvenience and difficulty in navigating medical details.

In another similar study, Gustafsson et al. (2005) sought to examine patients' comprehension of information about interactions and contraindications contained in the PILs of prescription drugs in Sweden. Thirty leaflets were randomly selected by experts for patient-based evaluation. Through a user-based comprehensibility questionnaire, the PILs were assessed by patients who were dispensed of the drug where the leaflets come from to get their feedback as actual users of the material. Findings revealed that while generally most parts of the information content of the PILs appear sound and comprehensible to the participants, the information items on 'risks of interaction' and 'contraindications' received mostly low scores among the participants indicating low comprehensibility. The researchers also found that the leaflets with the most low-scoring information are the ones which have texts of technical orientation or complex messages laced with medical jargons. The study raises concerns over the risk involved in relaying incomprehensible safety information to patients. Clinical contents, which are difficult to process and understand for patients, can make them more vulnerable to medical accidents or injuries.

Meanwhile, in the Philippines, one notable local linguistic research, which perhaps serves as one of the first few to look into the linguistic intricacies of Philippine safety texts, is the study of Dacumos (2016) on Philippine product warnings. The study was an intensive assessment of the legal adequacy, linguistic features, readability, and comprehensibility of product warnings indicated in the labeling content of products Filipino consumers would usually buy in the market. An interesting finding in the study is that of the three product categories – household chemicals, cosmetics, and medicine, that were subjected to a careful linguistic analysis, readability test, and comprehensibility assessment, medical warnings from the labels of over-the-counter (OTC) medications contained the most technical content and consequently, were the most difficult to understand for the participants. Feedback from a majority of the participants points to the frequent use of technical terms as source of the medical warnings' incomprehensibility.

A more recent study in Thailand by Pongpunna et al. (2018) explored patient's use of PILs, and the factors that affect their needs and expectations in using the healthcare leaflet. The study involved a survey of outpatients in two university hospitals in Thailand. Using self-administered questionnaires distributed over a three-month period, the researchers sought to examine patient's awareness and knowledge of PILs and their perception of PILs' importance and usefulness. The findings revealed that a majority of those who had not previously known PILs see such as important tools for disseminating essential medical information, and additionally, they had higher expectations of knowledge gains from using the leaflet. On the other hand, those with previous experience with PILs reported increased patient need scores, and even improved knowledge and level of awareness of medical safety information. Another major finding of the study points to the fact that while PILs are meant to relay essential safety information to patients in order to prevent medical accidents or injuries, a number of Thai patients claim that they only read PILs when they already start to experience

the side effects of the medical product they are taking. In addition, a majority of the patients do not even read the entire leaflet because of the anxiety and confusion triggered by the use of technical language.

In sum, despite the fact that in several countries, PILs have long been a required document for medical products, it is only in recent decades that linguistic or textual studies on PILs have proliferated (Asekhave & Zethsen, 2003). In the Philippines, the most notable local scholarly work that has so far made a comprehensive linguistic investigation of safety texts is the research of Dacumos (2016) on consumer product warnings. While it was an initial attempt by local language researchers to use linguistic theories and principles to address consumer safety issues in the country, local studies on the use of language in safety communication, particularly in Philippine healthcare texts, however, remain scarce. Research on PILs plays a significant role in improving practices in the healthcare industry as it can enlighten clinicians and legislators alike about how to better respond to healthcare consumers' needs and promote their general welfare.

1.1 Research Questions

The type of information healthcare consumers receive affects how they decide for their health or well-being. The language used in delivering healthcare information affects public health behavior and plays a crucial role in healthcare affairs; hence, addressing such an issue warrants an investigation on the readability and comprehensibility of healthcare texts. Currently, there exists a paucity of studies on PILs where the object of investigation is text readability and comprehensibility (Clerehan et al., 2005), especially in the Philippines. This paper sought to address the scant local research on PILs, and aimed to shed light on the importance of language in healthcare communication by probing into the following queries:

1. How readable are the existing PILs of Philippine nonprescription drugs based on the text readability and easability measures of Coh-Metrix?
2. How comprehensible are these PILs using consumer-testing?
3. What are the indices of difficulty found in the texts of these PILs?

1.1 Theoretical Framework

This study is informed by the tripartite model of communicative effectiveness of Garner et al. (2012), which is a user-based framework that evaluates the textual quality of written healthcare information based on three focal elements: *focus on the text*, *focus on the text and reader*, and *focus on the reader*. The proponents of the framework observe that most of the evaluation practices for written healthcare materials usually rely on readability tests and expert-based assessment, sidelining feedback from actual users. They argue that these may not be enough to measure the textual quality of healthcare texts as readability tests only measure the surface level of texts, and expert-based assessment does not translate to actual

assessment by users themselves. They emphasized the importance of integrating users (or healthcare consumers) in the process of drafting and evaluating written healthcare texts. In doing so, the proponents suggest that an effective evaluation of healthcare materials must follow a triangulation method where the elements validate the findings of one another.

In their framework, Garner et al. (2012) identify three key elements that must be focused on when conducting a proper evaluation of healthcare texts. First is *focus on the text* (readability), which concerns analysis of healthcare information at the textual/linguistic level. Bailin and Graftstein (2016) state that at its core, readability is a textual property that tells how easy or difficult it is to understand the information being conveyed by a text. In this paper, the researchers used the notion of reading by Garner et al. (2012) to understand the activity of reading as the process of proceeding through the text and making attempts at assigning meanings to words and phrases, and constructing coherent meanings from the text in general.

Second is *focus on the text and reader* (comprehensibility), which involves testing users' understanding of healthcare texts, i.e., how they react to or interact with the material. To understand how texts are comprehended, this paper subscribed to the notion of comprehension by Giliam et al. (2009, as cited in Dacumos, 2016) who postulated that comprehending a text means engaging in a complex set of processes that involve generating textual meaning by tapping into one's background knowledge or schema and connecting relevant information to make inferences of the text. In any evaluation of healthcare texts, users' comprehension of the material is crucial as it allows the evaluator to check if the existing information is understandable to the reader, and if there is a need to revise some parts of the material to improve its comprehensibility.

The final focal element is *focus on the reader* (communicative effectiveness), which refers to eliciting users' insights about the healthcare information they read. This element involves getting comments directly from the users, asking them about their experience of reading the material, and obtaining suggestions from them on how the material can be further improved. This paper draws on Fage-Butler's (2015) perspective of incorporating user-feedback whereby the viewpoints of healthcare consumers, being the end-users of healthcare materials themselves, are integrated in the evaluation of healthcare texts to help determine and meet their information needs and demands.

2. Method

2.1 Research Design

This paper conducted an evaluation of the readability and comprehensibility of PILs utilizing the mixed method of quantitative and qualitative approaches. The quantitative approach was employed in computing and tabulating for the readability measures, comprehensibility scores, and difficulty indices of PILs. Consequently, the qualitative approach was utilized in analyzing the descriptive feedback or comments of the participants about the PILs.

2.2 Research Corpus

The healthcare texts examined in this study were extracted from the PILs of nonprescription medications. It is important to note here that most of the nonprescription medications used in this study are manufactured in the Philippines; however, there are others that were manufactured outside but have obtained FDA approval to be distributed and sold in Philippine pharmaceutical market. Regardless of the place of manufacture, the PILs for both are mandated to comply with FDA's prescribed leaflet information requirements. Furthermore, PILs were chosen to be analyzed in this paper because these are medical documents that are most accessible to buyers since they come in the packaging of pharmaceutical medications that consumers can easily purchase with no necessary consultation and prescription from an attending physician (Reguero, 2017). Furthermore, since PILs come in medications bought in bundles/packs, an informal online survey was conducted among adult Filipino consumers to determine which nonprescription medicines do they usually keep in their household. The survey determined *fever reducers*, *cough and cold remedy*, and *anti-inflammatory drugs* as the top three household nonprescription medical necessities of the Filipino respondents. These nonprescription medications became the primary sources from which to get PIL samples.

Moreover, since there exists plenty of medicines for each of the top three common household nonprescription medications, this study saw it fit to limit the selection of medicines. Hence, in order to ensure that the chosen medicines fit the aim and scope of the paper, the researchers utilized the purposive sampling technique, i.e., the nonprescription medicines from which the PIL samples were to be taken underwent a stringent selection process informed by the following criteria: (1) the nonprescription drug must be a legally recognized OTC medical product by the FDA; (2) it must be currently distributed and sold in Philippine pharmaceutical stores; and (3) finally, it must belong to any of the top three nonprescription medications in the poll.

The PIL samples were obtained from a community health clinic, a pharmacy boutique, and from various Filipino households, which kept a supply of any of the three nonprescription medications. Over 50 PILs were initially obtained by the researchers; however, upon a closer inspection of the collected PILs, only 42 samples satisfied the paper's selection criteria. Table 1 shows the composition of the corpus specifying the number of PILs collected for each medication category, and the total number of words of the collected leaflets per category.

Table 1
Composition of the corpus

Medication Category	Number of PILs	Total Number of Words
Fever Reducers	20	19,942
Cough and Cold Remedy	12	7,729
Anti-inflammatory Drugs	10	15,012
Total	42	42,683

2.3 Research Participants

The paper utilized a purposive sampling technique in which it suitably chose adult parents (mother or father) as research participants for the user-testing. The choice of research participants is informed by previous consumer-related studies in which parents – mothers or fathers – have been observed to be typically holding financial control and making purchasing decisions for the household (Castillo, 2018; Dacumos, 2016). Furthermore, based on the user-testing guidelines of UK's Medicines and Healthcare products Regulatory Agency or the MHRA, and also based on the suggestion of Sless (2018), the standard number of participants in performing user-testing of healthcare materials must be at least 20. This study was able to double that number and have 41 adult participants for the user-testing who were selected based on the following specific qualifications: (1) he/she must be 20 years old and above, (2) a Filipino resident, (3) at least a high school graduate, and (4) able to understand English (as the PIL samples that were tested are drafted in English). Moreover, the 41 participants who were found eligible to take part in the study's user-testing were comprised of 26 mothers and 15 fathers. Twenty-four were college graduates; 12 were high school finishers, while the remaining five have either a second bachelor's degree or a graduate degree. A majority of the participants (n=20) were middle-aged adults belonging to the age bracket of 40 to 60. About 13 participants were in their 20s, while only eight were in their 30s. The participants come from Manila, Isabela, Batangas, and Samar – locations where the user-testing was likewise held. Additionally, in terms of their linguistic background, all 41 participants mentioned English as their second language. To make sure that the results of the comprehensibility assessment were accurate and valid, those who practice medicine or clinical science were opted out as it is assumed that compared to lay people, they have been exposed to a lot of medical texts and hence, are much more familiar with medical language.

2.4 Research Instrument

2.4.1 Multiple-choice Cloze Test

Following the user-testing methodologies of Dacumos (2016), Lintao (2015), and Surulivelrajan et al. (2013), this paper utilized a cloze-test questionnaire in multiple-choice format to assess how adult participants comprehend PIL textual information. Wilson L. Taylor (1953, as cited in Ajideh & Mozzafarzadeh, 2012) pioneered the use of cloze procedure as an empirical measure to evaluate reading comprehension. Cloze procedure is done by omitting words in a passage and requiring readers to restore the deleted words. The cloze test used in this study employed the rational deletion technique where omission is not fixed, and the gaps are rather determined by the examiner's preference. The researchers deemed the rational deletion technique to be more apt for the cloze procedure of PIL texts as PILs are documents that contain technical words/specific medical terms, which, if removed haphazardly from the passage, may confound or perplex the examinee (Kleijn et al., 2019). Moreover, the use of the multiple-choice test type for the cloze procedure is based on the claims of previous

assessment studies (Alderson, 1990; Haynie III, 1994; Mujeeb et al., 2010) that have found it to be not just reliable but also feasible and convenient in easing the examinee's memory load, lessening the time it would take to process information, and abating test anxiety. In this paper, the participants were tested for their comprehension of the PILs of the three nonprescription medications. Hence, the test questionnaire has three sets of PILs – one for fever reducers, one for cough and cold remedy, and one for anti-inflammatory drugs. Each set contains about 50 blanks following Taylor's (1953) suggestion.

2.4.2 Passage Selection

The passages used for the cloze test are actual textual extracts lifted from existing Philippine PILs. While efforts were exerted to make sure substantial texts from the leaflets would be covered for the cloze test, not everything was included. For instance, numerical information, which often appears under the leaflet's formulation section, or manufacturing or expiration details may not be appropriate items for a cloze test that aims to assess examinee's understandability of the text. Furthermore, in the selection process, although the goal is to include as much text as possible, special focus was given to sections about drug description, dosage instructions, drug administration, contraindications (i.e., what the drug should not be used for), medicine interaction, drug overdosage, and adverse/side effects as these are the parts of the leaflet that have been observed to be mostly read, and that have been found to contain essential risk information about the medical drug (Arsalan, et al., 2015; Garrud et al., 2001; Leemans et al., 2011; Webster et al., 2018).

2.4.3 Test Validation

To make sure that the test questionnaire to be administered is valid and to ensure the quality of the test, the research instrument underwent a validation examination/critique by two invited experts in the study. The first invited validator is a professor of educational assessment who is a Ph.D. candidate specializing in educational leadership in a private research university in Manila. The second is a registered psychometrician who works as a counseling mentor and a testing officer in the guidance department/mental health clinic of a private university. The comments/suggestions for improvement of the test by the validators were heeded and were used to improve the instrument.

2.5 Data Analysis

2.5.1 Readability Test

This paper determined the readability of the PILs of Philippine nonprescription drugs by using Coh-Metrix's computation of reading difficulty based on the tool's standard indices for text readability and easability. Coh-Metrix is a computational tool that examines text in multiple levels of language and discourse centering on cohesion. The current version of the

tool, Coh-Metrix 3.0, provides over a hundred indices grouped into different components that assess various aspects of textual cohesion such as co-reference, lexical diversity, latent semantic analysis, or syntactic complexity among others. Compared to traditional readability metrics that focus on surface-level analysis of sentence length and number of words, Coh-Metrix has been lauded for its ability to take into account the role of cohesion in the cognitive and psycholinguistic processes of reading and hence, yielding a more accurate and reliable readability estimate (Crossley et al., 2011). In this paper, the PIL samples were uploaded by batch in the Coh-Metrix web tool for the readability analysis. The Coh-Metrix readability tool can be accessed online through the following webpage address: http://tool.cohmetrix.com/. It is free and available for any researcher who wishes to use it to analyze any form of texts.

The Coh-Metrix measures specifically utilized in the current study were the indices for text readability and easability as these were the components of the tool that were identified to be closely relevant to reading difficulty (MacNamara et al., 2014). For the text-readability component, three standard reading indices were used, which include the Flesch Reading Ease, Flesch-Kincaid Grade Level, and Coh-Metrix L2 Reading Index. The Flesch Kincaid Grade Level and Flesch Readability Ease use the same core metrics in which they both rely on sentence length (syntactical complexity) and word length (vocabulary difficulty) to assess a text's readability. The former computes how many years of education it would likely require for a reader to understand a text, whereas the latter measures the extent to which a reader would likely find a text difficult to process. Furthermore, Coh-Metrix derived the formulas for Flesch Reading Ease and Flesch-Kincaid Grade Level from their original proponents, Flesch (1948), and Kincaid et al. (1975). The complete equations for these reading measures are shown below.

Flesch Reading Ease
= 206.835 - 1.015 x (words/sentences) - 84.6 x (syllables/words)

Flesch-Kincaid Grade Level
= 0.39 x (words/sentences) + 11.8 x (syllables/words) - 15.59

Meanwhile, the L2 Reading Index is a recent addition in Coh-Metrix. This reading measure was developed to address the lack of readability formula designed exclusively for second-language (L2) readers. Critics of traditional reading formulae observed that existing traditional reading metrics have been employing readability standards based on the reading skills of native (L1) readers, discounting the reading abilities and experiences of their L2 counterparts. Using identified aspects of cohesion related to readability, and accounting for cognitive and psychological models of L2 reading, Crossley et al. (2011) formulated the Coh-Metrix L2 Reading Index to account for L2 readability of texts. The Coh-Metrix L2 Reading Index is derived from three linguistic variables: **the CELEX word frequency** (the frequency of using common words, which concerns decoding), **syntax similarity** (the uniformity and consistency of sentence constructions, which concerns parsing), and **content word overlap**

(the connection of content words and ideas between sentences, which concerns text cohesion and meaning construction). The Coh-Metrix L2 Reading Index formula uses the following computation:

Coh-Metrix L2 Reading Index
= -45.032 + (22.205 x CELEX Frequency Value) + (61.306 x Sentence Syntax Similarity Value) + (52.230 x Content Word Overlap Value)

It is worth noting that the formula for the L2 Reading Index was based on Bormuth's (1971) corpus of academic texts utilized by Greenfield (1999) in formulating the Miyazaki EFL (English as Foreign Language) readability index. Based on Greenfield's (1991) collected cloze performance of 200 Japanese university students, it was found that compared with other reading formulae, the Coh Metrix L2 Reading Index obtained the highest correlation (0.93) with the Japanese students' cloze scores. This was a notable finding because as opposed to other reading formulae, which scale reading difficulty at only the word and sentence level, the Coh-Metrix L2 Reading Index proves that it can go beyond surface-level assessment, and can effectively address issues of text difficulty by assessing textual cohesion, particularly in L2 contexts (MacNamara et al., 2014).

While the indices for the text-readability component are useful in providing information about text difficulty, they only show half of the picture. MacNamara et al. (2014) observe that the readability component provides unidimensional measures, which may only account for *how* more or less difficult a text is, and not *why* it is difficult. This drove them to formulate another related component, which provides a multidimensional readability assessment of text. They referred to it as text easability, which consists of five indices designed to evaluate the specific types of words and/or sentences used in the text in order to identify potential areas in the material where the reading difficulty could be coming from. The five indices are (1) *narrativity*, which refers to how story-like and conversational the text is; (2) *syntactic simplicity*, the degree by which the text employs simple constructions and sentences with fewer words; (3) *word concreteness*, the use of concrete and meaningful words that trigger mental images for easy processing of text; (4) *referential cohesion*, the overlapping of concepts in the text that show explicit connection of ideas between sentences or paragraphs; and (5) *deep cohesion*, the use of connectives that make the text easy to follow and give a sense of cohesiveness of the material. The rule of thumb for these indices is that the higher the percentile score, the easier it is to read the text. These indices were likewise measured in this paper to determine the ease of reading PILs, and subsequently, identify the potential sources of text difficulty of the material.

When the results of the readability test were finally obtained, the researchers interpreted them against the Coh-Metrix Indices Norms formulated by MacNamara et al. (2014). The norms were based on a large corpus comprising over 120,000 paragraphs sourced from around 40,000 different texts from the Touchstone Applied Science Associates (TASA) considered to have the largest corpora of K-12 texts currently available for research. MacNamara et al. (2014) randomly selected 100 passages from the TASA corpora for

three general disciplines: language arts, social studies, and science texts. Furthermore, in assigning normative values of indices for different grade levels, MacNamara et al. (2014) adapted Questar Assessment Inc's readability measure of Degrees of Reading Power (DRP), which has been found to have high correlation with the standard reading measures of Flesch Reading Ease and Flesch Kincaid Grade Level used in the Coh-Metrix tool. As a readability metric, DRP examines text characteristics by computing for word length, sentence length, and word familiarity based on the Bormuth mean cloze readability formula. It assesses text complexity on a single scale ranging from 1-100 units, with the lower units being associated with easy reading, and higher units being related to complex reading. MacNamara et al. (2014) converted these DRP scales to their equivalent grade-level projection and broke them down into different grade bands. Ultimately, this allowed them to determine norms for each Coh-Metrix index in order to set the standard readability scores deemed appropriate and acceptable for certain grade bands. A detailed rundown of Coh-Metrix's 108 indices and their corresponding norms across different grade bands is presented in their book, *Automated Evaluation of Text and Discourse with Coh-Metrix*.

This paper used the norms specifically set for 6-8 grade band as this is considered to be the prescribed grade level for L2 readers of healthcare materials (Coco et al., 2017; Kher et al., 2017). The 6-8 grade band normative values or the standard percentile norms served as the minimum standards by which the text readability and easability scores of the existing PILs were measured against. Table 2 shows the 6-8 grade band normative values for the indices of the text-readability component of Coh Metrix.

Table 2
6-8 grade band normative values of the text readability indices of Coh-Metrix

Coh-Metrix Text Readability Indices	6-8 Grade Band Normative Values
Flesch Reading Ease	70.209
Flesch Kincaid Grade Level	8.381
L2 Reading Index	15.467

Meanwhile, the indices for the text-easability component, which consists of narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion, were assigned the following standard percentile norms under the 6-8 grade band as shown in Table 3.

Table 3
6-8 grade band standard percentile norms of the text easability indices of Coh-Metrix

Coh-Metrix Text Easability Indices	6-8 Grade Band Standard Percentile Norms
Narrativity	64.119 %
Syntactic Simplicity	38.152 %
Word Concreteness	74.252 %
Referential Cohesion	38.894 %
Deep Cohesion	54.417 %

2.5.2 Comprehensibility Assessment (User-testing)

A user-testing of the PIL samples was conducted among target users – adult healthcare consumers – to assess the existing PILs' comprehensibility. According to the MHRA, user-testing is the process of testing the actual text to its target users in order to identify the barriers that hamper users' ability to understand and use information in the material. Before the actual test, two pilot-runs of user-testing were conducted among ten faculty members and nonteaching staff (five for each pilot run) in the department of the university where the principal researcher teaches. The goal of the pilot-testing was to let the researchers fix possible weaknesses of the instrument and as much as possible, prevent some unfavorable scenarios or mishaps that could potentially happen during the actual conduct of the test. For the first pilot-run, a few of the participants' complaints include being confused with the instructions, and having a hard time navigating the questionnaire because the texts were originally printed back to back. The researchers heeded and fixed these problems by printing the texts in separate pages and providing a clear example on the questionnaire on how to answer the test. The second pilot-run tested the revised and improved questionnaire, and the testing went smoothly. Furthermore, after the pilot-testing, the researchers conducted a distractor analysis to improve the distractors used in the multiple-choice test.

When the questionnaire has been revised and improved, the researchers then looked for eligible participants and generated a list of those who were willing to partake in the research project. When the list was finalized and the participants have confirmed, the researchers then set the schedule and venue for the testing. The time and place depended on the availability of both the participants and the researchers. Depending on the preference of the participants, testing was either done in groups or individually, and was held in places favorable and convenient to them such as coffee shops, or their own household with their permission. When all was set, user-testing then proceeded. The researchers conducted the user-testing with the help of two research assistants who are graduate students. During the actual conduct of the user-testing, the participants were provided a brief orientation about their rights and duties in the study. Once everything was clear to them, they were asked to sign the informed consent form. They were provided a copy after signing. The next step was to give instructions on answering the questionnaire. The questionnaire has two parts – the

first part elicited baseline information from the respondents such as their age, educational attainment, linguistic background, and reading frequency/familiarity of PILs. The second part is the multiple-choice cloze test where the participants were made to read passages with missing words and have to circle the letter of the correct word that should be placed on the blank. A time limit of 40 minutes was given for the participants to finish the test; this was based on the suggestion of the registered psychometrician who validated the test. A time extension of 5 minutes, however, was given to those who could not finish the test within the time limit. Once they were done answering, the questionnaires were then collected for checking and scoring.

Since the test is a multiple-choice cloze type, it followed a criterion-referenced scoring. In interpreting the scores of the participants, this paper adopted the criterion-referenced scoring scheme used in the study of Miller et al. (2009) for a multiple-choice cloze procedure they conducted to assess the comprehensibility of pharmacy education materials for patients. Based on the scoring scheme of Miller et al. (2009), a cloze score of 61% and above is categorized under the 'independent' level. On the other hand, a cloze score range of 41% to 60% is interpreted as belonging to the 'instructional' level. Meanwhile, a cloze score of 40% and below is equivalent to the 'frustration' level. The scoring system with its corresponding interpretation is shown in Table 4.

Table 4
Cloze test scoring scheme

Correct Number of Responses in Percentage (%)	Level of Comprehension
≥60 to ≤100	Independent (*Leaflet is understood.*)
≥40 to <60	Instructional (*Leaflet requires supplementary teaching to be understood.*)
>0 to <40	Frustration (*Leaflet is not understood.*)

After getting the scores and determining the participants' level of comprehension, an item analysis was then conducted to get the difficulty index of the leaflets and determine which areas in the text proved to be easy, average, and difficult for participants to understand. The difficulty index is the percentage computation of the number of test takers who gave correct responses to an item. In general, the higher the percentage-value, the easier the test; whereas a lower percentage-value equates to a more difficult test. In describing the difficulty level of the leaflets, this study adopted the interpretation scheme for difficulty index used in the study of Dacumos (2016) whereby a value ranging between 0.76 and 0.99 represents low difficulty; a value ranging between 0.51 and 0.75, moderate difficulty; and a value ranging between 0.10 and 0.50, high difficulty.

2.6 Research Ethics

The user-testing protocols of this study, which involved human participants, were evaluated and approved by the Ethics Review Committee (ERC)[1] of a private research university in Manila, the Philippines. Moreover, as regards identifying the medicines used in this paper, only their generic identification was used. The specific brand names, manufacturers, pharmaceutical companies, or market-distributors were concealed through masking. Similarly, information about the participants who provided data for this study and the experts who were consulted for some areas in this paper remained confidential. To ensure that their identities stayed private, they were only referred to by their professional affiliation and/or occupation in this paper.

3. Results and Discussion

3.1 Readability of the PILs

As with previous studies that utilized Coh-Metrix in assessing the level of readability of healthcare materials (Govander et al., 2016; Liu et al., 2009; Peter et al., 2018), this paper determined the text difficulty of the PILs of Philippine nonprescription drugs by using the tool's standard text readability measures, and text easability indices. Given that the target audience of these PILs comprises L2 readers, Coh-Metrix operating on cognitive and psycholinguistic frameworks of L2 reading proves to be particularly useful in analyzing materials that take L2 readability into account. In this paper, the reading measures under the text readability component of Coh-Metrix were computed. Each PIL sample was analyzed based on the tool's three standard reading measures. The mean scores were obtained for each reading measure and were used to determine the reading difficulty of the PILs of each nonprescription medicine.

Table 5
Text readability measures of the PILs

Medication Category	Flesch Reading Ease	Flesch Kincaid Grade Level	L2 Reading Index
Fever Reducers	35.050	11.607	9.072
Cough and Cold Remedy	42.486	10.167	5.014
Anti-inflammatory Drugs	37.349	11.245	8.600

The results show that in terms of the Flesch Reading Ease, the three medications each obtained a mean score between the range of 30 and 50, with fever reducers having the

[1] ERC Protocol Number: GS2019 PN180

lowest score. These scores fell short of the normative value of 70.209 set by McNamara et al. (2014) for this reading index. Moreover, based on the Flesch Reading Ease difficulty scale, scores falling in the range between 30 and 50 indicate that the PILs are difficult to read and require almost the same skilled level of comprehension demanded in reading academic materials. As to the Flesch-Kincaid Grade Level score of the samples, it was revealed that the PILs of the three medicines are best understood by readers with at least ten to 11 years of schooling. This finding exceeded the sixth- to eighth-grade reading level suggestion of Coco et al. (2017) and Kher et al. (2017) for L2 readers of healthcare materials. Finally, with respect to the Coh-Metrix L2 Reading Index, the PILs of the three medicines fall below the normative value of 15.467, which is the suggested measure by which texts are readable to an average L2 adult (McNamara et al., 2014). Under the Coh-Metrix L2 Reading Index, higher values indicate more understandable texts; lower scores therefore mean difficult-to-understand texts.

To augment and complement the measures of the text-readability component, the PILs were also analyzed based on the five principal indices of the Coh-Metrix text easability. MacNamara et al. (2014) argue that the principal indices of the tool's text easability component, which cover narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion, add more details to the analysis and consider more layers of the text as opposed to the unidimensional measures scaled by the text-readability component.

Moreover, where the text readability metrics concern the degree of difficulty of reading the material, the text easability indices determine the potential areas in the material that account for its difficulty. This paper determined the mean percentile scores of the PILs of the three medicines for each index of the Coh-Metrix text easability component (see Table 6).

Table 6
Text easability indices of the PILs

Medication Category	Narrativity	Syntactic Simplicity	Word Concreteness	Referential Cohesion	Deep Cohesion
Fever Reducers	10.659 %	92.217 %	58.120 %	15.556 %	85.444 %
Cough and Cold Remedy	7.998 %	90.550 %	58.375 %	11.910 %	66.619 %
Anti-inflammatory Drugs	7.633 %	92.943 %	65.161 %	12.816 %	87.945 %

Findings reveal that in terms of text easability, the PILs of the three medicines obtained low percentile scores in three indices, namely narrativity, word concreteness, and referential cohesion. The narrativity percentile scores of the PILs range between 7% and 10%, far below the 64.119% standard percentile norm set for this index. Narrativity is concerned with how well a text presents a story, and therefore how spoken-like its orientation is. The existing PILs' low narrativity indicates that the leaflets appear to contain a relatively high

proportion of unfamiliar words/terms, which make the materials more formal and technical rather than being story-like and conversational. The following excerpts illustrate the lack of narrativity in the texts of existing PILs:

(1) Paracetamol (acetaminophen) has been reported to potentiate the effect of orally administered anticoagulants. (OTC #12)

(2) There is an increased risk of nephrotoxicity when Ciclosporin or Tacrolimus are given together with Ibuprofen. (OTC #26)

Excerpts 1 and 2 were lifted from the sections of the PILs, which discuss the harmful reactions or risks that come when a medical drug is mixed with other medicines. In excerpt 1, the term *potentiate* means to increase the action of something; meanwhile, the phrase, *orally administered* means to take the drug by mouth. *Anticoagulants* refer to medications used for the reduction of blood-clotting. Meanwhile, in excerpt 2, the term *nephrotoxicity* refers to toxicity induced by the kidneys. *Ciclosporin* and *Tacrolimus* are immunosuppressant drugs (i.e., drugs that reduce the strength of the immune system). For nonspecialists, reading texts with low narrativity may prove to be challenging as it requires a certain level of education and sufficient knowledge of the field to understand them. Consequently, the frequent use of specialized terms in the existing PILs may not bode well for healthcare consumers as it could compromise their knowledge about the safety use of a medical drug.

As for word concreteness, the existing PILs registered percentile scores between 58% and 65% as against the index's standard percentile norm of 74.252%. The existing PILs' low word concreteness suggests that the leaflets are riddled with a lot of abstract words that are hard to visualize and may require more processing time on the part of the reader. The recurrent use of abstract words in the existing PILs are shown in excerpts 3 and 4, which were lifted from the parts of the leaflet that convey information about the medical symptoms or conditions the drug treats, and also the side effects that come in taking the said drug.

(3) Paracetamol syrup is useful for reducing fever and for the relief of mild to moderate, aches, pains and discomfort associated with the common colds and flu, inoculation or immunizations. It is also valuable in reducing pain following tonsillectomy and adenoidectomy. (OTC #14)

(4) Phenylephrine HCl may cause sudden, persistent, severe headache, nervousness, restlessness, insomnia/sleeplessness, dizziness, anxiety, confusion, high blood pressure, palpitation, chest tightness, tremor, agitation, irritability, aggressiveness (particularly in young children), nausea, and blurred vision (OTC #11)

In excerpt 3, the words *fever*, *aches*, *pains*, *discomfort*, *common colds*, and *flu* refer to a state of affairs pertaining to clinical conditions, which the medical drug aims to relieve or treat. Additionally, the same excerpt also contains the terms *inoculation or immunization* (vaccination), *tonsillectomy* (surgical removal of tonsils), and *adenoidectomy* (surgical removal of adenoid glands), which are abstract words referring to clinical processes or surgical procedures. Meanwhile, the abstract terms in excerpt 4 are used to inform healthcare consumers about the list of potential ailments that may come as a consequence of taking the medical drug. The frequent occurrence of abstract words in the existing PILs is actually expected as medication texts commonly contain statements that mostly revolve around state of affairs and processes such as drug administration and precautions, clinical conditions/ailments, side effects, and medical procedures/treatments (Herber et al., 2014). However, documents possessing low word concreteness and containing too many abstract expressions may demand more cognitive effort from readers. In the case of the existing PILs, this could pose potential risks as it could hamper healthcare consumers' ability to respond to certain hazards or address urgent situations should these occasions arise.

For referential cohesion, the existing PILs did not meet the standard percentile norm of 38.894% having accomplished low percentile scores falling between 11% and 15%. Referential cohesion is the characteristic of a text that shows overlapping of concepts from one sentence to another. It is observed when sentences consistently refer to the same words or ideas, forming explicit conceptual connections throughout the text for the reader. Hence, when a text is low in referential cohesion such as in the case of the existing PILs, it suggests lack of consistent and specific conceptual references that make it hard for readers to identify related ideas in the text. The lack of referential cohesion in the existing PILs is shown by the following excerpts:

(5) Additive central nervous system (CNS) depression may occur when dextromethorphan is taken together with alcohol, antihistamines, psychotropics, and <u>other CNS depressants</u>. (OTC #19)

(6) Ask a doctor or pharmacist before use if you are taking <u>other NSAIDs</u>, anticoagulants or <u>any other drugs</u> or if you are taking aspirin for heart attack or stroke because Ibuprofen may decrease the effect of aspirin. (OTC #25)

In excerpt 5, *other CNS depressants* refer to other medicines that affect brain activity such as sedatives, tranquilizers, and hypnotics. Additionally, in excerpt 6, *other NSAIDs* points to other anti-inflammatory drugs, whereas the term *any other drugs* pertains to other medications. The use of these expressions could result in linguistic indeterminacy whereby texts in documents are characterized by vagueness of meanings and uncertainty of references (Coulthard & Johnson, 2010). In the given excerpts, readers may find themselves wondering what is particularly being referred to by *other CNS depressants*, *other NSAIDs,* and *any other drugs* since no specific references were mentioned for these expressions. This

leaves cohesion gaps that require readers to make unnecessary inferences, which can be challenging especially to those who have no sufficient background of the concepts discussed in the text. Dacumos (2016) note that safety texts with vague conceptual references could place consumers in a delicate situation where they are compelled to rely on their guesses or speculations to make sense of crucial safety information.

Interestingly, with respect to syntactic simplicity and deep cohesion, the PILs were found to have accomplished percentile scores that met the standard values for these indices. For syntactic simplicity, the PILs exceeded the percentile norm of 38.152 %, attaining high percentile scores ranging between 90% and 92%. Syntactic simplicity refers to the use of simple structures or sentences with fewer words/clauses in the text. In the following excerpt, it can be observed that the text contains no embedded clauses, which usually characterize long sentences in technical documents. Instead, in this text, what could have been a long sentence is split into three short imperative clauses for easy reading of the safety measures in taking or administering the medications. The use of the same sentence structure or function throughout the text also adds to its syntactic simplicity (see excerpt 7).

(7) These medicines should be given with care in patients with a history of stomach or intestinal ulcer and gastrointestinal bleeding. Do not take more than the recommended dose. Do not use after the expiry date on the label. (OTC #21)

One probable explanation, which accounts for the use of simple sentence constructions in the PILs, is that the information contents of these documents are bound by strict regulations (Frade, 2015). By regulatory practice, PILs are supposed to be short as they are meant for a layperson's reference; drafters are therefore obliged to produce these leaflets with texts that are less verbose, less lengthy, and less structurally complex in an effort not to intimidate and much so, discourage readers from browsing the material. It has to be clarified here though that while at the surface, shorter sentences appear to contribute to the ease of reading the text, they do not necessarily equate to making the text readable or comprehensible (Janan & Wray, 2012).

Finally, as regards the index for deep cohesion, the existing PILs reached the standard percentile norm of 54.417%, having obtained satisfactory percentile scores between 66% and 87%. Deep cohesion refers to the use of connectives that help readers associate concepts or recognize relationships of ideas in the text. This finding implies that the existing PILs have a high frequency of using connective devices or cohesive markers in the leaflets, which potentially help readers in making sense of the discourse organization of the texts. Deep cohesion in the existing PILs is illustrated in excerpt 8.

(8) Before taking this medication, tell your doctor if you have high blood pressure or any other type of heart problems, glaucoma, thyroid problems, diabetes, liver or kidney disease, an enlarged prostate, bladder problems, or difficulty urinating. Stop use and ask a doctor if fever gets worse or lasts more than 3 days, or if new symptoms occur. (OTC #16)

Excerpt 8 shows the use of a temporal cohesive marker (*before*) and additives (*and, or*) that help tie the concepts or information together in an attempt to convey the instruction easily and clearly to readers. The high deep cohesion value of the existing PILs may be attributed to the fact that these healthcare materials contain crucial contents pertaining to proper drug administration, side effects, stages or phases of overdosage, or safe handling of the drug, which require the use of connectives in an attempt to convey the safety information/ instructions clearly and accordingly to the readers. The use of connectives in the PILs make it easy for the readers to follow the order of information in the leaflets.

Overall, the readability analysis of the existing PILs of Philippine nonprescription drugs shows a striking revelation – that these existing healthcare leaflets being distributed to lay consumers do not seem to be written at a reading level that is suitable for the general public. It can be argued that the PILs' low readability could have also accounted for the reported occasional reading of the leaflets by the study's participants. Healthcare texts, which have high reading difficulty, have been shown to impede consumers' access to healthcare information, and consequently, lead to lower health literacy (Badarudeen & Sabharwal, 2010; Kher et al., 2017; Sand-Jecklin, 2007).

In addition, von Beusekom et al. (2016) found that healthcare materials with low readability tend to be discouraging for medical drug consumers, making it less likely for them to read and use them. In effect, this leads consumers to become more vulnerable to looking for and relying on alternative information sources that have questionable legitimacy.

3.2 Comprehensibility of the PILs

To address the need of integrating healthcare consumers in the evaluation of PILs, and also to validate the findings of the readability test done in this paper, this study conducted a comprehensibility assessment of existing PILs. Using actual healthcare materials, the researchers tested these existing PILs to the adult participants. The outcome of the user-testing is shown in the following figure:

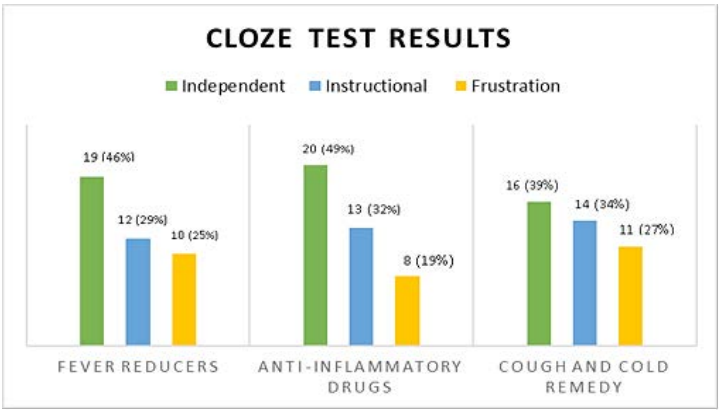

Figure 1. **Results of the PIL user-testing via cloze procedure**

Findings reveal that for each medication category, a majority of the participants appear to belong to the independent level. Similarly, in all three medications, those under instructional level ranked second in number, while the ones categorized under the frustration level came last. However, it has to be clarified here that while the independent level may seem to have the highest figure in all three medication categories, a careful observation of the results reveals that **a majority represents only less than half of the total number of participants (N=41)**. Generally, this means that only a few actually demonstrated an independent level of understanding of the leaflets – only 46% for fever reducers, 49% for anti-inflammatory drugs, and 39% for cough and cold remedy. Interestingly, if those under instructional and frustration levels are taken together, they constitute more than half of the participants that either need somebody else's assistance to understand the leaflet, or that completely fail to understand the healthcare material. Under the standards of user-testing set by David Sless (2018), an Australian document design specialist whose user-testing guidelines have been used and adopted by several international healthcare agencies, an effective PIL should be able to accomplish the success criteria of around 80%-90% of literate adults being able to demonstrate independent comprehension of the material. The findings of the present investigation reveal that with only less than 50% of the participants being able to understand the leaflets for all three medications, the existing PILs of Philippine nonprescription drugs appear to fall short of Sless's (2018) success criteria for an effective leaflet, and may therefore need to be reviewed and revised.

Furthermore, heeding the need for integrating patients' perspective in the evaluation process of healthcare materials (Fage-Butler, 2015; Garner et al., 2012), the researchers sought to get participants' feedback about the test by asking them to give a comment about the information they read from the actual PILs. Most of them pointed the excessive use of unfamiliar medical terminologies that make the information in the PILs difficult to understand. A few of these notable comments from the participants can be read below.

Participant #6
"Hindi naging madali ang pagbasa at pag-unawa ko sa mga nakasulat sa information leaflet. Malalim ang salita at 'di ako familiar."
(It was not easy for me to read and understand the content written on the information leaflet. The words are high-falutin which I am not familiar with.)

Participant #27
"The patient information leaflets are quite complicated especially when you have no medical background since the terms used here are for professionals only and not for regular people. It would've been easy to understand if they could deliver it in a way that regular people can understand in layman's terms."

Some participants reported that the leaflets' incomprehensibility leads them to make guesses and simply rely on previous experiences or common knowledge to make sense of the information. This could pose risk on the part of the healthcare consumer as poorly construed information could enable self-medicating behavior (Herber et al., 2014).

Participant #1
"I found it hard to understand due to some medical terminologies I'm not familiar with. I just made a good guess or used my common sense, and depended on my previous experiences I heard from my wife who happens to be a nurse."

Participant #17
Mahirap sagutan. Nagbase lang ako sa naging karanasan ko sa pagbibigay ng gamot sa aking asawa at anak sa tuwing nagkakasakit sila. Dapat na basahin din ang nakalagay sa kahon ng gamot.
(It was difficult to answer. I only relied on my experience of giving medicines to my spouse and children every time they get sick. It is important to read the information on the medical drug's packaging.)

Interestingly, there are participants who are aware of the consequences of reading incomprehensible information from the leaflet, and are worried about the potential harms difficult information can do to a clueless healthcare consumer. These are some of the concerns they raised:

Participant #5
Nahirapan po akong intindihin ang karamihan sa mga salita doon sa pinasasagutang mga tanong. Maaaring magkaroon ng maling pagkaintindi ang makakabasa at humantong sa paglala ng sakit or pagkamatay.
(I found it hard to understand most of the words in the questions I was asked to answer. There could be a chance that readers might have a wrong understanding of the information, which could worsen one's condition that might even lead to death.)

Participant #26
'Yong ibang terms, hindi ako nakakaintindi. Paano na lang 'yong ibang mga consumers na bumibili at kukuha ng leaflets?! Hinde sila magkakaroon ng panahon na basahin kung para sa kanila mahirap intindihin ang mga terms na ginagamit which are suited only for the doctors or those studying medical courses.
(I did not understand some of the terms. I can just imagine the consumers who buy medicines and pick up these leaflets. They will not have the time to read if the terms used are suited only for doctors or those studying medical courses.)

3.3 Difficulty Index of the Information Texts of PILs

To further validate participants' claim about the low comprehensibility of Philippine PILs, an item analysis was conducted to identify the difficulty index of the leaflets, and to determine which parts of the text were most challenging for users to comprehend. This paper found that most of the information items in the PILs of the three medications were difficult. A detailed summary of the difficulty index for each medication category is shown in Table 7.

Table 7
Difficulty index of the PILs of Philippine nonprescription drugs

Medication Category	Item Difficulty Level	F	%	Mean Difficulty Index
Fever Reducers	High	20	40%	0.37
	Moderate	16	32%	0.59
	Low	14	28%	0.80
Anti-inflammatory Drugs	High	18	36%	0.39
	Moderate	17	34%	0.61
	Low	15	30%	0.80
Cough and Cold Remedy	High	23	46%	0.35
	Moderate	19	38%	0.63
	Low	8	16%	0.79

The leaflet for cough and cold remedy was the most difficult to understand garnering a mean difficulty index of 0.35. This was followed by the leaflet for fever reducers with 0.37. Meanwhile, the least difficult leaflet was that of the anti-inflammatory drugs obtaining a mean difficulty index of 0.39. Furthermore, it has been observed that the items which got the least number of correct responses were mostly those that require an understanding of a medical term. Below is a sample item from a fever reducer leaflet that obtained the least number of correct responses.

WHAT IS THE MEDICINE USED FOR?

3. Paracetamol is used for the treatment of minor____(6)_____ conditions, and ____(7) relief____ of headaches and pains.

(6) A. allergic B. convulsive C. inflammatory (D.) febrile

The answer to item 6 is *D. febrile*, which refers to fever. Out of the 41 participants, only 12 got it correct. Most of those who answered it incorrectly chose *inflammatory*, which is not the main clinical use/function of paracetamol. It can be reckoned that perhaps, the participants know about paracetamol being generally used for fever treatment; however, the term used in the leaflet has probably confused them and led them to simply make guesses or choose a term they are more familiar with. It is suggested that the word *fever*, which is more likely to be recognized by laypeople rather than the word *febrile* should be used to make the text more understandable.

Meanwhile, only a few participants correctly identified the clinical word being referred to in the following sample item taken from the leaflet of an anti-inflammatory drug:

WHAT OTHER MEDICINE OR FOOD SHOULD BE AVOIDED WHILE TAKING THIS MEDICINE?

14. Ibuprofen may cause acute _(40)_____ in kidney function, and blood pressure in response to ACE inhibitors. Medicines which _(41) stimulate_ the enzymes responsible for metabolic activation of Paracetamol such as medicines for convulsion may increase (42) susceptibility_ to the harmful effects to the liver.

(40) (A.) reduction B. filtration C. retention D. fusion

The correct answer to this item is A. reduction; in medical contexts, it means the decrease of the clinical function of something. This item garnered only 14 correct responses. It is worth noting that the term *reduction* is nominalized in the text. While nominalization is a normal practice in medical writing, nominalized terms add to lay readers' cognitive load,

requiring them more time to process and understand the text (Maglie, 2015). This could probably explain why the item has a high level of difficulty; the use of nominalization in the text may have demanded more information-processing time, and may have also caused abstraction on the part of the participants. As much as possible, healthcare materials produced for lay patients must aim at lessening nominalized terms.

Moreover, in the following sample item lifted from a cough and cold medicine leaflet, a majority of the participants had trouble identifying correctly the part of the eye that dilates or widens because of overdosage:

SIGNS AND SYMPTOMS OF OVERDOSAGE

11. Mild (31) cases of diphenhydramine overdosage are mainly characterized by dry mouth, headache, nausea, tachycardia or (32) fast heart rate and urinary retention. In children, the clinical features include hallucinations, dilated ___(33)___, fever ataxia, and convulsions. Cardiorespiratory depression and coma may subsequently _(34) develop_, with death occurring between 2 to 18 hours.

(33) A. cornea B. iris (C.) pupil D. optic nerve

In this item, only 11 participants were able to identify the correct eye part that will be affected. It can be reckoned that participants who got it wrong may have difficulty understanding the word *dilate*; hence, they ended up giving the incorrect response. It can also be the case that some participants may not be familiar with the function of *pupil* as an anatomical part of the eye, raising concerns about health literacy. Smith et al. (2013) note that if healthcare texts are already difficult to understand for a person with adequate health literacy, comprehensibility will be much worse for patients with limited health literacy.

4. Conclusion

This study evaluated the textual quality of the PILs of Philippine nonprescription drugs based on readability measures and comprehensibility assessment. The present investigation revealed that based on a readability analysis, the existing PILs were found to be more academic-oriented rather than spoken-like, and hence require high-skilled reading. A further validation of this finding was made through a user-testing of the leaflets, which found that the existing PILs appear to have low comprehensibility as it was revealed that consumers reported experiencing difficulty in reading and understanding healthcare information contained in the leaflets. Moreover, an analysis of the difficulty index of the existing PILs revealed that most information items in the leaflets contain high difficulty, which raises concerns over the healthcare materials' suitability for lay readers.

The lack of local studies investigating the use of language in healthcare-related texts is an invitation for language scholars to engage in more research exploring the role of language in Philippine medical/healthcare contexts. Those who are interested in studying the language of healthcare or medicine may begin by expanding the current study and examining the areas that were not explored in the paper. For instance, this study dwelt on only PILs as corpus for analysis; further linguistic studies may adopt the methods utilized by this paper to assess texts of other patient education materials such as informed consent forms for clinical treatments/trials, or user manuals of medical devices. Furthermore, since the study has already accomplished the evaluation phase of the existing PILs, those wishing to extend it may use the findings of the present investigation to easify or simplify the leaflets to make them more user-centered. In doing so, one may learn from the document simplification techniques/ protocols of Lintao (2015), or the simplified safety directives/warning models of Dacumos (2016). Additionally, the work of Frade (2015) in easifying PILs and the proposed computer-generated simplification tool for healthcare materials designed by Proulx et al. (2013) may serve as helpful resources in revising, easifiying, and simplifying the PILs. It is also highly suggested that aside from healthcare consumers, further studies on these existing PILs may consider incorporating the perspective of other stakeholders such as drug manufacturers, medical practitioners, and legal experts in the process of designing and producing consumer healthcare materials.

Finally, the low readability and comprehensibility of the existing PILs suggests that the leaflets do not seem to be written at the level of a layperson's understanding, which AO 2016-0008 requires; hence, this finding warrants a revisit of the existing regulatory mandate to check if it is really strictly enforced. Moreover, as it stands, the existing regulatory mandate currently does not require manufacturers to conduct user-testing of these consumer healthcare documents. Philippine healthcare regulators may accommodate the idea of performing a linguistic review and mandatory user-testing of healthcare materials whereby consultations about the development of healthcare texts are made with not just clinical professionals but also linguists, and most importantly, healthcare consumers. In this regard, they can adopt the linguistic review protocols and user-testing guidelines set by the European Medicines Agency (EMA) whose drug manufacturing requirements interestingly include linguistic intervention and consumer consultation of consumer healthcare documents. Additionally, Philippine healthcare officials may consider providing manufacturers a more detailed PIL template that would guide them in drafting a leaflet that is more legally-compliant and more consumer-centered.

References

Ajideh, P., & Mozaffarzadeh, S. (2012). C-test vs. multiple-choice cloze test as tests of reading comprehension in Iranian EFL context: Learners' perspective. *English Language Teaching 5*, 143-150.

Alderson, J.C. (1990). Learner-centered testing through computers: Institutional issues in individual assessment. In J.H.A.L. de Jong, & D.K. Stevenson (Eds.), *Individualizing the assessment of language abilities* (p. 27). Multilingual Matters.

Arsalan, A., Qureshi, S., Alam, M., Ahmed, S., Shakeel, O., Ali, A., & Jabeen, W. (2015). Errors in patients' information leaflets of marketed medicines in Pakistan. *Journal of Applied Pharmaceutical Science*, *5*(5), 68-074. https://dx.doi.org/10.7324/JAPS.2015.50513

Askehave, I., & Zethsen, K. (2003). Communication barriers in public discourse: The patient package insert. *Document Design*, *4*(1), 22-41. https://doi.org/10.1075/dd.4.1.03ask

Badarudeen, S., & Sabharwal, S. (2010). Assessing readability of patient education materials: Current role in orthopaedics. *Clinical Orthopaedics and Related Research*, *468*(10), 2572-2580. https://doi.org/10.1007/s11999-010-1380-y

Bailin, A., & Grafstein, A. (2016). *Readability: Text and context*. Springer.

Bormuth, J.R. (1971). Development of standards of readability: Toward a rational criterion of passage performance. *Final report, U.S. Office of Education, Project No. 9–0237*. University of Chicago.

Castillo Jr, F.G. (2018). Consumer buying behavior: The multicultural influence in the Philippines. *The International Journal of Business Management and Technology*, *2*(2), 71-75.

Clerehan, R., Buchbinder, R., & Moodie, J. (2005). A linguistic framework for assessing the quality of written patient information: Its use in assessing methotrexate information for rheumatoid arthritis. *Health Education Research*, *20*(3), 334-344. https://doi.org/10.1093/her/cyg123

Coco, L., Colina, S., Atcherson, S.R., & Marrone, N. (2017). Readability level of Spanish-language patient-reported outcome measures in audiology and otolaryngology. *American Journal of Audiology*, *26*(3), 309-317. https://doi.org/10.1044/2017_aja-17-0018

Coulthard, M., & Johnson, A. (2010). *The Routledge handbook of forensic linguistics*. Routledge.

Crossley, S.A., Allen, D.B., & McNamara, D.S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, *23*(1), 84-101.

Dacumos, S. (2016). *Linguistic Features, Adequacy, Comprehensibility, and Readability of Philippine Consumer Product Warnings* (Doctoral dissertation). University of Santo Tomas, Manila, the Philippines.

Dickenson, D., Raynor, D., & Dunman, M. (2001). Patient information leaflets for medicines: Using consumer-testing to determine the most effective design. *Patient Education and Counseling*, *43*(2), 147-159. https://doi.org/10.1016/s0738-3991(00)00156-7

Fage-Bulter, A. (2011). The discursive construction of risk and trust in patient information leaflets. *Hermes – Journal of Language and Communication Studies*, *24*(46), 61-73. https://doi.org/10.7146/hjlcb.v24i46.97368

Fage-Butler, A. (2015). Package leaflets for medication in the EU: The possibility of integrating patients' perspectives in a regulated genre? *Medical Writing*, *24*(4), 210-214. https://doi.org/10.1179/2047480615Z.000000000325

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221-233. https://psycnet.apa.org/doi/10.1037/h0057532

Frade, C. (2015). Attempting the easification of patient information leaflets (PILs) in Brazil. In M. Gotti, S. Macci, & M. Sala (Eds.), *Insights into medical communication* (pp. 137-157). Petar Lang AG, International Academic Publishers. https://doi.org/10.3726/978-3-0351-0844-6

Garner, M., Ning, Z., & Francis, J. (2012). A framework for the evaluation of patient information leaflets. *Health Expectations*, *15*(3), 283-294. https://doi.org/10.1111/j.1369-7625.2011.00665.x

Garrud, P., Wood, M., & Stainsby, L. (2001). Impact of risk information in a patient education leaflet. *Patient Education and Counseling*, *43*(3), 301-304. https://doi.org/10.1016/s0738-3991(00)00168-3

Govender, D., Villafranca, A., Hamlin, C., Hiebert, B., Parveen, D., & Jacobsohn, E. (2016). Appropriateness of language used in patient educational materials from 24 national anesthesiology associations. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, *125*(6), 1221-1228. https://doi.org/10.1097/aln.0000000000001361

Greenfield, G. (1999). *Classic Readability Formulas in an EFL Context: Are They Valid for Japanese Speakers?* (Doctoral dissertation). Temple University, Philadelphia, PA, United States. (University Microfilms No. 99–38670).

Gustafsson, J., Kälvemark, S., Nilsson, G., & Nilsson, J.L.G. (2005). Patient information leaflets—patients' comprehension of information about interactions and contraindications. *Pharmacy World and Science*, *27*, 35-40. https://doi.org/10.1007/s11096-005-1413-x

Harwood, A., & Harrison, J. (2004). How readable are orthodontic patient information leaflets*? Journal of Orthodontics*, *31*(3), 210-219. https://doi.org/10.1179/146531204225022425

Haynie III, W.J. (1994). Effects of multiple-choice and short-answer tests on delayed retention learning. *Journal of Technology Education*, *6*(1), 32-44.

Herber, O.R., Gies, V., Schwappach, D., Thürmann, P., & Wilm, S. (2014). Patient information leaflets: Informing or frightening? A focus group study exploring patients' emotional reactions and subsequent behavior towards package leaflets of commonly prescribed medications in family practices. *BMC Family Practice*, *15*, 163. https://doi.org/10.1186/1471-2296-15-163

Hughman, B. (2009) *Healthcare communication*. Pharmaceutical Press.

Janan, D., & Wray, D. (2012) *Readability: The limitations of an approach through formulae.* http://www.leeds.ac.uk/educol/documents/213296.pdf

Kessels, R. (2003). Patients' memory for medical information. *Journal of the Royal Society of Medicine*, *96*(5), 219-222. https://dx.doi.org/10.1258%2Fjrsm.96.5.219

Kher, A., Johnson, S., & Griffith, R. (2017). Readability assessment of online patient education material on congestive heart failure. *Advances in Preventive Medicine*. https://doi.org/10.1155/2017/9780317

Kincaid, J.P., Fishburne, R.P., Rogers, R.L., & Chissom, B.S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*, Research Branch Report 8-75, Millington, TN: Naval Technical Training, U. S. Naval Air Station, Memphis, TN.

Kleijn, S., Pander Maat, H., & Sanders, T. (2019). Cloze testing for comprehension assessment: The HyTeC-cloze. *Language Testing*, *36*(4), 553-572. https://doi.org/10.1177%2F0265532219840382

Leemans, L., Heylen, N., Quanten, A., & Deferme, S. (2011). Consumer study on the use of patient information leaflets. *Journal de Pharmacie de Belgique*, *4*, 109-116.

Lintao, R.B. (2015). *Simplifying a Philippine Consumer-finance Contract: Towards the Development of a User-comprehensible, Readable and Legally Acceptable Document* (Doctoral dissertation). University of Santo Tomas, Manila, the Philippines.

Liu, C.J., Kemper, S., & Bovaird, J.A. (2009). Comprehension of health-related written materials by older adults. *Educational Gerontology*, *35*(7), 653-668. https://doi.org/10.1080/03601270902885504

Maglie, R. (2015). Can you read this leaflet?: User-friendliness of patient information leaflets in the UK and Italy. In M. Gotti, S. Macci, & M. Sala (Eds.), *Insights into medical communication* (pp. 137-157). Petar Lang AG, International Academic Publishers.

McNamara, D.S., Graesser, A.C., McCarthy, P.M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Miller, M.J., DeWitt, J.E., McCleeary, E.M., & O'keefe, K.J. (2009). Application of the cloze procedure to evaluate comprehension and demonstrate rewriting of pharmacy educational materials. *Annals of Pharmacotherapy*, *43*(4), 650-657. https://doi.org/10.1345/aph.1l642

Mujeeb, A., Pardeshi, M., & Ghongane, B. (2010). Comparative assessment of multiple choice questions versus short essay questions in pharmacology examinations. *Indian Journal of Medical Sciences*, *64*(3), 118-124. https://doi.org/10.4103/0019-5359.95934

Peter, S.C., Whelan, J.P., Pfund, R.A., & Meyers, A.W. (2018). A text comprehension approach to questionnaire readability: An example using gambling disorder measures. *Psychological Assessment*, *30*(12), 1567-1580. https://doi.org/10.1037/pas0000610

Pongpunna, S., Pratipanawatr, T., & Jarernsiripornkul, N. (2018). Survey of outpatients' use and needs of patient medicine information leaflets in Thailand. *Journal of Clinical Pharmacy*, *41*, 141-150. https://doi.org/10.1007/s11096-018-0748-z

Proulx, J., Kandula, S., Hill, B., & Zeng-Treitler, Q. (2013). Creating consumer friendly health content: Implementing and testing a readability diagnosis and enhancement tool. In *2013 46th Hawaii International Conference on System Sciences* (pp. 2445-2453). IEEE.

Reguero, T. (2017). *Please Consult Your Pharmacist: A Linguistic Analysis of Patient Information Leaflets* (Master's thesis). Vallodolid University, Faculty of Arts and Philosophy, Spain.

Sand-Jecklin, K. (2007). The impact of medical terminology on readability of patient education materials. *Journal of Community Health Nursing*, *24*(2), 119-129. https://doi.org/10.1080/07370010701316254

Sless, D. (2018). Designing documents for people to use. *She Ji: The Journal of Design, Economics, and Innovation*, *4*(2), 125-142. https://doi.org/10.1016/j.sheji.2018.05.004

Smith, S.G., Curtis, L.M., Wardle, J., von Wagner, C., & Wolf, M.S. (2013). Skill set or mind set? Associations between health literacy, patient activation and health. *PloS one*, *8*(9), e74373. https://doi.org/10.1371/journal.pone.0074373

Surulivelrajan, M., Joy, M., Suhaj, A., Rajesh, V., & Manu, M.K. (2013). Study to assess the utility of cloze test in readability assessment of patient information leaflets in English and Indian languages. *Asian Journal of Pharmaceutical and Health Sciences*, *3*(4), 824-829.

Taylor, W. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism & Mass Communication Quarterly*, *30*(4), 415-433. https://doi.org/10.1177%2F107769905303000401

von Beusekom, M.M., Grootens-Wiegers, P., Bos, M.J., Guchelaar, H.J., & van den Broek, J.M. (2016). Low literacy and written drug information: information-seeking, leaflet evaluation and preferences, and roles for images. *International Journal of Clinical Pharmacy*, *38*(6), 1372-1379. https://doi.org/10.1007/s11096-016-0376-4

Webster, R.K., Weinman, J., & Rubin, G.J. (2018). Positively framed risk information in patient information leaflets reduces side effect reporting: A double-blind randomized controlled trial. *Annals of Behavioral Medicine*, *52*(11), 920-929. https://doi.org/10.1093/abm/kax064

Yamamoto, M. (2014). Patient drug information leaflets for risk/benefit communication. *Journal of Pharmacovigilance*, *3*(1), e132.

Anthony Lemuell S. Cabañas is a licensed professional teacher and language researcher. He earned his Bachelor's Degree in Secondary Education Major in English at Far Eastern University (FEU), Manila where he graduated summa cum laude. He completed his Master of Arts in English Language Studies degree at the University of Santo Tomas Graduate School, Manila. He has served as a lecturer handling English courses in Far Eastern University (FEU) and Chiang Kai Shek College. His language research interests are in the areas of discourse studies, pop culture and language, second language acquisition, medical linguistics, and forensic linguistics.

Prof. Marilu Rañosa-Madrunio, Ph.D. currently serves as the dean of the faculty of arts and letters of the University of Santo Tomas (UST). She was the previous dean of the UST Graduate School. She earned her Doctor of Philosophy in Applied Linguistics degree with academic distinction at De La Salle University (DLSU), Manila. She obtained her Master of Arts in English Language Studies degree at the UST Graduate School where she graduated summa cum laude. She has held leadership positions in several local and international linguistic organizations. She presently serves as a member of the executive board of the International Association of Forensic Linguists (IAFL).